

Remarks on Filtering Error due to Quantisation of a 2-state Hidden Markov Model¹

Louis Shue², Subhrakanti Dey, Brian D. O. Anderson, Franky De Bruyne

Dept of Systems Engineering, RSISE, Australian National University, Canberra ACT 0200, Australia

Abstract

This paper considers the problem of quantisation from the perspective of minimising filtering error when quantised instead of continuous measurements are used as inputs to the nonlinear filter, specialising to a discrete-time 2-state hidden Markov model (HMM) with continuous output.

1 Introduction

In communication systems, a fundamental problem is the question of how much degradation in performance a receiver suffers when incoming data are quantised. In this brief note we will consider the effects of quantisation on the performance of a nonlinear filter, with the signal source assuming the form $Y_k = h(X_k) + v_k$. Here X_k is a 2-state Markov chain, with $h(X_k) \in \{-1, 1\}$ and v_k is independent and identically distributed (i.i.d.) noise. The function of the filter is to estimate X_k given quantised versions of Y_k .

Much of previous work dealing with quantisation aims to reduce the distortion between the original and quantised signals under some minimum mean-squared-error criteria [2, 3, 6]. In contrast, the primary focus here is to minimise the probability of error of a nonlinear filter with quantised observations as inputs. Some of the questions we address are: (1) Is there an optimal strategy for choosing the quantisation levels which minimises the filtering error? (2) How does filtering error vary with noise for different number of quantisation levels M ?

At this stage we have only partial answers to both questions. For the first question, we provide two quantisation schemes, based on the maximisation of the mutual information and Kullback-Leibler (KL) divergence measures respectively. A comparison of these schemes will be made by simulations. We also provide an analysis for computing the maximum *a posteriori* (MAP) error as a function of signal-to-noise ratio (SNR), for a 2-state discrete-time HMM with continuous-range output (this is a discrete-time version of a continuous-time result obtained by Wonham [5]).

2 Signal Model

Consider a first order, homogeneous discrete-time 2-state Markov chain $X_k \in \{1, 2\}$, the subscript k denoting time.

¹The authors wish to acknowledge the funding of the activities of the Cooperative Research Centre for Robust and Adaptive Systems by the Australian Commonwealth Government under the Cooperative Research Centres Program.

²Email: Louis.Shue@anu.edu.au

The actual state levels are denoted as $h(X_k) \in \{-1, 1\}$. At each k , a corresponding signal $Y_k \in \mathbb{R}$ is observed. The associated quantised observations obtained via an M -level quantiser will be denoted as $Y_k^q \in \{1, 2, \dots, M\}$. The recursive filtering equations using Y_k^q can be found in e.g. [4].

The evolution of the Markov chain is governed by the state transition probability matrix $A = (a_{ij})$, where $a_{ij} = \Pr(X_{k+1} = i | X_k = j)$. We will further assume that A is symmetric (or doubly stochastic) and has the form $A = \begin{bmatrix} 1-\alpha & \alpha \\ \alpha & 1-\alpha \end{bmatrix}$ where $0 < \alpha < 1$. The observation process Y_k is defined as $y_k = h(x_k) + v_k$ where v_k is i.i.d. and $\mathcal{N}(0, \sigma^2)$. Consequently, Y_k is conditionally distributed as $\mathcal{N}(x_k, \sigma^2)$, given the state x_k at time k .

The objective of quantisation is to partition the observations Y_k into M distinct levels. The quantised outputs are then related to the Markov states by the measurement matrix $C = (c_{mi})$, such that $c_{mi} = \Pr(Y_k^q = m | X_k = i)$:

$$c_{mi} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\ell_{m-1}}^{\ell_m} \exp\left[-\frac{(y-h(i))^2}{2\sigma^2}\right] dy \quad (1)$$

where ℓ_{m-1} and ℓ_m are the limits of the m -th quantisation level, with the terminating levels $\ell_0 = -\infty$ and $\ell_M = \infty$.

3 Filtering Performance

3.1 Continuous Measurements

In this section we will derive the filtering error for a discrete-time and 2-state HMM with continuous-range observations as a function of SNR by considering the joint process $\{X_k, q_k\}$, where $q_k = \Pi_{k|k}(1) - \Pi_{k|k}(2)$, and $\Pi_{k|k} = \begin{bmatrix} \Pi_{k|k}(1) \\ \Pi_{k|k}(2) \end{bmatrix} = \begin{bmatrix} \Pr(X_k = 1 | Y_0, Y_1, \dots, Y_k) \\ \Pr(X_k = 2 | Y_0, Y_1, \dots, Y_k) \end{bmatrix}$ is the filtered probability vector at time k . Denoting the stationary densities¹ as $\pi^\pm(q) dq = \Pr(h(x_k) = \pm 1, q_k \in (q, q + dq))$, $-1 \leq q \leq 1$, we have (since $h(x_k) = -\text{sign}(q_k)$)

$$\Pr(\text{Filter Error}) = \int_0^1 \pi^+(q) dq + \int_{-1}^0 \pi^-(q) dq. \quad (2)$$

It can be shown that the densities $\pi^\pm(q)$ are the steady-state solutions to the following integral equation:

$$\int_{X,Q}^{k+1} (X_{k+1} = i, q) dq = \Pr(X_{k+1} = i, q_{k+1} \in (q, q + dq))$$

¹We omit here the proof of the existence and uniqueness of solution due to page constraints

$$= \sum_{j=1}^2 a_{ij} \left[\int_{-1}^1 S_i(q, \bar{q}) f_{X, Q}^k(X_k = j, \bar{q}) d\bar{q} \right] dq. \quad (3)$$

where

$$S_i(q, \bar{q}) dq = \Pr \{q_{k+1} \in (q, q + dq) | X_{k+1} = i, q_k = \bar{q}\}.$$

In this paper, we will solve (3) numerically.

3.2 Filtering Using Quantised Observations

We will now propose schemes for determining the quantisation levels using information-theoretic ideas; uniform quantisation will be used as the basis of comparisons.

For random variables X and Y , we have (see [1]):

Definition 3.1 (Entropy and Conditional Entropy)

$$H(X) = -\sum_x P(x) \log P(x)$$

$$H(X|Y) = -\sum_{x,y} P(x,y) \log P(x|y).$$

Definition 3.2 (Mutual Information)

$$I(X; Y) = \sum_{x,y} P(x,y) \log \frac{P(x,y)}{P(x)P(y)} = H(X) - H(X|Y).$$

Definition 3.3 (KL divergence)

$$K(X_k = i, X_k = j) = \sum_k c_{y_k^i} \log \frac{c_{y_k^i}}{c_{y_k^j}} \triangleq E_i \log \frac{c_{y_k^i}}{c_{y_k^j}}.$$

Method 1. Maximisation of Mutual Information The quantisation levels are evaluated by maximisation of $H(X_k; Y_0^q, \dots, Y_k^q)$ via a minimisation of $H(X_k | Y_0^q, \dots, Y_k^q)$, or an approximate version, namely $H(X_k | Y_k^q)$.

Method 2. Maximisation of KL divergence We will 'separate' the distributions $C_{\bullet 1} = \Pr(Y_k^q | X_k = 1)$ and $C_{\bullet 2}$ so that the observations are as informative as possible, by maximising the KL divergence between $C_{\bullet 1}$ and $C_{\bullet 2}$.

4 Simulation Studies

Fig. 1 compares the theoretical error curve (2) and simulation error (continuous measurements). There is high fit at $\sigma \geq 0.7$; the discrepancy occurs because at small σ the theoretical error computed is moderately sensitive to the number of discretising steps used.

Fig 2 shows a comparison of different quantisation schemes. From our simulations, one can conclude that maximisation of mutual information leads to significant improvement over the uniform case, particularly when $\sigma \leq 1$. However, KL-maximisation did not lead to better filter performance.

From Fig. 3, it can be seen that when $M = 4$, the resulting filtering error (when maximising mutual information) is comparable to the theoretical curve irrespective of the SNR.

5 Conclusion

In this paper, we have derived an expression for the filtering error for a 2-state HMM with continuous output. Our simulation studies indicate that maximisation of mutual information can lead to significant improvement in filtering performance over uniform quantisation. However, it seems that KL-maximisation does not in general lead to better filter performance.

References

- [1] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications. John Wiley & Sons, Inc., 1991.
- [2] S. P. Lloyd. Least squares quantisation in PCM. *IEEE Transactions on Information Theory*, IT-28(2):129–137, Mar. 1982.
- [3] J. Max. Quantising for minimum distortion. *IRE Transactions on Information Theory*, IT-6:7–12, Mar. 1960.
- [4] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–285, Feb. 1989.
- [5] W. M. Wonham. Some applications of stochastic differential equations to optimal nonlinear filtering. *SIAM Journal on Control*, 2(3):347–369, 1965. Series A.
- [6] J. Ziv. On universal quantisation. *IEEE Transactions on Information Theory*, IT-31(3):344–347, May 1985.

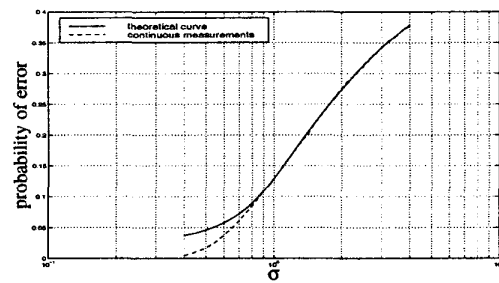


Figure 1: Comparison of theoretical and simulated filtered error (using continuous measurements).

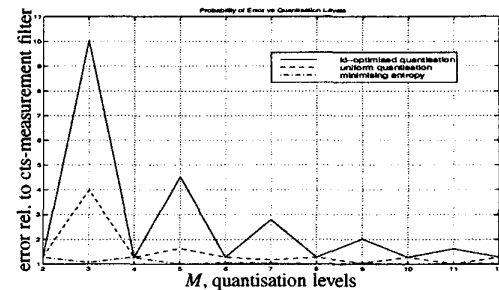


Figure 2: Comparison of filtering error for different quantisation schemes, $\sigma = 0.4$.

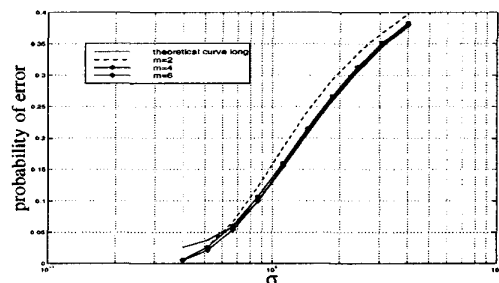


Figure 3: Filtering error using maximisation of mutual information for different number of quantisation levels.