The University of Melbourne
Department of Mathematics and Statistics

Hidden Markov Models with Multiple Observation Processes

Honours Thesis

November 2, 2007

James Zhao

supervised by

Bill Moran Peter Taylor

Contents

1 Introduction

A *Hidden Markov Model* is a statistical model consisting of two stochastic processes—the underlying process which describes the system in question, and the observation process through which observations are made. The key notion is that the underlying process is a Markov chain which cannot be directly observed. Thus, all available information comes through the observation process, which is a probabilistic function of the underlying process.

Heuristically, such models can be thought of as "Markov chains with observation errors", and have commonly been used as a model for languages, with notable applications including speech recognition [18], musical transcription [19], cryptanalysis [11], and DNA and protein sequencing [6].

One question of interest is: "What happens if there are different ways of observing the underlying process?" If we can simultaneously observe in all of these ways and combine the observations, this would clearly be the optimal solution. However, this is not always possible. Even if neither time nor computing power is limited, some systems do not easily lend themselves to multiple simultaneous observations. For example, a radar can choose only one electromagnetic pulse to transmit within a certain time interval [13].

It is this problem of choosing the best method of observation to use which will be the primary focus of our work. In particular, we will consider policies for selecting the observation method based on past observations, with the optimality criterion that one policy is considered superior to another if the best estimate under that policy has a lower long term expected information entropy.

Our aim will be to describe an optimal policy for selecting the method of observation, show that this policy is indeed optimal, and show that in certain situations, it is significantly better than naïve ways of choosing a method of observation, such as only using one method or alternating between two methods.

1.1 History

Markov chains are named after Andrei Andreievich Markov (1856–1922), who first studied them in an attempt to discredit his lifelong adversary, P.A. Nekrasov, who had asserted that independence is a necessary condition for the weak law of large numbers [1]. Markov constructed his *chains* as a counterexample to Nekrasov's claim, publishing his results in a series of papers beginning in 1906 [22]. While his initial motivation was purely theoretical, Markov also provided one of the earliest applications of his model, using it to analyse the relationship between consonants and vowels in Alexander Pushkin's *Eugene Onegin* in 1913 [15].

Work on Markov chains continued throughout the twentieth century, among the most notable results being the generalisation to countably infinite state spaces by A.N. Kolmogorov in 1937. Markov chains are now used in an enormous range of mathematical applications, including statistical mechanics, communications, mathematical biology, life insurance, and the Google search engine.

The *Hidden Markov Model* was first introduced in a series of papers from 1966 to 1972 by L.E. Baum and others, who called them by the more descriptive name of *Probabilistic Functions of Markov Chains* [3]. The initial major application for this theory was in speech recognition in the early 1970s, with the main results in this area described particularly concisely in L.R. Rabiner's influential 1989 tutorial paper [18]. Hidden Markov Models saw another major application in bioinformatics from the 1980s onwards, with extensive usage in the analysis of sequences of proteins and nucleic acids [6].

Algorithms for analysing Hidden Markov Models were discovered very early in their development. In 1967, Baum described an algorithm, now known as the forward-backward algorithm, which allows for the efficient calculation of observation probabilities and the estimation of individual states [2, 18]. Also in 1967, A.J. Viterbi formulated an algorithm for the decoding of convolution codes in telecommunications, which is directly applicable to Hidden

Markov Models in estimation of sample paths [23, 10]. Then in 1970, Baum and L. Welch discovered an algorithm for parameter estimation in a Hidden Markov Model [4, 18].

These are the three classical algorithms for solving problems in Hidden Markov Models. As we are primarily interested in state estimation, we will only make use of the forward-backward algorithm in our results, but we will also describe the Viterbi and Baum-Welch algorithms for the sake of completeness.

1.2 Recent Developments

The problem we are interested in, that of selecting an observation process for a Hidden Markov Model in which there are multiple choices, was first considered by J.S. Evans and V. Krishnamurthy in 1998 as a sensor scheduling problem [8, 9]. They considered the system not strictly as a Hidden Markov Model, but rather, as a standard Markov chain observed using one of several sensors, each with Gaussian noise depending on both the current state of the Markov chain and the observation method used.

Evans and Krishnamurthy defined optimality for such a system as the minimisation of a cost function, given by the sum of the observation error, as defined by a standard measure of uncertainty such as variance or information entropy, and an explicitly defined usage cost associated with each sensor. With such a broad definition, they did not attempt to solve the problem in general, instead describing a numeric example for which they presented an intuitive solution.

Krishnamurthy followed up this work with a paper in 2002 [12], this time describing the system as a *Markov Decision Process*. This is a model in which one has limited control over the transition probabilities of a Markov chain. Each action results in a different set of transition probabilities, with the aim being to decide which action to take so as to minimise a cost function based on the state of the Markov chain. If the Markov chain is not directly observable, the model is called a *Partially Observed Markov Decision Process*, for which there exists substantial literature concerning algorithmic solutions [17, 14].

The difference between a Partially Observed Markov Decision Process and a Hidden Markov Model with multiple observation processes is that in the former, one can alter the outcome and does so in order to achieve the optimal such outcome, while in the latter, the outcome cannot be altered and the aim is to achieve the optimal measurement of that outcome. However, it is possible to transform a Hidden Markov Model with multiple observation processes into a Markov Decision Process by considering it as a Markov chain in the *information state*—the posterior distribution of the underlying state given the sequence of observations—which is a sufficient statistic for the underlying state.

The Markov chain obtained by this transformation is fully observable, but its state space is uncountable. Krishnamurthy acknowledged this in his 2002 paper [12], presenting several algorithms which solved the problem in the sense that they would theoretically yield an optimal solution, but were analytically and computationally intractable due to the difficulty in working with an uncountable state space.

We note that Evans and Krishnamurthy always considered the primary trade-off as between the precision of estimation and the explicitly given cost of using each sensor. While such an approach is very appropriate in an engineering setting, the problem is interesting in itself without attaching a cost of usage to each sensor. Differences in sensors will arise from some being more suited to observing particular states than others, and, in a theoretical sense, this is a more natural trade-off than the costs involved with using them.

This was the approach taken by M. Rezaeian in 2007 [20], who also considered the problem as a Markov Decision Process in the information state, with the minimisation of long term expected information entropy as the optimality criteria, and no usage costs for sensors. However, Rezaeian did not proceed further than describing the problem, and gave an incorrect equation for the evolution of the information state; we present the correct version here as Equation (16).

In consideration of the work described above, it would appear that a straightforward application of the theory of Markov Decision Processes is unlikely to lead to a solution. Hence, we will not build upon the work of Evans, Krishnamurthy and Rezaeian, but will instead work towards an analytic rather than algorithmic solution. This was the basis of unpublished work by W. Moran and S. Suvorova—our approach is similar, but derived independently.

2 Hidden Markov Models

2.1 Markov Chains

At the centre of any results involving Hidden Markov Models is the underlying Markov chain, which, observable or not, describes the state of the system in question. Thus, we begin by introducing the Markov chain.

Definition 2.1. A **stochastic process** is a collection of random variables $\{X_t\}_{t\in\mathcal{T}}$ on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, indexed by a totally ordered set $\mathcal{T}$.

A **discrete time** stochastic process is one in which the index set $\mathcal{T}$ is countable; in this case, let $\mathcal{T} = \mathbb{Z}^+ = \{0, 1, 2, \dots\}$. A **discrete state** stochastic process is one in which the random variables $\{X_t\}_{t\in\mathcal{T}}$ take values in a countable set, which we again take to be $\mathbb{Z}^+$. A **finite state** stochastic process is defined analogously, taking the n distinct values $0, \dots, n-1$.

Definition 2.2. A **Markov chain** is a discrete time stochastic process $\{X_t\}_{t\in\mathbb{Z}^+}$ satisfying the **Markov property**, which is that the next state depends only on the current state and not on the path taken to reach the current state. For a discrete state Markov chain, the Markov property can be expressed by the equation

$$\mathbb{P}\big(X_{t+1} = x_{t+1} \big| X_t = x_t, \dots, X_0 = x_0\big) = \mathbb{P}\big(X_{t+1} = x_{t+1} \big| X_t = x_t\big). \qquad (1)$$

This holds for all time t, and all states $x_0, \dots, x_{t+1}$. If, in addition, this probability does not depend on t, the Markov chain is called *time homogeneous.*

Definition 2.3. The **transition matrix** of a finite state, time homogeneous Markov chain $\{X_t\}_{t\in\mathbb{Z}^+}$ is the $n \times n$ matrix T, where n is the number of states of the Markov chain, with entries given by the transition probabilities, that is, $T_{i,j} = \mathbb{P}(X_{t+1} = j | X_t = i)$.

We now take a moment to introduce some convenient notation. First, we will use a superscript in parentheses to denote the vector of all the values of a sequence up to the number within the parentheses. For example, we will write $X^{(t)}$ to mean the vector $(X_0, X_1, \ldots, X_t)$.

Second, when we write a random variable $X : \Omega \to \mathbb{R}$ in the context of an event, we mean the event $\{\omega \in \Omega : X(\omega) = x\}$ for some real number x, with the understanding that the result should hold no matter which x we pick. For example, we will write $\mathbb{P}(X|Y) = \mathbb{P}(X)$ to mean the statement

$$\mathbb{P}\big(X = x \big| Y = y\big) = \mathbb{P}\big(X = x\big) \text{ for all real numbers } x \text{ and } y.$$

Using this notation, we can write the discrete state Markov property as

$$\mathbb{P}\big(X_{t+1} \big| X^{(t)}\big) = \mathbb{P}\big(X_{t+1} \big| X_t\big). \tag{2}$$

If the Markov chain is also time homogeneous, we can write the Markov property in terms of the transition matrix, giving the equation

$$\mathbb{P}\big(X_{t+1} = j \big| X_t = i, X^{(t-1)}\big) = \mathbb{P}\big(X_{t+1} = j \big| X_t = i\big) = T_{i,j}. \tag{3}$$

2.2 Observation Processes

As originally described by L.E. Baum [3], the observation process is simply a probabilistic function of the underlying Markov chain. We now define this formally.

Definition 2.4. Given a finite state, time homogeneous Markov chain $\{X_t\}_{t \in \mathbb{Z}^+}$, an **observation process** for this Markov chain is a discrete time, finite state stochastic process $\{Y_t\}_{t \in \mathbb{Z}^+}$ on the same probability space, such that the current observation Y_t depends only on the current state X_t and not on any previous states or observations, and the next state X_{t+1} depends only on the current state X_t and not on either the current observation or any previous states or observations.

This can be written as the equations

$$\begin{aligned}\mathbb{P}\big(Y_t \big| X^{(t)}, Y^{(t-1)}\big) &= \mathbb{P}\big(Y_t \big| X_t\big),\\ \mathbb{P}\big(X_{t+1} \big| X^{(t)}, Y^{(t)}\big) &= \mathbb{P}\big(X_{t+1} \big| X_t\big).\end{aligned} \tag{4}$$

This must hold for all t, and we require that the probabilities in both equations do not depend on t.

Definition 2.5. The **observation matrix** is the matrix that describes the transition probabilities from X_t to Y_t in the same manner that the transition matrix describes the transition probabilities from X_t to X_{t+1}. More precisely, it is the $n \times m$ matrix M, where m is the number of states of the observation process, with entries

$$M_{i,j} = \mathbb{P}\big(Y_t = j \big| X_t = i\big) = \mathbb{P}\big(Y_t = j \big| X_t = i, Y^{(t-1)}, X^{(t-1)}\big). \tag{5}$$

We will use these symbols to represent the Markov chain and the observation process for the remainder of our discourse. In particular, $\{X_t\}_{t\in\mathbb{Z}^+}$ will be a Markov chain with state space $\{0, \dots, n-1\}$ and transition matrix T on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $\{Y_t\}_{t\in\mathbb{Z}^+}$ will be an observation process for $\{X_t\}_{t\in\mathbb{Z}^+}$ with state space $\{0, \dots, m-1\}$ and observation matrix M.

We also define π to be the distribution of the initial state X_0 of the Markov chain. With these definitions, the parameters of a Hidden Markov Model are fully specified by the triple (T, M, π). We will usually consider all three parameters to be known. However, as we are primarily interested in limiting behaviour, where the value of π is often unimportant, we will not give a particular value for π but instead aim to prove our results for all possible π.

With this construction, there are three basic problems of interest for Hidden Markov Models, which are those concerning the calculation of probabilities, the estimation of states or sample paths, and the estimation of parameters, as stated by L.R. Rabiner [18].

2.3 Calculation of Probabilities

The probability calculation problem is that of computing the probability of observing a particular sequence given the parameters of the model. Formally, given a Hidden Markov Model with parameters (T, M, π) and a sequence of observations $y^{(t)}$, we wish to calculate the probability of observing this sequence, that is,

$$\mathbb{P}\big(Y^{(t)} = y^{(t)}\big) \equiv \mathbb{P}\big(Y^{(t)} = y^{(t)} \big| (T, M, \pi)\big).$$

In theory, one can simply apply the law of total probability to obtain

$$\begin{aligned}\mathbb{P}\big(Y^{(t)} = y^{(t)}\big) &= \sum_{x^{(t)}} \Big[\mathbb{P}\big(Y^{(t)} = y^{(t)} \big| X^{(t)} = x^{(t)}\big) \cdot \mathbb{P}\big(X^{(t)} = x^{(t)}\big)\Big] \\ &= \sum_{x^{(t)}} \left[\pi(x_0) \cdot \prod_{s=0}^{t-1} T_{x_s, x_{s+1}} \cdot \prod_{s=0}^{t} M_{x_s, y_s}\right]. \end{aligned} \tag{6}$$

The sum is over the finitely many sample paths $x^{(t)}$.

While this is theoretically a very easy solution, calculating the probability using this formula is computationally infeasible, as it requires, at the very least, evaluation of n^{t+1} sample paths, which grows exponentially in time. Fortunately, there is a fairly simple recursive algorithm, the forward-backward algorithm first described by Baum in 1967 [2], which can be used to solve this problem.

The mathematical basis of this algorithm is similar, but the trick is to keep track of the state of the Markov chain.

Definition 2.6. The **forward variable** at time s is the vector valued function $\alpha_s : \{0, \ldots, m-1\}^{s+1} \to \mathbb{R}^n$ taking a sequence of observations $y^{(s)}$ to the probability of observing that sequence, for each current state of the system. More precisely, the entries of $\alpha_s \equiv \alpha_s[y^{(s)}]$ are given by

$$\alpha_s(j) \equiv \alpha_s\big[y^{(s)}\big](j) = \mathbb{P}\big(Y^{(s)} = y^{(s)}, X_t = j\big). \tag{7}$$

We will understand that α_s is a function of $y^{(s)}$ without writing the dependence explicitly. In particular, we will assume there is a fixed sequence of observations $y^{(t)}$ for some time t, and let s vary from 0 to t.

Then, we can apply the finite additivity property of probability and Equation (4) to obtain a recurrence relation for α_{s+1} in terms of α_s,

$$\begin{aligned}
\alpha_{s+1}(j) &= \mathbb{P}\big(Y^{(s+1)} = y^{(s+1)}, X_{s+1} = j\big) \\
&= \sum_{i=0}^{n-1} \mathbb{P}\big(Y^{(s+1)} = y^{(s+1)}, X_{s+1} = j, X_s = i\big) \\
&= \sum_{i=0}^{n-1} \mathbb{P}\big(Y^{(s)} = y^{(s)}, X_s = i\big) M_{j,y_{s+1}} T_{i,j} \\
&= M_{j,y_{s+1}} \sum_{i=0}^{n-1} \alpha_s(i) T_{i,j}. \qquad (8)
\end{aligned}$$

We also require an initial value α_0, which we obtain by

$$\alpha_0(j) = \mathbb{P}\big(Y_0 = y_0, X_0 = j\big) = M_{j,y_0}\pi(j). \qquad (9)$$

Then, the required probability can be extracted from the forward variable by summing over all states of the Markov chain, giving

$$\mathbb{P}\big(Y^{(t)} = y^{(t)}\big) = \sum_{j=0}^{n-1} \mathbb{P}\big(Y^{(t)} = y^{(t)}, X_t = j\big) = \sum_{j=0}^{n-1} \alpha_t(j). \qquad (10)$$

Thus, to calculate the probability of a sequence of observations, we need to calculate the $t+1$ variables α_s, $s = 0, \ldots, t$, where each α_s is a vector of n probabilities, each one calculated using n addition and $n+1$ multiplication operations. Importantly, the number of calculations required increases linearly with time, and the number of values that need to be stored is constant, since for each $s = 0, \ldots, t$, we only need to know the current forward variable α_s and can discard the previous ones.

Hence, this is a computationally feasible method for calculating the probability of a particular sequence of observations. We summarise this algorithm as follows.

Theorem 2.1. *The probability of observing a particular sequence $y^{(t)}$ is $\sum_{j=0}^{n-1} \alpha_t(j)$, where $\alpha_t(j)$ are the values of the forward variable, which can be calculated by the recurrence*

$$\begin{aligned} \alpha_0(j) &= M_{j,y_0}\pi(j), \\ \alpha_s(j) &= M_{j,y_{s+1}} \sum_{i=0}^{n-1} \alpha_s(i) T_{i,j}. \end{aligned} \tag{11}$$

Note that we have introduced the forward variable but no corresponding *backward variable*—such an object does exist, and is used in the Baum-Welch algorithm to solve the problem of parameter estimation. As such, it will be described in Section 2.6.

2.4 State Estimation

The state estimation problem is that of estimating the current state of the underlying Markov chain given a sequence of observations. However, rather than estimating the single most likely state, we will instead estimate the posterior distribution of the Markov chain given a sequence of observations, which is known as the *information state*. Note that we can easily determine the most likely state from the information state by simply taking the state with greatest probability.

Definition 2.7. The **information state** $z_t \equiv z_t[y^{(t)}]$ at time t given by a sequence of observations $y^{(t)}$ is the posterior distribution of the current state X_t given by those observations. The probabilities of this distribution are

$$z_t(j) \equiv z_t\big[y^{(t)}\big](j) = \mathbb{P}\big(X_t = j \big| Y^{(t)} = y^{(t)}\big). \tag{12}$$

As with the forward variable, we will understand that the information state depends on the sequence of observations, without writing this dependence explicitly in cases where we are considering a fixed such sequence.

Note that the information state is a probability distribution on $\{0, \dots, n-1\}$, and in particular, the space of information states is the set of all probability measures on a finite set of cardinality n, which can be considered as the simplex Δ of dimension $n-1$ contained in $\mathbb{R}^n$ given by

$$\Delta = \{(x_0, \dots, x_{n-1}) \in \mathbb{R}^n : x_0, \dots, x_{n-1} \geq 0, x_0 + \cdots + x_{n-1} = 1\}. \tag{13}$$

To evaluate the information state z_t arising from a sequence of observations $y^{(t)}$, we can relate z_t to the forward variable α_t, giving

$$\begin{aligned} z_t(j) &= \frac{\mathbb{P}\big(X_t = j, Y^{(t)} = y^{(t)}\big)}{\mathbb{P}\big(Y^{(t)} = y^{(t)}\big)} \\ &= \frac{\mathbb{P}\big(X_t = j, Y^{(t)} = y^{(t)}\big)}{\sum_{k=0}^{n-1} \mathbb{P}\big(X_t = k, Y^{(t)} = y^{(t)}\big)} \\ &= K_t \alpha_t(j), \end{aligned} \tag{14}$$

where K_t is some constant since the denominator does not depend on the state j. Then, we can use the recurrence relation for α_t, as written in Equation (8), to obtain one for z_t, giving

$$\begin{aligned} z_{t+1}(j) = K_{t+1}\alpha_{t+1}(j) &= K_{t+1} M_{j,y_{t+1}} \sum_{i=0}^{n-1} \alpha_t(i) T_{i,j} \\ &= \frac{K_{t+1}}{K_t} M_{j,y_{t+1}} \sum_{i=0}^{n-1} z_t(i) T_{i,j} \\ &= \frac{M_{j,y_{t+1}} \sum_{i=0}^{n-1} z_t(i) T_{i,j}}{\sum_{k=0}^{n-1} M_{k,y_{t+1}} \sum_{i=0}^{n-1} z_t(i) T_{i,k}}. \end{aligned} \tag{15}$$

Note that we have evaluated the constant K_{t+1}/K_t in terms of z_t from our knowledge that $\sum_{j=0}^{n-1} z_{t+1}(j) = 1$.

We can write this in matrix form by considering z_t as a row vector of probabilities. Let $D[y_{t+1}]$ be the $n \times n$ diagonal matrix with entries given by $D[y_{t+1}]_{i,i} = M_{i,y_{t+1}}$, and let $\mathbf{1}$ be a $n \times 1$ column vector of ones. This allows us to write the recurrence relation for the information state z_t in a succinct manner, as follows.

Theorem 2.2. *The information state z_t follows the recurrence relation*

$$z_{t+1} = f(z_t, y_{t+1}) = \frac{z_t \cdot T \cdot D[y_{t+1}]}{z_t \cdot T \cdot D[y_{t+1}] \cdot \mathbf{1}}, \tag{16}$$

defining $f : \Delta \times \{0, \dots, m-1\} \to \Delta$ to be the function which takes the current information state z_t and the next observation y_{t+1} to the next information state z_{t+1}.

Remark. If we know the initial distribution π of the underlying Markov chain, then clearly we should initialise the information state to $z_0 = \pi$. However, this method generalises to the case where we do not know the initial state; we simply pick $z_0 \in \Delta$ to be any prior distribution on the state space of the Markov chain.

Corollary 2.3. *The maximum likelihood estimate of the current state is*

$$\hat{x}_t = \underset{i}{\operatorname{argmax}} \Big\{ z_t(i) \Big\}. \tag{17}$$

2.5 Sample Path Estimation

One potential problem with state estimation occurs when we wish to know the sequence of most likely states. If we take the most likely state given by each information state, it is possible to obtain an estimated sequence which is impossible, in that it contains transitions which have probability zero. Thus, if we are interested in the sequence of states, it makes sense to find the sample path of maximum likelihood, rather than the sequence of individually most likely states.

Formally, we wish to find the sample path $x^{(t)}$, where t is the current time, which maximises

$$\mathbb{P}\big(X^{(t)} = x^{(t)} \big| Y^{(t)} = y^{(t)}\big).$$

The solution to this problem is the *Viterbi algorithm* [23, 10].

Definition 2.8. The **Viterbi score** $v_s(j)$ of a state $j = 0, \ldots, n-1$ at time s, given an observed sequence $y^{(t)}$, is the maximal probability of observing that sequence when the chain is currently in state j. This is given by the equation

$$v_s(j) \equiv v_s\big[y^{(s)}\big](j) = \max_{x^{(s-1)}} \Big\{ \mathbb{P}\big(X_s = j, X^{(s-1)} = x^{(s-1)}, Y^{(s)} = y^{(s)}\big) \Big\}. \quad (18)$$

As with the forward variable and the information state, we will understand the dependence on the observed sequence and not write this dependence explicitly when considering a fixed sequence of observations.

We can evaluate these probabilities to obtain

$$v_s(x_s) = \max_{x^{(s-1)}} \left\{ \pi(x_0) \cdot \prod_{k=0}^{s-1} T_{x_k, x_{k+1}} \cdot \prod_{k=0}^{s} M_{x_k, y_k} \right\}, \quad (19)$$

for each $x_s = 0, \ldots, n-1$. Note that we have chosen to use x_s as our dummy variable instead of i in order to simplify the form of the product.

Then, a recurrence for v_{s+1} in terms of v_s is given by

$$\begin{aligned} v_{s+1}(x_{s+1}) &= \max_{x^{(s)}} \left\{ \pi(x_0) \cdot \prod_{k=0}^{s} T_{x_k, x_{k+1}} \cdot \prod_{k=0}^{s+1} M_{x_k, y_k} \right\} \\ &= \max_{x_s} \left\{ \max_{x^{(s-1)}} \left\{ \pi(x_0) \cdot \prod_{k=0}^{s-1} T_{x_k, x_{k+1}} \cdot \prod_{k=0}^{s} M_{x_k, y_k} \right\} T_{x_s, x_{s+1}} \right\} M_{x_{s+1}, y_{s+1}} \\ &= \max_{x_s} \Big\{ v_s(x_s) T_{x_s, x_{s+1}} \Big\} M_{x_{s+1}, y_{s+1}}. \end{aligned} \quad (20)$$

This can be written more concisely as

$$v_{s+1}(j) = \max_i \Big\{ v_s(i) T_{i,j} \Big\} M_{j, y_{t+1}}. \quad (21)$$

We also need an initial value for this recurrence, which is given by

$$v_0(j) = \mathbb{P}(X_0 = j, Y_0 = y_0) = \pi(j) M_{j, y_0}. \quad (22)$$

In fact, we could have derived this recurrence by comparing the Viterbi score to the forward variable. We can write the forward variable as

$$\alpha_s(j) = \mathbb{P}\big(Y^{(s)}\!=\!y^{(s)}, X_s\!=\!j\big) = \sum_{x^{(s-1)}} \mathbb{P}\big(X_s = j, X^{(s-1)} = x^{(s-1)}, Y^{(s)} = y^{(s)}\big), \tag{23}$$

which is exactly the definition of the Viterbi score, with the maximum replaced by a sum. This means we can simply replace the sum with a maximum in Equation (8), the recurrence for the forward variable, to obtain the recurrence for the Viterbi score. By inspection, this indeed gives Equation (21).

Having defined the Viterbi score and derived a recurrence for it, it remains to find the most likely sample path, which we denote by $\hat{x}^{(t)}$. For a fixed sequence of observations $y^{(t)}$, $v_t(i)$ is proportional to

$$\max_{x^{(t-1)}} \Big\{\mathbb{P}\big(X_t = i, X^{(t-1)} = x^{(t-1)} \big| Y^{(t)} = y^{(t)}\big)\Big\},$$

by definition of conditional probability. But this is simply the probability of the most likely sample path ending in state i, hence the probability of the overall most likely sample path is given by the maximum of this probability across all states $i = 0, \ldots, n-1$. In particular, the final state in this most likely sample path is

$$\hat{x}_t = \operatorname*{argmax}_i \big\{v_t(i)\big\}. \tag{24}$$

In order to determine the rest of the sample path, we need to keep track of which state we picked at time s to reach state j at time $s+1$ with maximum likelihood. If we let that state be $u_{s+1}(j)$, it is clear that it is given by

$$u_{s+1}(j) = \operatorname*{argmax}_i \Big\{v_s(i) T_{i,j}\Big\}. \tag{25}$$

In particular, if j is the state at time $s+1$ in the most likely sample path, then $u_{s+1}(j)$ gives the state at time s. Then, we can work out all previous states by backtracking in this way, giving the recurrence

$$\hat{x}_s = u_{s+1}(\hat{x}_{s+1}). \tag{26}$$

The resulting sequence $\hat{x}^{(t)}$ is the most likely sample path calculated by the Viterbi algorithm. We summarise this result below.

Theorem 2.4. *Given a sequence of observations $y^{(t)}$, the most likely sample path is $\hat{x}^{(t)}$, which can be calculated by the following recurrences:*

$$\begin{aligned}
v_0(j) &= \pi(j) M_{j,y_0};\\
v_{s+1}(j) &= \max_i \Big\{ v_s(i) T_{i,j} \Big\} M_{j,y_{t+1}};\\
u_{s+1}(j) &= \operatorname*{argmax}_i \Big\{ v_s(i) T_{i,j} \Big\};\\
\hat{x}_t &= \operatorname*{argmax}_i \big\{ v_t(i) \big\};\\
\hat{x}_s &= u_{s+1}(\hat{x}_{s+1}).
\end{aligned} \tag{27}$$

2.6 Parameter Estimation

Finally, the parameter estimation problem is that of estimating the parameters of the Hidden Markov Model given a sequence of observations. Formally, given a sequence of observations $y^{(t)}$, we wish to pick the vector of parameters (T, M, π) which maximises

$$\mathbb{P}\big(Y^{(t)} = y^{(t)} \big| (T, M, \pi)\big).$$

As noted by Rabiner [18], this is by far the hardest of the problems, and one for which there is a multitude of imperfect solutions. While we will not be interested in parameter estimation for our work, we present for completeness a classic solution, the Baum-Welch algorithm [4, 18], which finds a local optimum for (T, M, π). We begin with a definition of the *backward variable* which we briefly mentioned in Section 2.3.

Definition 2.9. The **backward variable** at time s with respect to a final time $t \geq s$ is the vector $\beta_s \in \mathbb{R}^n$ with entries given by

$$\beta_s(i) = \mathbb{P}\big(Y_{s+1} = y_{s+1}, \ldots, Y_t = y_t \big| X_s = i\big). \tag{28}$$

Note that as with the forward variable, the backward variable depends on the sequence of observations, but since we are considered a fixed such sequence, this dependence will not be written explicitly.

We can derive a recurrence for the backward variable similar to that for the forward variable. Clearly,

$$\beta_t(i) = \mathbb{P}\big(\Omega \big| X_t = i\big) = 1, \tag{29}$$

and for $s < t$,

$$\begin{aligned}
\beta_s(i) &= \sum_{j=0}^{n-1} \mathbb{P}\big(Y_{s+1} = y_{s+1}, \ldots, Y_t = y_t, X_{s+1} = j \big| X_s = i\big) \\
&= \sum_{j=0}^{n-1} \mathbb{P}\big(Y_{s+2} = y_{s+2}, \ldots, Y_t = y_t \big| Y_{s+1} = y_{s+1}, X_{s+1} = j, X_s = i\big) \\
&\qquad\qquad \times \mathbb{P}\big(Y_{s+1} = y_{s+1}, X_{s+1} = j \big| X_s = i\big) \\
&= \sum_{j=0}^{n-1} \beta_{s+1}(j) T_{i,j} M_{j,y_{s+1}},
\end{aligned} \tag{30}$$

by Equation (4), which gives the independence relations between the underlying process and the observation process. This gives the following result.

Lemma 2.5. *The backward variable $\beta_s(i)$ with respect to the final time t satisfies the recurrence relation*

$$\begin{aligned}
&\beta_t(i) = 1, \\
&\beta_s(i) = \sum_{j=0}^{n-1} \beta_{s+1}(j) T_{i,j} M_{j,y_{s+1}}.
\end{aligned} \tag{31}$$

We also define the quantity $\xi_s(i,j)$ to be the probability of a transition from state i to state j at time s, given the sequence of observations $y^{(t)}$ for $t > s$. More precisely,

$$\xi_s(i,j) = \mathbb{P}\big(X_s = i, X_{s+1} = j \big| Y^{(t)} = y^{(t)}\big). \tag{32}$$

We can relate this to the forward and backward variables. By definition of conditional expectation, for some constant L_s,

$$
\begin{aligned}
\xi_s(i,j) &= L_s \times \mathbb{P}\big(X_s = i, X_{s+1} = j, Y^{(t)} = y^{(t)}\big) \\
&= L_s \times \mathbb{P}\big(Y_{s+2} = y_{s+2}, \ldots, Y_t = y_t \big| Y^{(s+1)} = y^{(s+1)}, X_s = i, X_{s+1} = j\big) \\
&\qquad \times \mathbb{P}\big(Y_{s+1} = y_{s+1}, X_{s+1} = j \big| Y^{(s)} = y^{(s)}, X_s = i\big) \\
&\qquad \times \mathbb{P}\big(Y^{(s)} = y^{(s)}, X_s = i\big) \\
&= L_s \times \beta_{s+1}(j) T_{i,j} M_{j,y_{s+1}} \alpha_s(i) \\
&= \frac{\alpha_s(i) T_{i,j} M_{j,y_{s+1}} \beta_{s+1}(j)}{\sum_{k=0}^{n-1} \sum_{l=0}^{n-1} \alpha_s(k) T_{k,l} M_{l,y_{s+1}} \beta_{s+1}(l)},
\end{aligned} \tag{33}
$$

where we have evaluated the constant L_s in terms of α_s, β_{s+1}, T and M from our knowledge that $\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \xi_s(i,j) = 1$.

Note that the variables $\xi_s(i,j)$ depend both on $y^{(t)}$, which is a known quantity, and on (T, M, π), which is an unknown quantity. However, we can guess some parameters (T, M, π), and then use the variables $\xi_s(i,j)$ to estimate new parameters $(\hat{T}, \hat{M}, \hat{\pi})$. We formalise this process as follows.

Definition 2.10. The **Baum-Welch reestimate** of the triple of parameters (T, M, π) is the triple of parameters $(\hat{T}, \hat{M}, \hat{\pi})$ given by

$$
\begin{aligned}
\hat{\pi}(i) &= \mathbb{P}(X_0 = i) = \sum_{k=0}^{n-1} \xi_0(i,k), \\
\hat{T}_{i,j} &= \frac{\sum_{s=0}^{t-1} \mathbb{P}(X_s = i, X_{s+1} = j)}{\sum_{s=0}^{t-1} \mathbb{P}(X_s = i)} = \frac{\sum_{s=0}^{t-1} \xi_s(i,j)}{\sum_{s=0}^{t-1} \sum_{k=0}^{n-1} \xi_s(i,k)}, \\
\hat{M}_{i,j} &= \frac{\sum_{s:y_s=j} \mathbb{P}(X_s = i)}{\sum_{s=0}^{t} \mathbb{P}(X_s = i)} = \frac{\sum_{s:y_s=j} \sum_{k=0}^{n-1} \xi_s(i,k)}{\sum_{s=0}^{t} \sum_{k=0}^{n-1} \xi_s(i,k)}.
\end{aligned} \tag{34}
$$

Note that all probabilities given above are conditional on the event $Y^{(t)} = y^{(t)}$ representing the known sequence of observations—we have opted not to write this explicitly so as to make the equations less cluttered. The sum in the numerator of the expression for $\hat{M}_{i,j}$ is over all times s such that the observation at that time, y_s, is equal to j.

The intuitive justification for this reestimation procedure is that $\hat{\pi}(i)$ is the expected probability that X_0 is in state i, $\hat{T}_{i,j}$ is the expected number of transitions from state i to state j divided by the expected total number of transitions away from state i, and $\hat{M}_{i,j}$ is the expected number of observations j made in state i divided by the expected total number observations made in state i.

Having estimated new parameters $(\hat{T}, \hat{M}, \hat{\pi})$, we can use them to calculate new values for the variables $\xi_s(i,j)$, and thereby iterate the process to obtain a sequence of estimated parameters. Baum proved that this process of reestimation of parameters improves the estimate, in the sense described below.

Theorem 2.6. *The Baum-Welch reestimate $(\hat{T}, \hat{M}, \hat{\pi})$ is better than the original estimate (T, M, π) in the sense that*

$$\mathbb{P}\big(Y^{(t)} = y^{(t)} \big| (\hat{T}, \hat{M}, \hat{\pi})\big) \geq \mathbb{P}\big(Y^{(t)} = y^{(t)} \big| (T, M, \pi)\big). \tag{35}$$

Furthermore, a sequence of parameter estimates obtained by the Baum-Welch reestimation procedure converges to a local maximum of

$$\mathbb{P}\big(Y^{(t)} = y^{(t)} \big| (T, M, \pi)\big),$$

considered as a function of (T, M, π).

This theorem is the theoretical basis for the Baum-Welch algorithm, and is named for Baum and L. Welch who first proposed the algorithm in an unpublished paper; the actual first publication was in a paper by Baum and others in 1970 [4]. Since we will not be considering the problem of parameter estimation in our work, we omit the proof of this theorem.

3 Multiple Observation Processes

3.1 Motivations

In the previous chapter, we discussed some of the well-known theory of Hidden Markov Models, in which an underlying Markov chain cannot be observed directly, but is instead observed with error. In particular, we have assumed that there is only one such way in which to observe the chain. However, it will often be the case that there is more than one way to observe, each being associated with different errors.

This situation often arises when there are multiple sensors in an area, each attempting to measure the same quantity, but producing different errors due to their positioning or other differences. If there are sufficient resources available to operate all sensors simultaneously and process their outputs, it is possible to consider them as a single collective sensor. For example, if k sensors each produce one of m possible outputs, such a collective sensor would produce m^k possible outputs, which is still finite, hence the problem is essentially identical.

The more interesting situation in which multiple observation methods arise is when a single sensor is able to operate in several different modes, each with its advantages and disadvantages. It is highly unlikely that one mode of operation is superior to all others in every situation, which means the optimal choice of sensor will depend on the current state of the underlying system being observed. A simple but commonly occurring example is radar, in which one must choose an electromagnetic pulse to transmit at each time interval—for example, a low frequency pulse is better able to detect targets at long range, while a high frequency pulse is better suited to targets at short range [13].

We formulate this problem by supposing there are l observation methods, each represented by a $n \times m$ observation matrix $M[i]$, $i \in \{0, \ldots, l-1\}$. Note that we have assumed a common observation space; this is possible since we can embed all observation spaces within the one with largest cardinality, noting that this will present no problems since such spaces are finite sets without any additional constructions such as a metric or norm.

The choice of observation matrix will be based on the previous information state. Recall from Section 2.4, that the space of information states can be considered as a simplex Δ of dimension $n-1$ contained in $\mathbb{R}^n$. This leads to the following definition.

Definition 3.1. A **policy** is a function $g : \Delta \to \{0, \ldots, l-1\}$ which takes an information state to the index of an observation matrix. Alternatively, we can define a policy by the of preimages of g, which forms an ordered partition $\{A_0, \ldots, A_{l-1}\}$ of the space of information states Δ, with the property that g maps A_i to i for each $i \in \{0, \ldots, l-1\}$.

This leads directly to the problem which we will be considering. Given a sequence of observations from a corresponding sequence of observation matrices, it is not difficult to calculate the information state, which requires only a minor adjustment to the recurrence relation presented in Equation (16). The difficult part of the problem is deciding which observation matrix to use, that is, determining a policy, and in particular, one that is optimal under some criterion.

As mentioned previously, this criterion will be the minimisation of the long term expected information entropy of the information state, where the information entropy of an information state z_t is the quantity

$$H(z_t) = -\sum_{i=0}^{n-1} z_t(i) \log z_t(i). \tag{36}$$

We will use the natural logarithm here, as opposed to the base 2 logarithm commonly used in the information theory literature, noting all choices of base are equivalent for the minimisation problem, as can be seen from the equation $\log_a(b) = \frac{\log(b)}{\log(a)}$. We will also use the word entropy to mean information entropy, where convenient.

By convention, we define $0 \log 0 = 0$, on the basis that

$$\lim_{x \to 0^+} \big(x \log x\big) = 0.$$

Lemma 3.1. *The information entropy $H(z_t)$ is minimised when z_t is a non-random information state, that is, one in which the probability of one state is 1 and the probabilities of the other states are 0. It is maximised when z_t is the uniformly distributed information state.*

Proof. Since $-x \log x \geq 0$ with equality at $x = 0$ and $x = 1$, it follows that $H(z_t)$ is minimised when all of the $z_t(i)$ are either 0 or 1. Since the sum of the $z_t(i)$ is 1, it follows that all of them are 0 except one which is 1.

To maximise $H(z_t) = -\sum_i z_t(i) \log z_t(i)$, with the constraint $\sum_i z_t(i) = 1$, we can use the method of Lagrange multipliers. This gives the equation $\nabla H(z_t) = \lambda \nabla \sum_i z_t(i)$. Hence, $-\log z_t(i) - 1 = \lambda$, so the $z_t(i)$ are equal, and therefore equal to $\frac{1}{n}$.

It is not difficult to see that this is the unique global maximum. In particular, the maximum entropy for a n state distribution is

$$H(z_t) = -\sum_{i=0}^{n-1} \tfrac{1}{n} \log \tfrac{1}{n} = \log n. \tag{37}$$

An information theoretic proof of this lemma can be found in the book by T.M. Cover and J.A. Thomas [5]. □

This result shows that entropy is a suitable measure of uncertainty—in fact, it can be defined axiomatically by certain desirable properties of such a measure [5]. However, this is beyond the scope of our work. We will merely note that in this context, entropy is much more suitable than the other common measure of uncertainty, variance, since our state space is a finite set without well-defined operations of addition and scalar multiplication, nor any natural notion of distance other than the discrete metric.

Having decided upon information entropy for our measure of uncertainty, it remains to define what we mean by long term expected entropy. In particular, it is necessary to define Z_t as the distribution on z_t, that is, a distribution which takes value $z_t \equiv z_t[y^{(t)}]$ with probability $\mathbb{P}(Y^{(t)} = y^{(t)})$. If we let $\delta(x)$ represent the unit point mass at x, then we can write

$$Z_t = \sum_{y^{(t)}} \mathbb{P}\big(Y^{(t)} = y^{(t)}\big)\, \delta\big(z_t[y^{(t)}]\big). \tag{38}$$

This is a probability measure. In particular, for any measurable set $A \subseteq \Delta$,

$$Z_t(A) = \sum_{y^{(t)}} \mathbb{P}\big(Y^{(t)} = y^{(t)}\big)\, \mathbb{I}_A\big(z_t[y^{(t)}]\big), \tag{39}$$

where $\mathbb{I}_A(x)$ is the indicator function which takes value 1 if $x \in A$ and 0 otherwise.

Then, the expected information entropy of the information state at time t is

$$H(Z_t) \equiv \int_\Delta H(z)\, dZ_t(z), \tag{40}$$

and the long run expected information entropy, if it exists, is

$$\lim_{t\to\infty} H(Z_t) = \lim_{t\to\infty} \int_\Delta H(z)\, dZ_t(z). \tag{41}$$

Note that the information state z_t is given by a recurrence based on the previous information state z_{t-1}, the new observation y_t and the observation matrix used. This means the distribution of information states Z_t over all possible sequences of observations $y^{(t)}$ depends only on the prior information state z_0, which may be generalised to a prior distribution of information states Z_0, and the policy g. This leads us to the following definition.

Definition 3.2. A policy $g : \Delta \to \{0, \ldots, l-1\}$ is called **optimal** if, for any prior information state distribution Z_0, the long term expected information entropy resulting from the combination of g and Z_0 is equal to the infimum of the set of achievable such entropies.

Hence, we can state our problem as:

- Under what conditions does the long term expected entropy exist?
- Under what conditions does an optimal policy exist?
- If an optimum policy exists, how can we describe it?

3.2 Information State

We now describe some of the properties of the information state. Ultimately, we wish to understand the limiting behaviour of its distribution. We begin by arguing that it is the correct variable to be investigating, due to the fact that it is a *sufficient statistic*, which we mentioned in passing in the Introduction.

Definition 3.3. A statistic T is a **sufficient statistic** for the parameter λ given data X, if T is a deterministic function of X, and no further information about λ can be inferred from X than what is contained in T.

Lemma 3.2. *The information state $z_t(i)$ is a sufficient statistic for the parameter $\mathbb{P}(X_t = i)$ given the sequence of observations $y^{(t)}$.*

Proof. By definition, $z_t(i) = \mathbb{P}(X_t = i | Y^{(t)} = y^{(t)})$. This clearly satisfies the properties required for a sufficient statistic. □

This shows that the information state z_t is a suitable choice of variable. Having seen this, we now investigate its behaviour. We saw in Theorem 2.2 that the evolution of the information state is given by a recurrence, which is easily adapted to reflect the observation matrix used. This gives us the equation

$$z_{t+1} = f(z_t, y_{t+1}) = \frac{z_t \cdot T \cdot D[g(z_t), y_{t+1}]}{z_t \cdot T \cdot D[g(z_t), y_{t+1}] \cdot \mathbf{1}}, \tag{42}$$

where $D[g(z_t), y_{t+1}]$ is the diagonal matrix corresponding to the chosen observation matrix $M[g(z_t)]$, with entries $D[g(z_t), y_{t+1}]_{i,i} = M[g(z_t)]_{i,y_{t+1}}$, and $\mathbf{1}$ is a vector of ones.

This gives us a natural way to derive the form of the recurrence F_t on the distribution of information state, defined by

$$Z_{t+1} = F_t(Z_t).$$

In particular, a unit point mass $Z_t = \delta(z)$ evolves to

$$\begin{aligned} Z_{t+1} &= \sum_{y=0}^{m-1} \mathbb{P}(Y_{t+1} = y)\, \delta\big(f(z, y)\big) \\ &= \sum_{y=0}^{m-1} \mathbb{P}(Y_{t+1} = y)\, \delta\left(\frac{z \cdot T \cdot D[g(z), y]}{z \cdot T \cdot D[g(z), y] \cdot \mathbf{1}}\right) \\ &= \sum_{y=0}^{m-1} \mathbb{P}(Y_{t+1} = y) \sum_{i=0}^{l-1} \mathbb{I}_{A_i}(z)\, \delta\left(\frac{z \cdot T \cdot D[i, y]}{z \cdot T \cdot D[i, y] \cdot \mathbf{1}}\right). \end{aligned} \tag{43}$$

Note that $\mathbb{P}(Y_{t+1} = y)$ depends implicitly on the chosen observation matrix and hence z, and is given by

$$\mathbb{P}(Y_{t+1} = y) = \big(\pi \cdot T^{t+1} \cdot M[g(z)]\big)_y, \tag{44}$$

where π is the distribution of X_0, considered as a row vector of probabilities.

Having found the form of F_t for a unit point mass, we can integrate to find its general form, which is

$$\begin{aligned} F_t(Z_t) &= \int_\Delta \delta(z)\, dZ_t(z) \\ &= \int_\Delta \sum_{y=0}^{m-1} \mathbb{P}(Y_{t+1} = y) \sum_{i=0}^{l-1} \mathbb{I}_{A_i}(z)\, \delta\left(\frac{z \cdot T \cdot D[i, y]}{z \cdot T \cdot D[i, y] \cdot \mathbf{1}}\right) dZ_t(z) \\ &= \sum_{i=0}^{l-1} \sum_{y=0}^{m-1} (\pi \cdot T^{t+1} \cdot M[i])_y \int_{A_i} \delta\left(\frac{z \cdot T \cdot D[i, y]}{z \cdot T \cdot D[i, y] \cdot \mathbf{1}}\right) dZ_t(z) \end{aligned} \tag{45}$$

Note that F_t has some time dependence, which means it will not in general have any fixed points. A convergent sequence of iterations of F_t must instead converge to a fixed point of the limiting form F of F_t, given by

$$F(Z) = \sum_{i=0}^{l-1} \sum_{y=0}^{m-1} (\pi \cdot T^{\infty} \cdot M[i])_y \int_{A_i} \delta\left(\frac{z \cdot T \cdot D[i, y]}{z \cdot T \cdot D[i, y] \cdot \mathbf{1}} \right) dZ(z). \tag{46}$$

Note that as long as the Markov chain $\{X_t\}_{t \in \mathbb{Z}^+}$ is not periodic, the limiting distribution $\pi \cdot T^{\infty}$ exists by direct application of the Perron-Frobenius Theorem [21]. We give the precise relation between F_t and F as follows.

Definition 3.4. A **global attractor** of a sequence of functions $\{G_t\}_{t \in \mathbb{Z}^+}$, from a space X to itself, is a set $A \subset X$ such that for any sequence $\{x_t\}_{t \geq 0}$ contained in X satisfying $x_{t+1} = G_t(x_t)$, any open neighbourhood of A contains a tail of the sequence. For a single function G, we simply consider the constant sequence $\{G\}_{t \in \mathbb{Z}^+}$.

We note that while this is not the usual definition of an attractor—which usually requires it be invariant under the evolution function—but this is a necessary generalisation in order to consider attractors of non-stationary but convergent systems. In particular, if we can show that a single point $\{x\} \subset X$ is a global attractor, it will have similar properties to a single point global attractor of a stationary system under the usual definition.

Lemma 3.3. *A subset of the set of distributions on Δ is a global attractor of the process F_t if and only if it is a global attractor of the process F.*

Proof. It suffices to show that F_t converges to F uniformly. This is true since their supports are identical, and the convergence of the probability $(\pi \cdot T^{t+1} \cdot M[i])_y$ to $(\pi \cdot T^{\infty} \cdot M[i])_y$ is uniform since y takes only finitely many values. □

Hence, it suffices to look for stationary points of F. Note that these points are actually measures on the space of information states. For this reason, we will also call them *invariant measures* when we wish to emphasise this fact.

3.3 Two Dimensions

For the remainder of our results, the examples presented in Section 3.4 and the special case presented in Chapter 4, we will only be dealing with the two dimensional case. This is the case where the underlying Markov chain takes two states, the observation process takes two states, and there is a choice between at most two observation matrices. In terms of quantities we have defined, we are setting $n = m = l = 2$. Many of the equations we have derived take much simpler forms in this case.

In particular, it is redundant to consider the information state z_t as a probability vector $(z_t(0), z_t(1))$, since we know that $z_t(1) = 1 - z_t(0)$. Thus, for the remainder of our discourse, we will write the information state as a scalar $z_t \in [0, 1]$, representing what was previously called $z_t(0)$, that is,

$$z_t \equiv z_t\big[y^{(t)}\big] = \mathbb{P}\big(X_t = 0\big|Y^{(t)} = y^{(t)}\big). \tag{47}$$

The entropy associated with this information state is then

$$H(z_t) = -z_t \log z_t - (1 - z_t) \log(1 - z_t). \tag{48}$$

The recurrence for the information state, as stated in Theorem 2.2, becomes

$$\begin{aligned} z_{t+1} &= f(z_t, y_{t+1}) \\ &= \frac{M_{0,y_{t+1}}\big(z_t T_{0,0} + (1 - z_t)T_{1,0}\big)}{M_{0,y_{t+1}}\big(z_t T_{0,0} + (1 - z_t)T_{1,0}\big) + M_{1,y_{t+1}}\big(z_t T_{0,1} + (1 - z_t)T_{1,1}\big)}. \end{aligned} \tag{49}$$

From this, we can see that Z_t, the distribution of the information state, follows the recurrence F_t, where

$$\begin{aligned} Z_{t+1} &= F_t(Z_t) \\ &= \mathbb{P}(Y_t = 0) \int_{[0,1]} \delta\Big(f(z, 0)\Big)\, dZ_t(z) + \mathbb{P}(Y_t = 1) \int_{[0,1]} \delta\Big(f(z, 1)\Big)\, dZ_t(z). \end{aligned} \tag{50}$$

The limiting behaviour of this process, F, is given by

$$F(Z) = \mathbb{P}(Y_\infty = 0) \int_{[0,1]} \delta\Big(f(z, 0)\Big)\, dZ(z) + \mathbb{P}(Y_\infty = 1) \int_{[0,1]} \delta\Big(f(z, 1)\Big)\, dZ(z). \tag{51}$$

In the case of multiple observation matrices, where we are using a policy to choose which one to use, we can adapt Equation (46) to obtain

$$F(Z) = \sum_{i=0}^{1} \sum_{y=0}^{1} \mathbb{P}(Y_\infty = y) \int_{A_i} \delta\Big(f(z, y)\Big)\, dZ(z). \tag{52}$$

3.4 Examples

We now give some examples to show that, with this definition of optimality, the choice of policy significantly affects the long-run performance. For simplicity, we will assume the two dimensional case, as in Section 3.3.

We begin by picking an appropriate Markov chain to use as the underlying process throughout this section. It is clear that some Markov chains do not readily lend themselves to a Hidden Markov Model with multiple observation processes. In particular, we need the dependence between consecutive states to be neither too high nor too low, as shown by the following examples.

Example 3.1. Consider a Hidden Markov Model with transition matrix

$$T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{53}$$

In this case, picking a sample path of the Markov chain is equivalent to picking an initial state. Since the chain does not evolve in time, the observation process Y_t is a sequence of independent and identically distributed Bernoulli random variables, with probability of success

$$\mathbb{P}(Y_t = 0) = \begin{cases} M_{0,0} & \text{if } X_0 = 0, \\ M_{1,0} & \text{if } X_0 = 1. \end{cases} \tag{54}$$

Note that this probability does not depend on t, hence the identical distribution of the Y_t. This means that in the long run, we can determine $\mathbb{P}(Y_t = 0)$ with arbitrary precision.

Thus, unless $M_{0,0} = M_{1,0}$, which is an uninteresting scenario since it means the observation process is independent of the underlying process, we can determine X_0 and hence X_t for all t with arbitrary precision. This means the long run expected information entropy for any non-trivial observation process is zero, hence there is no benefit to having multiple observation processes.

Example 3.2. Consider a Hidden Markov Model with transition matrix

$$T = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}. \tag{55}$$

This means the underlying Markov chain is in fact a sequence of independent and identically distributed Bernoulli random variables with probability of success $\frac{1}{2}$. Then, the information state, by Theorem 2.2 and Equation (49), is given by

$$\begin{aligned} z_{t+1} &= \frac{M_{0,y_{t+1}}\big(\frac{1}{2}z_t + \frac{1}{2}(1-z_t)\big)}{M_{0,y_{t+1}}\big(\frac{1}{2}z_t + \frac{1}{2}(1-z_t)\big) + M_{1,y_{t+1}}\big(\frac{1}{2}z_t + \frac{1}{2}(1-z_t)\big)} \\ &= \frac{M_{0,y_{t+1}}}{M_{0,y_{t+1}} + M_{1,y_{t+1}}}. \end{aligned} \tag{56}$$

Hence, the next information state does not depend on the current information state, so any policy based on the information state must be a constant one, that is, one in which we always use the same observation matrix. Again, there can be no benefit to having multiple observation processes.

For the remainder of the examples, we will use the transition matrix

$$T = \begin{bmatrix} \frac{4}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{4}{5} \end{bmatrix}. \tag{57}$$

The gives a fixed point, or stationary distribution, of $X_\infty = (\frac{1}{2}, \frac{1}{2})$. In particular, we can use this to calculate the long term one-step autocorrelation,

$$\begin{aligned} \rho &= \lim_{t\to\infty} \frac{\mathbb{E}[X_t X_{t+1}] - \mathbb{E}[X_t]\mathbb{E}[X_{t+1}]}{\mathbb{E}[X_t^2] - \mathbb{E}[X_t]^2} \\ &= \lim_{t\to\infty} \frac{\mathbb{P}(X_t = 1, X_{t+1} = 1) - \mathbb{P}(X_t = 1)\mathbb{P}(X_{t+1} = 1)}{\mathbb{P}(X_t = 1) - \mathbb{P}(X_t = 1)^2} \\ &= \frac{\frac{1}{2} \times \frac{4}{5} - \frac{1}{2} \times \frac{1}{2}}{\frac{1}{2} - (\frac{1}{2})^2} = \frac{3}{5}. \end{aligned} \tag{58}$$

Roughly speaking, this is not too high and not too low, and should allow for interesting examples. We now consider a variety of observation processes.

Example 3.3. Consider the Hidden Markov Model with transition matrix T and observation matrix

$$M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{59}$$

This corresponds to the case where we can observe the Markov chain perfectly. In this case, $Y_t = X_t$ for all t, and in fact, this is just an ordinary Markov chain and not a Hidden Markov Model. The information state is

$$z_t = \mathbb{P}\big(X_t = 0 \big| Y^{(t)} = y^{(t)}\big) = \begin{cases} 1 & \text{if } y_t = 0, \\ 0 & \text{if } y_t = 1. \end{cases} \tag{60}$$

This means the information entropy is always zero, and in particular, the long term expected information entropy is zero. Hence, this is the best observation matrix possible.

Example 3.4. Consider the Hidden Markov Model with transition matrix T and observation matrix

$$M = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}. \tag{61}$$

This corresponds to the case where we cannot make any useful observation. In particular, $Y^{(t)}$ is independent of $X^{(t)}$, and the information state is

$$z_t = \mathbb{P}\big(X_t = 0 \big| Y^{(t)} = y^{(t)}\big) = \mathbb{P}(X_t = 0). \tag{62}$$

Note that the information state z_t does not depend on the observations, which means the distribution of information states Z_t is simply the unit point mass at z_t. Since $X_\infty = (\frac{1}{2}, \frac{1}{2})$, the information state converges to $z_\infty = \frac{1}{2}$, and the long run average information entropy is

$$H(Z_\infty) = H(z_\infty) = -\tfrac{1}{2}\log\tfrac{1}{2} - (1 - \tfrac{1}{2})\log(1 - \tfrac{1}{2}) = \log 2 \approx 0.693147. \tag{63}$$

This is the highest possible information entropy for a random variable with two values, as seen in Lemma 3.1, hence this is the worst observation matrix possible.

Example 3.5. Consider the Hidden Markov Model with transition matrix T and observation matrix

$$M = M_0 = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}. \tag{64}$$

This corresponds to the case where we can observe state 0 perfectly, but cannot make any useful observation when the Markov chain is in state 1. In this case, the derivation of the information state is not trivial; we will proceed via the recurrence relation in Theorem 2.2, in the form stated in Equation (49), which we reproduce here:

$$z_{t+1} = \frac{M_{0,y_{t+1}}\big(z_t T_{0,0} + (1-z_t)T_{1,0}\big)}{M_{0,y_{t+1}}\big(z_t T_{0,0} + (1-z_t)T_{1,0}\big) + M_{1,y_{t+1}}\big(z_t T_{0,1} + (1-z_t)T_{1,1}\big)}. \tag{49}$$

If $y_{t+1} = 1$, then the numerator above contains a factor of $M_{0,1} = 0$, hence $z_{t+1} = 0$. Otherwise, if $y_{t+1} = 0$,

$$\begin{aligned} z_{t+1} &= \frac{M_{0,0}\big(z_t T_{0,0} + (1-z_t)T_{1,0}\big)}{M_{0,0}\big(z_t T_{0,0} + (1-z_t)T_{1,0}\big) + M_{1,0}\big(z_t T_{0,1} + (1-z_t)T_{1,1}\big)}. \\ &= \frac{\big(\frac{4}{5}z_t + \frac{1}{5}(1-z_t)\big)}{\big(\frac{4}{5}z_t + \frac{1}{5}(1-z_t)\big) + \frac{1}{2}\big(\frac{1}{5}z_t + \frac{4}{5}(1-z_t)\big)} \\ &= \frac{6z_t+2}{3z_t+6}. \end{aligned} \tag{65}$$

Defining a fractional linear transformation r_0 by

$$r_0(z) = \frac{6z+2}{3z+6}, \tag{66}$$

we can simplify the recurrence for the information state z_t to

$$z_{t+1} = \begin{cases} r_0(z_t) & \text{if } y_{t+1} = 0, \\ 0 & \text{if } y_{t+1} = 1. \end{cases} \tag{67}$$

We can also calculate

$$\mathbb{P}(Y_\infty = 0) = \mathbb{P}(X_\infty = 0)M_{0,0} + \mathbb{P}(X_\infty = 1)M_{1,0} = \frac{1}{2} \times 1 + \frac{1}{2} \times \frac{1}{2} = \frac{3}{4}. \tag{68}$$

This gives $Y_\infty = (\frac{3}{4}, \frac{1}{4})$, which we can substitute into Equation (51) to determine the limiting behaviour of the distribution of the information state, which is

$$F(Z) = \frac{3}{4}\int_{[0,1]} \delta\Big(r_0(z)\Big)\, dZ(z) + \frac{1}{4}\delta(0). \tag{69}$$

By inspection, the fixed point of this dynamical system is

$$Z_\infty = \frac{1}{4}\left(\delta(0) + \frac{3}{4}\delta\big(r_0(0)\big) + \left(\frac{3}{4}\right)^2 \delta\big(r_0^2(0)\big) + \left(\frac{3}{4}\right)^3 \delta\big(r_0^3(0)\big) + \dots\right). \tag{70}$$

We will prove in Section 4.2 that this fixed point is unique. Given that it is unique, we can calculate the long term expected information entropy as

$$H(Z_\infty) = \frac{1}{4}\left(H(0) + \frac{3}{4}H\big(r_0(0)\big) + \left(\frac{3}{4}\right)^2 H\big(r_0^2(0)\big) + \left(\frac{3}{4}\right)^3 H\big(r_0^3(0)\big) + \dots\right). \tag{71}$$

Since this is an example, it suffices approximate the answer numerically, for which we obtain

$$H(Z_\infty) \approx 0.438157. \tag{72}$$

In calculating this figure, we computed the sum of the first 1000 terms; since the value of $H(r_0^k(0))$ is bounded above by $\log 2 < 1$, this means we can bound the error by the geometric series $\frac{1}{4}((\frac{3}{4})^{1000} + (\frac{3}{4})^{1001} + \dots) = (\frac{3}{4})^{1001}$, which is much smaller than the last significant digit quoted above.

Example 3.6. Consider the Hidden Markov Model with transition matrix T and observation matrix

$$M = M_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{bmatrix}. \tag{73}$$

This corresponds to the case where we can observe state 1 perfectly, but cannot make any useful observation when the Markov chain is in state 0. By symmetry with Example 3.5, we see that the long run average information entropy must again be

$$H(Z_\infty) \approx 0.438157. \tag{74}$$

While we already have the numerical answer, it will be useful later on for us to calculate the corresponding fractional linear transformation that determines the evolution of the information state. Again using Equation (49), we see that $z_{t+1} = 1$ when $y_{t+1} = 0$, and when $y_{t+1} = 1$,

$$\begin{aligned} z_{t+1} &= \frac{M_{0,1}\big(z_t T_{0,0} + (1 - z_t)T_{1,0}\big)}{M_{0,1}\big(z_t T_{0,0} + (1 - z_t)T_{1,0}\big) + M_{1,1}\big(z_t T_{0,1} + (1 - z_t)T_{1,1}\big)}. \\ &= \frac{\frac{1}{2}\big(\frac{4}{5}z_t + \frac{1}{5}(1 - z_t)\big)}{\frac{1}{2}\big(\frac{4}{5}z_t + \frac{1}{5}(1 - z_t)\big) + \big(\frac{1}{5}z_t + \frac{4}{5}(1 - z_t)\big)}. \\ &= \frac{3z_t + 1}{-3z_t + 9} \end{aligned} \tag{75}$$

Hence, defining the fractional linear transformation

$$r_1(z) = \frac{3z + 1}{-3z + 9}, \tag{76}$$

and noting that by symmetry, $Y_\infty = (\frac{1}{4}, \frac{3}{4})$, we see that the recurrence for the distribution of the information state is

$$F(Z) = \frac{3}{4}\int_{[0,1]} \delta\Big(r_1(z)\Big)\, dZ(z) + \frac{1}{4}\delta(1). \tag{77}$$

Example 3.7. Now consider a Hidden Markov Model with transition matrix T, with a choice between two observation matrices $M[0] = M_0$ and $M[1] = M_1$, which are the same matrices defined in Examples 3.5 and 3.6. Clearly, the best way to combine the information from M_0 and M_1 is to use both observation methods at every point in time, but let us suppose this is not possible due to either computational or physical constraints. Hence, we must choose an observation matrix for each time t.

Note that Example 3.5 is equivalent to always choosing M_0, and Example 3.6 is equivalent to always choosing M_1. For this example, suppose that at each time $t \in \mathbb{Z}^+$, we will choose M_0 if t is even and M_1 if t is odd. Note that this is not a policy as defined in Definition 3.1, in that it is not based on the previous information state, but it is a plausible way of combining the information obtained from M_0 and M_1.

For t even, the two-step time evolution of the information state can be calculated by inspection, giving

$$z_{t+2} = \begin{cases} 1 & \text{if } y_{t+2} = 0, \\ r_1(0) = \frac{1}{9} & \text{if } y_{t+2} = 1 \text{ and } y_{t+1} = 1, \\ r_1 \circ r_0(z_t) & \text{if } y_{t+2} = 1 \text{ and } y_{t+1} = 0. \end{cases} \tag{78}$$

We need to calculate the limiting probabilities of each of these cases, while continuing to insist that t be even.

Since Y_{t+2} is a measurement using M_1, and $Y_\infty = (\frac{1}{4}, \frac{3}{4})$ when using M_1, it follows that $\mathbb{P}(Y_{t+2} = 0) \to \frac{1}{4}$ as $t \to \infty$.

Similarly, $\mathbb{P}(Y_{t+1} = 1) \to \frac{1}{4}$ as $t \to \infty$. On the other hand, since it is not possible to observe 1 while in state 0 using M_0, it follows that $Y_{t+1} = 1$ necessarily implies $X_{t+1} = 1$, hence $\mathbb{P}(Y_{t+1} = 1) = \mathbb{P}(Y_{t+1} = 1, X_{t+1} = 1)$. Then,

$$\begin{aligned} \mathbb{P}(Y_{t+1} = 1, Y_{t+2} = 1) &= \sum_{i=0}^{1} \mathbb{P}(Y_{t+1} = 1, X_{t+1} = 1, X_{t+2} = i, Y_{t+2} = 1) \\ &= \sum_{i=0}^{1} \mathbb{P}(Y_{t+1} = 1) T_{1,i} [M_1]_{i,1} \\ &\to \frac{1}{4} \times \frac{1}{5} \times \frac{1}{2} + \frac{1}{4} \times \frac{4}{5} \times 1 \quad \text{as } t \to \infty \\ &= \frac{9}{40}. \end{aligned} \tag{79}$$

We can then use the fact that exactly one of the three cases in Equation (78) must occur to deduce that

$$\mathbb{P}(Y_{t+1} = 0, Y_{t+2} = 1) \to \frac{21}{40} \quad \text{as } t \to \infty. \tag{80}$$

Then, the two step evolution process for the distribution of the information state of this system is given by

$$F^2(Z) = \frac{1}{4}\delta(1) + \frac{9}{40}\delta\left(\tfrac{1}{9}\right) + \frac{21}{40} \int_{[0,1]} \delta\Big(r_1 \circ r_0(z)\Big)\, dZ(z). \tag{81}$$

Again, we can write down the fixed point by inspection, which is

$$\begin{aligned} Z_\infty = \frac{1}{4}\Bigg(&\delta(1) + \frac{21}{40}\delta\big(r_1 \circ r_0(1)\big) + \left(\frac{21}{40}\right)^2 \delta\big([r_1 \circ r_0]^2(0)\big) + \dots\Bigg) \\ &+ \frac{9}{40}\Bigg(\delta(\tfrac{1}{9}) + \frac{21}{40}\delta\big(r_1 \circ r_0(\tfrac{1}{9})\big) + \left(\frac{21}{40}\right)^2 \delta\big([r_1 \circ r_0]^2(\tfrac{1}{9})\big) + \dots\Bigg). \end{aligned} \tag{82}$$

We again numerically approximate the long term expected information entropy, by adding up the first 1000 terms, noting that the error is even smaller in this case, since the geometric series with ratio $\frac{21}{40}$ converges faster than the geometric series with ratio $\frac{3}{4}$. This gives

$$H(Z_\infty) \approx 0.424249. \tag{83}$$

This only calculates limiting entropy when t is even, but by symmetry, the limiting entropy for t odd must be the same. While the sequence of information state distributions Z_t will oscillate between two points, these two points have the same entropy and hence the entropy converges. Technically speaking, this is a fixed point of F^2, and thus a two-cycle of F; we omit the proof that it is unique.

Thus, we have slightly improved upon the long term expected information entropy achieved by using only one observation matrix, and conclude that it is indeed possible to do better with two observation matrices. However, the improvement is quite small, and intuitively, one might suspect that a substantial improvement can be made by choosing between observation matrices in a manner that is not predetermined.

Example 3.8. In our final example, we show that a policy based on the previous information state can do significantly better. Suppose the transition matrix T and the observation matrices $M[0] = M_0$ and $M[1] = M_1$ are the same as in Example 3.7, but instead of alternating between the two, we make a choice based on the information state, using observation matrix M_0 if $z_0 \geq \frac{1}{2}$, and using observation matrix M_1 if $z_0 < \frac{1}{2}$.

By Equation (52), the limiting behaviour of the distribution of the information state is given by the function

$$\begin{aligned}F(Z) &= \sum_{i=0}^{1}\sum_{y=0}^{1}\mathbb{P}(Y_\infty = y)\int_{A_i}\delta\Big(f(z,y)\Big)\,dZ(z)\\ &= \frac{3}{4}\int_{[\frac{1}{2},1]}\delta\Big(r_0(z)\Big)\,dZ(z) + \frac{1}{8}\delta(0) + \frac{1}{8}\delta(1) + \frac{3}{4}\int_{[0,\frac{1}{2})}\delta\Big(r_1(z)\Big)\,dZ(z).\end{aligned} \tag{84}$$

Note that in calculating the expression above, we have used the fact that, by symmetry,

$$\int_{[0,\frac{1}{2})} dZ(z) = \int_{[\frac{1}{2},1]} dZ(z) = \frac{1}{2}. \tag{85}$$

The fixed point of r_0 is

$$r_0\left(\sqrt{\tfrac{2}{3}}\right) = \frac{6\sqrt{\frac{2}{3}}+2}{3\sqrt{\frac{2}{3}}+6} = \frac{2\sqrt{6}+2}{\sqrt{6}+6} = \frac{2}{\sqrt{6}} = \sqrt{\tfrac{2}{3}}, \tag{86}$$

and the derivative of r_0 is

$$r_0'(z) = \frac{d}{dz}\left(\frac{6z+2}{3z+6}\right) = \frac{6\times 6 - 2\times 3}{(3z+6)^2} = \frac{10}{3(z+2)^2} > 0. \tag{87}$$

Since $\sqrt{\frac{2}{3}} > \frac{1}{2}$, and r_0 is monotonically increasing, it follows that

$$1 > r_0(1) > r_0^2(1) > \cdots > \tfrac{1}{2}. \tag{88}$$

By symmetry, we also have

$$0 < r_1(0) < r_1^2(0) < \cdots < \tfrac{1}{2}. \tag{89}$$

Putting these facts together, a fixed point for the process F is given by

$$\begin{aligned}Z_\infty = \frac{1}{8}\Bigg(&\delta(0) + \frac{3}{4}\delta\big(r_1(0)\big) + \left(\frac{3}{4}\right)^2\delta\big(r_1^2(0)\big) + \dots\\ &+ \delta(1) + \frac{3}{4}\delta\big(r_0(1)\big) + \left(\frac{3}{4}\right)^2\delta\big(r_0^2(1)\big) + \dots\Bigg).\end{aligned} \tag{90}$$

This fixed point is unique, which will be proved in Section 4.2. The entropy of this fixed point can be approximated numerically as

$$H(Z_\infty) \approx 0.323979. \tag{91}$$

As before, we have enumerated the first 1000 terms of the series, for which the error is much less than the last significant digit quoted.

We summarise our results in the table below.

Method	Example	Entropy
Best case	3.3	0.00
Worst case	3.4	0.69
M_0 only	3.5	0.44
M_1 only	3.6	0.44
Alternate	3.7	0.42
Policy	3.8	0.32

This shows that a policy can significantly outperform other methods of choosing an observation matrix. Theoretically speaking, an entropy of 0.44 corresponds to a Bernoulli random variable with probability of success approximately $\frac{1}{6}$, while an entropy of 0.32 corresponds to a Bernoulli random variable with probability of success approximately $\frac{1}{10}$. If we estimate the state of the Markov chain by maximum likelihood, these correspond exactly to the probability of an incorrect estimate.

Another argument for this being a significant improvement draws from potential applications. If we are concerned about the precision of state estimates, the information state is almost certainly something we would already have calculated as part of that estimation. In this case, if a choice must be made between one of the observation methods, it takes very little additional work to select one based on the information state, for a significant gain over selecting one on some other basis.

4 A Special Case

4.1 Definitions

We now consider a special case of the general Hidden Markov Model with multiple observation matrices, for which there exists an analytic solution to the problem of the existence of limiting distributions and optimal policies. We make an explicit note that while the previous chapters contain known results and relatively simple extensions thereof, this chapter comprises original work.

The first assumption we make in formulating this special case is that, as in Section 3.3, there are exactly two underlying states, two possible outcomes of observation, and two observation matrices from which to choose. This is the smallest case for which the problem is still interesting, as we can only infer partial knowledge of the underlying process, and it is not obvious how we should pick the observation matrix.

Next, we assume that the Markov chain $\{X_t\}_{t\in\mathbb{Z}^+}$ is well-behaved, by which we mean the associated transition matrix T contains no zeros. In two dimensions, this is equivalent to requiring that there be no deterministic transitions. While a deterministic transition would make the problem easier, we wish to avoid them so as not to have to consider separate cases in our working.

Our final assumption will be the strongest—that each observation matrix observes one state perfectly, in that it returns one observation outcome deterministically when the underlying process is in a particular state. This presents a trade-off between the two observation matrices which, as seen in Example 3.8, allows a policy to be effective. While we cannot expect perfect observations in general, this is often not too bad an approximation, and will allow us to solve the system analytically.

In summary, we assume that $n = m = l = 2$, and

$$T = \begin{bmatrix} a & 1-a \\ 1-b & b \end{bmatrix} \qquad M[0] = \begin{bmatrix} 1 & 0 \\ 1-p & p \end{bmatrix} \qquad M[1] = \begin{bmatrix} q & 1-q \\ 0 & 1 \end{bmatrix},$$

where $a, b, p, q \in (0, 1)$.

As in Section 3.3, we will consider an information state z_t to be the posterior probability that X_t equals 0, rather than the posterior distribution, which will be given by the vector of probabilities $(z_t, 1 - z_t)$.

Hence, we can define our policy $g : [0, 1] \to \{0, 1\}$ by an ordered partition $\{A_0, A_1\}$ of $[0, 1]$, such that g takes value 0 on A_0 and value 1 on A_1. Recalling that a Hidden Markov Model is specified by (T, M, π), and the multiple observation processes are specified by $M[0], \ldots, M[l-1]$ and g, we have completely defined the system except for the initial distribution π. As we will wish to prove our results for any initial value, we will leave this undefined.

Finally, we define a *threshold policy*, as one in which we specify a threshold $c \in [0, 1]$, and use observation matrix $M[0]$ to find y_{t+1} if z_t is one side of this threshold, and use observation matrix $M[1]$ if z_t is on the other side. Such a policy has a very simple form, and we will aim to show that these are the only policies we need to consider. Formally, we can define this as follows.

Definition 4.1. A policy $\{A_0, \ldots, A_{l-1}\}$ is called a **threshold policy** if the sets $A_0, \ldots, A_{l-1}$ are connected. If the underlying state space has dimension two, this is equivalent to their being any of path-connected, star-shaped, or convex. It is arguable which of these is the appropriate generalisation for higher dimensions, but since we are only considering the two dimensional case, that is irrelevant here.

4.2 Stationary Point

We now prove that there is always a stationary point in the process F corresponding to the estimate of this Hidden Markov Model, and that in most cases, including all cases where the policy is a threshold policy, this stationary point is unique and does not depend on the initial distribution Z_0.

First, we calculate the evolution process f for the estimate z_t, given by $f(z_t, y_{t+1}) = z_{t+1}$. Recalling Equation (49), this has form

$$\begin{aligned} f(z,y) &= \frac{M[g(z)]_{0,y}\big(zT_{0,0} + (1-z)T_{1,0}\big)}{M[g(z)]_{0,y}\big(zT_{0,0} + (1-z)T_{1,0}\big) + M[g(z)]_{1,y}\big(zT_{0,1} + (1-z)T_{1,1}\big)} \\ &= \frac{\big(az + (1-b)(1-z)\big)M[g(z)]_{0,y}}{\big(az + (1-b)(1-z)\big)M[g(z)]_{0,y} + \big((1-a)z + b(1-z)\big)M[g(z)]_{1,y}}. \end{aligned} \tag{92}$$

Hence, z_{t+1} is related to z_t by a function f which is piecewise a fractional linear transformation, and in particular, this fractional linear transformation is sometimes constant, as we will see below.

When $y_{t+1} = 1$ and $z_t \in A_0$, the term $M[g(z)]_{0,y}$ becomes $M[0]_{0,1}$, which is zero, so $z_{t+1} = 0$. When $y_{t+1} = 0$ and $z_t \in A_1$, the term $M[g(z)]_{1,y}$ becomes $M[1]_{1,0}$, which is zero, so $z_{t+1} = 1$. Denoting the two other fractional linear transformations by r_0 and r_1, we obtain

$$z_{t+1} = \begin{cases} r_0\big(z_t\big) & \text{if } z_t \in A_0 \text{ and } y_{t+1} = 0, \\ 0 & \text{if } z_t \in A_0 \text{ and } y_{t+1} = 1, \\ 1 & \text{if } z_t \in A_1 \text{ and } y_{t+1} = 0, \\ r_1\big(z_t\big) & \text{if } z_t \in A_1 \text{ and } y_{t+1} = 1. \end{cases} \tag{93}$$

In particular, r_0 and r_1 can be explicitly written down as

$$\begin{aligned} r_0(z) &= \frac{\big(az + (1-b)(1-z)\big)}{\big(az + (1-b)(1-z)\big) + (1-p)\big((1-a)z + b(1-z)\big)} \\ &= \big(1 + (1-p)u(z)\big)^{-1}, \\ r_1(z) &= \frac{(1-q)\big(az + (1-b)(1-z)\big)}{(1-q)\big(az + (1-b)(1-z)\big) + \big((1-a)z + b(1-z)\big)} \\ &= \big(1 + \tfrac{1}{1-q}u(z)\big)^{-1}, \end{aligned} \tag{94}$$

$$\text{where } u(z) = \frac{(1-a)z + b(1-z)}{az + (1-b)(1-z)} = \frac{(\vec{z}\cdot T)_1}{(\vec{z}\cdot T)_0}, \quad \vec{z} = [z, 1-z].$$

Recalling our expression for F from Equation (52), we can replace the integrand with the values of $f(z, y)$ calculated above, to obtain the following.

Lemma 4.1. *The limiting behaviour of the evolution of the information state distribution is given by the function F, where*

$$\begin{aligned} F(Z) = (\pi \cdot T^\infty \cdot M[0])_0 \int_{A_0} \delta\big(r_0(z)\big)\, dZ(z) + (\pi \cdot T^\infty \cdot M[0])_1\, Z(A_0)\, \delta(0) \\ + (\pi \cdot T^\infty \cdot M[1])_0\, Z(A_1)\, \delta(1) + (\pi \cdot T^\infty \cdot M[1])_1 \int_{A_1} \delta\big(r_1(z)\big)\, dZ(z). \end{aligned} \tag{95}$$

Note that by definition, each of the fractional linear transformations r_0 and r_1 are defined on $\Delta = [0, 1]$, and also take values in $[0, 1]$. Then, define the function $r : [0, 1] \to [0, 1]$ by

$$r(z) = r_0(z)\mathbb{I}_{A_0}(z) + r_1(z)\mathbb{I}_{A_1}(z), \tag{96}$$

where $\mathbb{I}_A(x)$ is the indicator function which takes value 1 if $x \in A$ and 0 otherwise. Intuitively, this function determines the next position of a point mass under F by choosing the appropriate fractional linear transformation. We see that a point mass $\delta(z)$ evolves to an affine combination of two point masses, $\delta(r(z))$ and one of $\delta(0)$ or $\delta(1)$.

Let $R \subset [0, 1]$ be the set

$$R = \{0, r(0), r^2(0), \ldots, 1, r(1), r^2(1), \ldots\}.$$

That is, R is the union of the orbits of the two points 0 and 1 under the function r. From the form of the function F, we have the following result.

Definition 4.2. An **absorbing set** is a set $A \subseteq \Delta$ such that for all $z_t \in A$, $z_{t+1} \in A$ with probability 1, where $z_{t+1} = f(z_t, y_{t+1})$ for some y_{t+1}.

Lemma 4.2. *Let Z_0 be a probability measure on Δ, and let the sequence $\{Z_t\}_{t\in\mathbb{Z}^+}$ be generated by $Z_{t+1} = F(Z_t)$. Then, the support of Z_t converges to a subset of the closure of R, or more precisely,*

$$\lim_{t\to\infty} Z_t(\Delta \setminus R) = 0. \tag{97}$$

Proof. By Equation (95), for any point $r^k(0) \in R$, $f(r^k(0))$ is a probabilistic combination of $r^{k+1}(0)$ and either 0 or 1, all of which are in R. Similarly for any point $r^k(1) \in R$, hence R is an absorbing set. This means the mass of Z_t which lies in R must remain in R under F. On the other hand, again by Equation (95), the proportion of the mass of Z_t not in R, which enters R under F, is bounded below by

$$\epsilon = \min\Big\{(\pi \cdot T^\infty \cdot M[0])_1\, Z(A_0), (\pi \cdot T^\infty \cdot M[1])_0\, Z(A_1)\Big\}, \tag{98}$$

which is a positive quantity. This means that

$$\frac{Z_{t+1}(\Delta \setminus R)}{Z_t(\Delta \setminus R)} \leq 1 - \epsilon, \tag{99}$$

for $\epsilon > 0$ given by Equation (98). Hence, the sequence $\{Z_t(\Delta \setminus R)\}_{t\in\mathbb{Z}^+}$ is bounded above by the geometric sequence $\{(1-\epsilon)^t\}_{t\in\mathbb{Z}^+}$, and therefore converges to zero. $\square$

In particular, the limiting behaviour of F is determined by the limiting behaviour of its restriction to the set of probability measures on R, $F|_R$, which can be considered as a Markov chain with state space R.

Define the sets $R_0 = \{0, r(0), r^2(0), \dots\}$ and $R_1 = \{1, r(1), r^2(1), \dots\}$, noting $R_0 \cup R_1 = R$, although R_0 and R_1 are not necessarily disjoint. We now introduce a result from the theory of discrete state Markov chains.

Theorem 4.3. *Consider a Markov chain with discrete state space S. If there is a state $x \in S$ such that the expected time of first entry into x is finite for **any** starting state, then the chain has a unique invariant distribution, and any initial distribution converges to this invariant distribution.*

Proof. The conditions imply that the chain contains exactly one positive-recurrent communication class, with all other classes transient and having positive probability of eventually entering this class. This is sufficient to guarantee the required results. For futher details, see the book by S.P. Meyn and R.L. Tweedie [16]. □

Lemma 4.4. *Suppose R_0 and R_1 are disjoint, $R_0 \subseteq A_0$ and $R_1 \subseteq A_1$. Then, the Markov chain $F|_R$ can be decomposed into two independent chains $F|_{R_0}$ and $F|_{R_1}$, each with a unique invariant distribution, and the invariant distributions of $F|_R$ are affine combinations of those invariant distributions.*

Proof. Note that for any point $z_t = r^k(0) \in R_0 \subseteq A_0$, the next observation matrix used is $M[0]$, hence the next state z_{t+1} is either 0 or $r_0(z_t) = r^{k+1}(0)$. Therefore, a point in R_0 can never leave R_0, and similarly for R_1, so both R_0 and R_1 are absorbing sets. In addition, they partition the state space R, so given any deterministic starting point, the entire Markov chain from that point is either contained entirely in R_0 or contained entirely in R_1. This means the Markov chain $F|_R$ can be decomposed into two independent chains $F|_{R_0}$ and $F|_{R_1}$, with state spaces R_0 and R_1 respectively. It is clear that the invariant distributions of $F|_R$ are affine combinations of the invariant distributions of $F|_{R_0}$ and $F|_{R_1}$, with this combination representing the probabilities that the chain is in R_0 and R_1.

Now consider the process $F|_{R_0}$. For any $z \in R_0$, z makes a transition to 0 with probability $\mathbb{P}(Y_\infty = 1)$, by Equation (93). Since this is a positive, constant probability from any starting state z, it follows that the time of first entry into state 0 is bounded above by a geometric distribution, and thus has finite expected value. Hence, $F|_{R_0}$ has a unique invariant distribution, and similarly for $F|_{R_1}$, by Theorem 4.3. $\square$

Lemma 4.5. *Suppose R_0 and R_1 are disjoint, and at least one of $R_0 \cap A_1$ and $R_1 \cap A_0$ is nonempty. Then, there exists a unique invariant distribution of $F|_R$.*

Proof. Without loss of generality, there is some $k \in \mathbb{Z}^+$ such that $r^k(0) \in A_1$. This means 1 can be reached from 0 in $k+1$ steps with positive probability γ. Furthermore, every point in R enters either 0 or 1 with some probability which is bounded below by $\epsilon > 0$—in fact, ϵ is equal to the value given in Equation (98). In particular, from any starting state, the chain enters state 1 within $k+2$ steps with probability at least $\gamma\epsilon > 0$.

Since k is fixed, the time of first entry into state 1 is bounded above by $k+2$ times a geometric distribution, and has finite expected value. In particular, this means the first entry time into state 1 has finite expected value, and hence, there is a unique invariant distribution, by Theorem 4.3. $\square$

Lemma 4.6. *Suppose R_0 and R_1 are not disjoint. Then, there exists a unique invariant distribution of $F|_R$.*

Proof. Supposed $r^k(0) = r^l(1) \in R_0 \cap R_1$; without loss of generality suppose this point is contained in A_1. But this k is fixed, so by the argument in Lemma 4.5, the first entry time into state 1 has finite expected value, and hence there exists a unique invariant distribution. $\square$

These results together show that there is always a invariant distribution, and this invariant distribution is independent of the prior information state distribution unless $R_0 \subseteq A_0$ and $R_1 \subseteq A_1$. We now show that this is impossible for a threshold policy.

Lemma 4.7. *If $R_0 \subseteq A_0$ and $R_1 \subseteq A_1$, then A_0 and A_1 cannot both be connected.*

Proof. Suppose that $R_0 \subseteq A_0$, $R_1 \subseteq A_1$, and A_0 and A_1 are both connected.

We have seen that in this case, the Markov chain $F|_R$ decomposes into two Markov chains $F|_{R_0}$ and $F|_{R_1}$, each with a unique invariant distribution, which we shall denote by $Z_\infty^{[0]}$ and $Z_\infty^{[1]}$ respectively. Note that these invariant distributions are probability measures on R_0 and R_1 respectively. In particular, $Z_\infty^{[0]}$ is the limiting information state distribution when the prior information state distribution $Z_0^{[0]}$ is a probability measure on R_0.

On the other hand, any information state distribution satisfies

$$
\begin{aligned}
\mathbb{E}[Z_t] &\equiv \int_\Delta z dZ_t(z) \\
&= \sum_{y^{(t)}} \mathbb{P}\big(Y^{(t)} = y^{(t)}\big) z_t\big[y^{(t)}\big] \\
&= \sum_{y^{(t)}} \mathbb{P}\big(Y^{(t)} = y^{(t)}\big) \mathbb{P}\big(X_t = 0 \big| Y^{(t)} = y^{(t)}\big) \\
&= \mathbb{P}(X_t = 0) \to \frac{1-b}{1-a+1-b} \in (0,1) \quad \text{as } t \to \infty.
\end{aligned} \tag{100}
$$

Hence,

$$
\mathbb{E}\big[Z_\infty^{[0]}\big] = \frac{1-b}{1-a+1-b}. \tag{101}
$$

Since $Z_\infty^{[0]}$ is a distribution on R_0, $\frac{1-b}{1-a+1-b}$ lies in the convex hull of R_0, which is contained in A_0 since A_0 contains R_0 and is connected and therefore convex. Hence, $\frac{1-b}{1-a+1-b} \in A_0$.

On the other hand, since Equation (100) does not depend on the prior information state distribution or the observation matrix, it follows that the same argument would hold for $Z_\infty^{[1]}$, and so, $\frac{1-b}{1-a+1-b} \in A_1$. This is a contradiction, which means that if $R \subseteq A_0$ and $R_1 \subseteq A_1$, then A_0 and A_1 cannot both be connected. $\square$

We summarise our results in the following theorem.

Theorem 4.8. *Consider a Hidden Markov Model with the transition matrix and the two observation matrices given by*

$$T = \begin{bmatrix} a & 1-a \\ 1-b & b \end{bmatrix} \qquad M[0] = \begin{bmatrix} 1 & 0 \\ 1-p & p \end{bmatrix} \qquad M[1] = \begin{bmatrix} q & 1-q \\ 0 & 1 \end{bmatrix},$$

where $a, b, p, q \in (0, 1)$.

For any policy, the process F representing the limiting behaviour of the information state distribution has a invariant distribution. If the policy is a threshold policy, this invariant distribution is unique and does not depend on the prior information state distribution. In either case, the set of invariant distributions is a global attractor of F considered as a dynamical system.

Proof. The existence and uniqueness of an invariant distribution follow from combining the previous lemmata. The invariant distribution being a global attractor follows from Lemma 4.2, which shows that the limiting distribution is a distribution on R, and Theorem 4.3, which shows that any distribution on R converges to an invariant distribution. $\square$

Corollary 4.9. *With the conditions of Theorem 4.8, the set of invariant measures corresponding to a policy is a global attractor of the time-dependent process F_t.*

Proof. Follows immediately from Lemma 3.3. $\square$

4.3 Optimality of Threshold Policies

We have shown that for every threshold policy, there is a unique invariant distribution which is a global attractor of the evolution process for the information state distribution. We will now prove that in certain cases, it is also optimal, in the sense that it minimises the long term expected information entropy

$$H(Z_\infty) = -\int_{[0,1]} z \log z + (1-z)\log(1-z)\, dZ_\infty(z), \tag{102}$$

where Z_∞ is the limiting information state distribution.

Recall that Z_∞ is a probability distribution on $R \subset \Delta$, where $R = R_0 \cup R_1$, $R_0 = \{0, r(0), r^2(0), \dots\}$ and $R_1 = \{1, r(1), r^2(1), \dots\}$.

It is often convenient to consider a policy not as a partition $\{A_0, A_1\}$ of Δ, but instead as a partition $\{B_0, B_1\}$ of R. Then, for each $r^k(0) \in R$, $r^{k+1}(0) = r_0(r^k(0))$ if $r^k(0) \in B_0$, and $r^{k+1}(0) = r_1(r^k(0))$ if $r^k(0) \in B_1$, and similarly for points $r^l(1) \in R$. In order for $\{B_0, B_1\}$ to be a valid policy, it suffices for the points in R to be disjoint in $[0,1]$.

We now consider the conditions under which an optimal policy is a threshold policy. In particular, we require that the underlying process is more likely to stay in the same state than oscillate between the two states. This will be a valid assumption in most cases, since a system which alternates—and more specifically, alternates at exactly the same frequency with which observations are made—is usually not a sensible model.

In the cases where this alternation does occur, we can transform the system into a non-oscillating chain by defining the stochastic process $\bar{X}_t$ which is equal to X_t if t is even and $1 - X_t$ if t is odd. Note that for the transformed process, the observation matrices $M[0]$ and $M[1]$ also alternate, so unless $p = q$, this transformed system no longer technically fits the model we have been working with.

For now, we define this concept more precisely.

Definition 4.3. A **positive autocorrelation** Markov chain is one in which the limiting one-step autocorrelation ρ is positive, where

$$\begin{aligned}
\rho &= \lim_{t\to\infty} \frac{\mathbb{E}[X_t X_{t+1}] - \mathbb{E}[X_t]\mathbb{E}[X_{t+1}]}{\mathbb{E}[X_t^2] - \mathbb{E}[X_t]^2} \\
&= \lim_{t\to\infty} \frac{\mathbb{P}(X_t = 1, X_{t+1} = 1) - \mathbb{P}(X_t = 1)\mathbb{P}(X_{t+1} = 1)}{\mathbb{P}(X_t = 1) - \mathbb{P}(X_t = 1)^2} \\
&= \frac{\frac{1-a}{1-a+1-b} \times b - \left(\frac{1-a}{1-a+1-b}\right)^2}{\frac{1-a}{1-a+1-b} - \left(\frac{1-a}{1-a+1-b}\right)^2} \quad \text{since } \Pr(X_\infty = 1) = \tfrac{1-a}{1-a+1-b} \\
&= \frac{(1-a)b(1-a+1-b) - (1-a)^2}{(1-a)(1-a+1-b) - (1-a)^2} \\
&= \frac{(1-a)(1-b)(a+b-1)}{(1-a)(1-b)} \\
&= a+b-1. \qquad (103)
\end{aligned}$$

Lemma 4.10. *The functions r_0 and r_1 are monotonic. They are strictly increasing if the Markov chain has positive autocorrelation, strictly decreasing if the Markov chain has negative autocorrelation, and constant if the Markov chain has zero autocorrelation.*

Proof. Recall Equation (94), which gives $r_0(z) = (1 + (1-p)u(z))^{-1}$ and $r_1(z) = (1 + \frac{1}{1-q}u(z))^{-1}$, where

$$u(z) = \frac{(1-a)z + b(1-z)}{az + (1-b)(1-z)} = \frac{(1-a-b)z + b}{(a+b-1)z + (1-b)}.$$

Noting that

$$u'(z) = \frac{1-a-b}{\big((a+b-1)z + (1-b)\big)^2}, \qquad (104)$$

we see that

$$r_0'(z) = (1 + (1-p)u(z))^{-2} \times \frac{a+b-1}{\big((a+b-1)z + (1-b)\big)^2}, \qquad (105)$$

which is positive for $\rho = a+b-1 > 0$, negative for $\rho < 0$, and zero for $\rho = 0$. Since r_1 has the same form as r_0 with the $1-p$ term replaced with $\frac{1}{1-q}$, this also holds for r_1. $\square$

Lemma 4.11. *The fractional linear transformations r_0 and r_1 satisfy the inequality $r_0(z) > r_1(z)$ for all $z \in [0, 1]$.*

Proof. Again from Equation (94), we have $r_0(z) = (1 + (1 - p)u(z))^{-1}$ and $r_1(z) = (1 + \frac{1}{1-q}u(z))^{-1}$. In particular, $u(z)$ is a ratio of probabilities and is always positive, and $1-p < 1 < \frac{1}{1-q}$, hence $r_0(z) > r_1(z)$ for all $z \in [0, 1]$. $\square$

Lemma 4.12. *The fractional linear transformations r_0 and r_1 have fixed points μ_0 and μ_1 respectively, which are contained in $(0, 1)$ and satisfy the inequality $\mu_0 > \mu_1$.*

Proof. Since r_0 maps $[0, 1]$ to $(0, 1)$, the function $r_0(z) - z$ is positive when $z = 0$ and negative when $z = 1$. It is also continuous since r_0 is continuous, so by the intermediate value theorem, it is zero somewhere in $(0, 1)$, which gives a fixed point of r_0. Similarly, r_1 also has a fixed point in $(0, 1)$.

If the chain has positive autocorrelation, then $\mu_0 > \mu_1$ follows from Lemma 4.10 and Lemma 4.11. If the chain has negative or zero autocorrelation, suppose that $\mu_0 \leq \mu_1$; then $\mu_0 = r_0(\mu_0) > r_1(\mu_0) \geq r_1(\mu_1) = \mu_1$ by Lemma 4.10, but this is a contradiction, so again we have $\mu_0 > \mu_1$. $\square$

Lemma 4.13. *Without loss of generality, we can assume that*

$$1 - \mu_0 \leq \mu_1. \tag{106}$$

Proof. Suppose that $1 - \mu_0 > \mu_1$. Now let us swap the labels 0 and 1 in both the state space and the observation space, as well as in the indices of the observation matrices. Then, an information state $z_t = \Pr(X_t = 0|Y^{(t)} = y^{(t)})$ is equal to $\Pr(\bar{X}_t = 1|\bar{Y}^{(t)} = \bar{y}^{(t)}) = 1 - \bar{z}_t$, where the bar denotes a relabelled quantity. Since the indices of the observation matrices are also swapped, we have

$$1 - \bar{\mu}_0 = \mu_1 < 1 - \mu_0 = \bar{\mu}_1, \tag{107}$$

which gives us the required inequality under this relabelling. $\square$

Lemma 4.14. *The information entropy of a point is proportional to its distance to the edge of the interval* $[0, 1]$*, that is,* $H(x) < H(y)$ *if and only if*

$$\min\{x, 1-x\} < \min\{y, 1-y\}. \tag{108}$$

Proof. This is clear from the fact that information entropy is strictly concave and symmetric about $\frac{1}{2}$. □

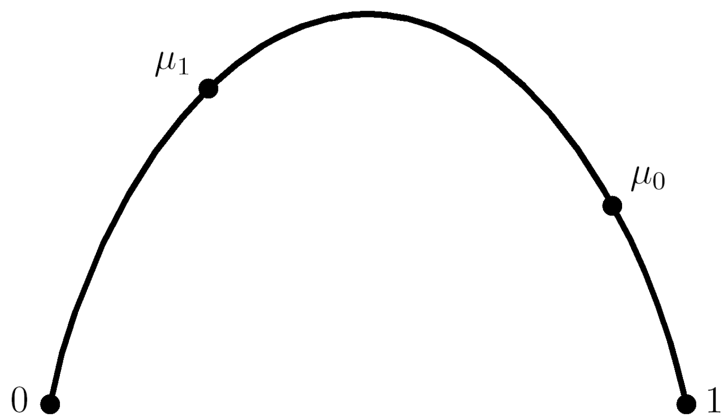

Figure 1: Graph of information entropy at several points. By Lemma 4.13, we can assume that μ_0 is closer to the boundary of $[0, 1]$ than μ_1, without loss of generality. By Lemma 4.14, this means $H(\mu_0) < H(\mu_1)$.

Lemma 4.15. *Suppose the Markov chain has positive autocorrelation, and without loss of generality, suppose* $1 - \mu_0 \leq \mu_1$*. Let* Z_∞ *be an invariant distribution given by the policy* $\{B_0, B_1\}$ *on* R*, and* $Z_\infty|_{R_1}$ *the restriction of this distribution to* R_1*. Then, the quantity*

$$H(Z_\infty|_{R_1}) = \int_{R_1} H(z)\, dZ_\infty(z) \tag{109}$$

is minimised when $R_1 \subseteq B_0$*.*

Proof. By Lemma 4.10 and Lemma 4.11, if $x \geq y$, then $r_0(x) \geq r(y)$ where r is either r_0 or r_1. Then, it is clear by induction that $r_0^k(1) \geq r^k(1)$, for all r^k a composition of k fractional linear transformations r_0 or r_1. Hence, $1 - r_0^k(1) \leq 1 - r^k(1)$.

On the other hand, r_0 and r_1 are monotonic by Lemma 4.10, so $r_0^k(1) \geq \mu_0$, and $r^k(1) \geq \min\{\mu_0, \mu_1\} = \mu_1$. Hence, $1 - r_0^k(1) \leq 1 - \mu_0 \leq \mu_1 \leq r^k(1)$.

Putting these together gives $1 - r_0^k(1) \leq \min\{r^k(1), 1 - r^k(1)\}$. Then, by Lemma 4.14, $H(r_0^k(1)) < H(r^k(1))$, that is, the entropy of at individual point $r^k(1) \in R_1$ is minimised when $r^k(1) = r_0^k(1)$.

Thus, if $R_1 \subseteq B_0$, that is, $r = r_0$ everywhere in R_1, then the entropy at each point in R_1 is individually minimised, hence the overall entropy of the limiting distribution restricted to R_1, $H(Z_\infty|_{R_1})$, is minimised. □

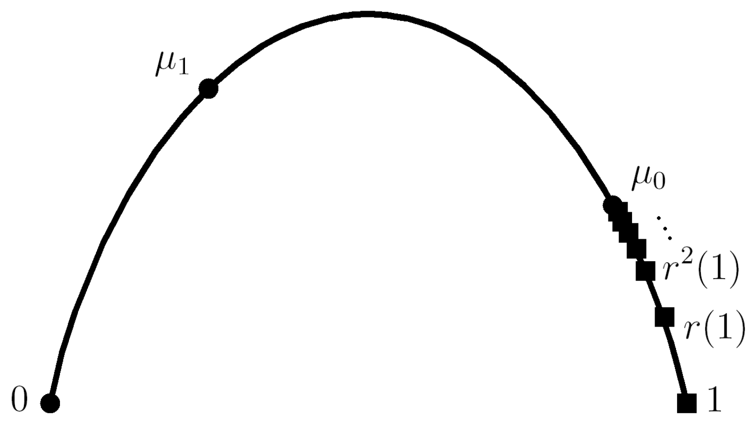

Figure 2: Graph of optimal information entropy on R_1 for a positive autocorrelation chain. By Lemma 4.15, each point approaches μ_0 rather than μ_1. Graphically, this clearly minimises the entropy.

Lemma 4.16. *Suppose the Markov chain has positive autocorrelation, and without loss of generality, suppose $1 - \mu_0 \leq \mu_1$. Let Z_∞ be an invariant distribution given by the policy $\{B_0, B_1\}$ on R, and $Z_\infty|_{R_0}$ the restriction of this distribution to R_0. Then, the quantity*

$$H(Z_\infty|_{R_0}) = \int_{R_0} H(z)\, dZ_\infty(z) \tag{110}$$

is minimised when, for some $k \in \mathbb{Z}^+ \cup \{\infty\}$, $\{0, r(0), r^2(0), \dots, r^{k-1}(0)\} \subseteq B_1$ and $\{r^k(0), r^{k+1}(0), \dots\} \subseteq B_0$.

Proof. Suppose some fixed partition $\{B_0, B_1\}$ minimises $H(Z_\infty|_{R_0})$. Let k be the first element in the orbit of 0 which is in B_0, that is,

$$k = \min\left\{ j \in \mathbb{Z}^+ : r^j(0) \in A_0 \right\}. \tag{111}$$

Then by definition, $\{0, r(0), \dots, r^{k-1}(0)\} \subseteq B_1$. If $k = \infty$, there is nothing left to prove, hence it suffices to show that $\{r^k(0), r^{k+1}(0), \dots\} \subseteq A_0$ in the case when $k < \infty$.

We claim that after $r^k(0)$, only $M[0]$ can be used. Let $c = r^k(0)$. Since $c \in B_0$ and the policy $\{B_0, B_1\}$ is optimal, the entropy is lower when we use r_0 at c than when we use r_1 at c. Noting that $r_0(c) > r_1(c)$, we see that in particular, the entropy is lowered by shifting everything after c to the right. Since entropy is unimodal, it follows that entropy is decreased by shifting the tail to the right for every point after c. In particular, since $r_0(x) > r_1(x)$, the entropy is reduced by using $M[0]$ rather than $M[1]$ at every point $x > c$.

Since $r^{k+1}(0) = r_0(c) > c$, this means we should use $M[0]$ for $r^{k+1}(0)$, that is, $r^{k+1}(0) \in B_0$. But then we can continue by induction to see that the entire sequence $\{r^k(0), r^{k+1}(0), \dots\}$ is contained in B_0. □

Remark. It is acknowledged that the proof of Lemma 4.16 is rather imprecise. While all evidence points to the statement being true, a rigorous proof remains elusive. A graphical depiction of this result is presented in Figure 3.

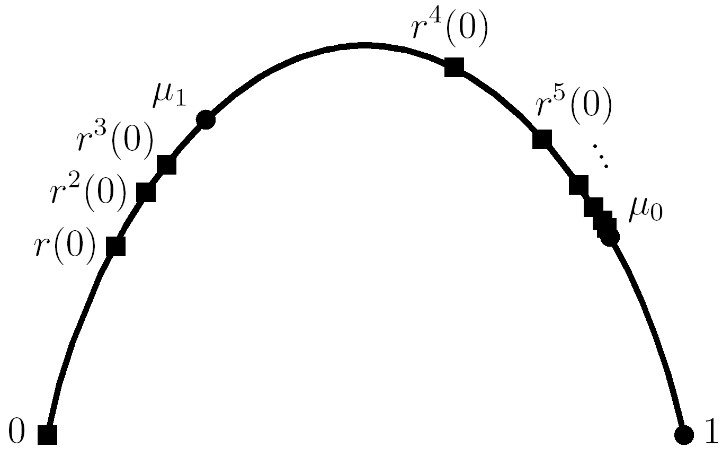

Figure 3: Graph of optimal information entropy on R_0 for a positive autocorrelation chain. By Lemma 4.16, each point approaches μ_1 until $r^k(0)$; from then onwards, they approach μ_0. The diagram depicts the case $k = 3$; $r^4(0)$ is the first point to approach μ_0. If any point after $r^4(0)$ were to approach μ_1, the resulting entropy would be higher.

Theorem 4.17. *For a Markov chain with positive autocorrelation, there exists an optimal threshold policy.*

Proof. By Lemma 4.15 and Lemma 4.16, there is some partition $\{B_0, B_1\}$ of R which simultaneously minimises $H(Z_\infty|_{R_0})$ and $H(Z_\infty|_{R_1})$, given by

$$\begin{aligned} B_0 &= \{r^k(0), r^{k+1}(0), \dots\} \cup \{1, r(1), r^2(1), \dots\}, \\ B_1 &= \{0, r(0), r^2(0), \dots, r^{k-1}(0)\}. \end{aligned} \tag{112}$$

We can explicitly evaluate the function r as either r_0 or r_1, to obtain

$$\begin{aligned} B_0 &= \{r_1^k(0), r_0(r_1^{k+1}(0)), r_0^2(r_1^{k+1}(0)), \dots\} \cup \{1, r_0(1), r_0^2(1), \dots\}, \\ B_1 &= \{0, r_1(0), r_1^2(0), \dots, r_1^{k-1}(0)\}. \end{aligned} \tag{113}$$

By inspection, this partition is equivalent to the threshold policy given by $A_0 = [r_1^k(0), 1]$ and $A_1 = [0, r_1^k(0))$. Furthermore, this policy simultaneously minimises $H(Z_\infty|_{R_0})$ and $H(Z_\infty|_{R_1})$, and hence minimises $H(Z_\infty)$. $\square$

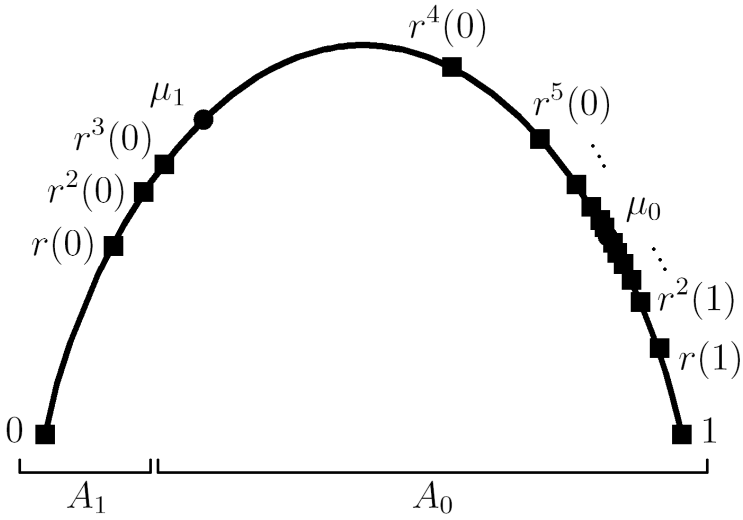

Figure 4: Graph of optimal information entropy on R for a positive autocorrelation chain. The threshold policy $\{A_0, A_1\}$ is shown.

While we have shown an optimal policy in the positive autocorrelation case is a threshold policy of the form $A_0 = [c, 1]$ and $A_1 = [0, c)$, we have not described a method of finding the threshold c. However, we have already made the problem much easier—at worst, we could simulate the long term expected entropy for a large number of values of c to determine one that is approximately optimal. We could also use a greedy approach, setting

$$c = \inf \Big\{ x \in [0, 1] : (\pi T^\infty M[0])_1 H\big(r_0(x)\big) < (\pi T^\infty M[1])_0 H\big(r_1(x)\big) \Big\}, \quad (114)$$

which calculates the position of the first point mass for which using $M[0]$ results in a lower information entropy at the next step than using $M[1]$.

Thus, while we have not specified how to find the optimal policy in the positive autocorrelation case, we have essentially solved the problem by restricting the form of the optimal policy sufficiently so that it is easily determined. We now briefly discuss the cases where autocorrelation is zero or negative.

In a two state Markov chain, zero autocorrelation is equivalent to independence, since $\rho = 0$ means $a = 1 - b$, so the rows of the transition matrix are identical. This is clearly an uninteresting scenario. However, we note that the optimal policy is still a threshold policy, albeit the trivial one in which one always uses the same observation matrix, as seen in Example 3.2.

The negative autocorrelation case is not as easy to understand. In this case, instead of monotonically approaching their fixed point, iterations of r_0 or r_1 converge to their fixed point by oscillation. This can be seen in Figure 5.

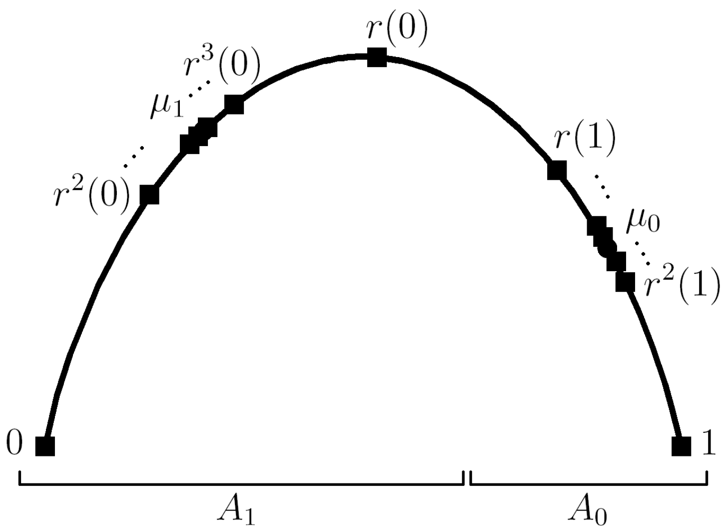

Figure 5: Graph of information entropy on R for a negative autocorrelation model, with a threshold policy $\{A_0, A_1\}$.

As discussed previously, it is possible to transform a negative autocorrelation chain into a positive autocorrelation chain by considering the sequence of random variables $\bar{X}_t$, equal to X_t if t is even and $1 - X_t$ if t is odd. This does not fit our model because the observation matrices $M[0]$ and $M[1]$ also alternate, hence the observation process is no longer stationary. However, since this time dependence is a very simple one, this may possibly lead to a solution using similar methods.

5 Conclusion

We considered situations in which the state of a system is described by a Markov chain which is not directly observable, but is instead only observable with error via an associated observation process. In the case where there is only one observation process, or there are multiple observation processes but we are able to use all of them simultaneously, the problem is well-understood, and we described some of the basic theory involved in analysing such systems.

We then considered the case where there are multiple observation processes, but they cannot all be used simultaneously, which is an interesting but mostly unexplored problem. In particular, we considered the state estimation problem, and showed that the classical forward-backward algorithm was readily adaptable to this situation, demonstrating with several numerical examples. We then restricted the problem to the simplest interesting case, and showed that in this case, the state estimation problem has an optimal solution which takes a very simple form.

The general problem, that of Hidden Markov Models with multiple observation matrices, is one which is potentially very useful to understand. Further work in this area would involve fully solving the state estimation problem in the general two dimensional case, from which it should not be difficult to generalise to higher dimensions. The parameter estimation problem is likely to be much harder; a solution may not exist beyond special cases and approximations.

References

[1] G.P. Basharin, A.N. Langville, and V.A. Naumov. The life and work of A.A. Markov. *Linear Algebra and its Applications*, 386:3–26, 2004.

[2] L.E. Baum and J.A. Eagon. An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. *Bulletin of the American Mathematical Society*, 73:360–363, 1967.

[3] L.E. Baum and T. Pétrie. Statistical inference for probabilistic functions of finite state Markov chains. *Annals of Mathematical Statistics*, 37(6):1554–1563, 1966.

[4] L.E. Baum, T. Pétrie, G. Soules, and N. Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Annals of Mathematical Statistics*, 41(1):164–171, 1970.

[5] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley & Sons, New York, 1991. ISBN: 0471062596.

[6] R. Durbin, S. Eddy, A. Krogh, and G. Mitchison. *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. Cambridge University Press, 1998. ISBN: 0521629713.

[7] R.J. Elliott, L. Aggoun, and J.B. Moore. *Hidden Markov Models: Estimation and Control*. Springer-Verlag, New York, 1995. ISBN: 0387943641.

[8] J.S. Evans and V. Krishnamurthy. Optimal sensor scheduling for hidden Markov models. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2161–2164 (Volume 4), Seattle, 1998.

[9] J.S. Evans and V. Krishnamurthy. Optimal sensor scheduling for hidden Markov model state estimation. *International Journal of Control*, 74(18):1737–1742, 2001.

[10] G.D. Forney Jr. The Viterbi algorithm. *Proceedings of the IEEE*, 61(3):268–278, 1973.

[11] C. Karlof and D. Wagner. *Hidden Markov Model Cryptanalysis*. Springer, Berlin, 2003. ISBN: 9783540408338.

[12] V. Krishnamurthy. Algorithms for optimal scheduling and management of hidden Markov model sensors. *IEEE Transactions on Signals Processing*, 50(6):1382–1397, 2002.

[13] N. Levanon. *Radar Principles*. John Wiley & Sons, Toronto, 1988. ISBN: 0471858811.

[14] W.S. Lovejoy. A survey of algorithmic methods for partially observed Markov decision processes. *Annals of Operations Research*, 28(1):47–66, 1991.

[15] A.A. Markov. Classical text in translation: An example of statistical investigation of the text *Eugene Onegin* concerning the connection of samples in chains. *Science in Context*, 19(4):591–600, 2006. Translated from a lecture given in Russian in 1913.

[16] S.P. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability*. Springer-Verlag, London, 1993. ISBN: 0387198326.

[17] G.E. Monahan. A survey of partially observable Markov decision processes: Theory, models, and algorithms. *Management Science*, 28(1):1–16, 1982.

[18] L.R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.

[19] C. Raphael. Automatic segmentation of acoustic musical signals using hidden Markov models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(4):360–370, 1999.

[20] M. Rezaeian. Sensor scheduling for optimal observability using estimation entropy. In *IEEE International Conference on Pervasive Computing and Communications Workshops*, pages 307–312, New York, 2007.

[21] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer-Verlag, New York, 1981. ISBN: 0387905987.

[22] E. Seneta. Markov and the birth of chain dependence theory. *International Statistical Review*, 64(3):255–263, 1996.

[23] A.J. Viterbi. Error bounds for convolution codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, 13(2):260–269, 1967.