

# Report: Synthesis of optimal filter for MHOQ

June 14, 2024

## Contents

<b>1</b>	<b>Quantisation</b>	<b>1</b>
<b>2</b>	<b>Noise shaping quantiser</b>	<b>1</b>
<b>3</b>	<b>Moving horizon optimal quantiser (MHOQ)</b>	<b>2</b>
<b>4</b>	<b>Frequency response: LPF and NTF</b>	<b>2</b>
4.1	Perception filter . . . . .	2
<b>5</b>	<b>Noise Transfer Function(NTF)</b>	<b>3</b>
<b>6</b>	<b>Spectrum of quantisation noise</b>	<b>4</b>
<b>7</b>	<b>Synthesis of optimal noise-shaping filter</b>	<b>5</b>
7.1	Solution 1: Optimal noise shaping filter without constraint . . . . .	5
7.2	Solution 2: Optimal noise shaping filter with constraint . . . . .	6
7.2.1	System . . . . .	6
7.2.2	Constraint 1 . . . . .	6
7.2.3	Constraint 2 . . . . .	6
7.2.4	Optimization problem: State Space formulation . . . . .	7

## 1 Quantisation

Let  $w \in \mathbb{R}$  be the input,  $\mathbf{Q}$  be the quantiser and  $y \in \mathbb{U}$  the quantiser output. The quantiser output Let us define the quantisation error as

$$q = \mathbf{Q}(w) - w = y - w. \quad (1)$$

The quantisation requires the signal to be mapped to a finite signal where each value of the output  $y$  is restricted to belong to a finite set  $\mathbb{U}$ . The elements of the set  $\mathbb{U}$  represent the quantiser levels and depends on the word-size of the quantiser.

## 2 Noise shaping quantiser

Noise-shaping quantisers can reduce the effective quantisation error by moving quantisation noise to higher frequencies through oversampling and feedback. The reconstruction filter is then used to attenuate the frequency-shaped quantisation noise. It operates by estimating the uniform quantisation error and employing a feedback filter to shape the noise power at the output of the DAC. A block diagram for a noise-shaping quantiser is shown in Fig. 2. The feedback filter  $F(z)$  is designed such that the transfer function  $y = (1 - F(z))\epsilon$  is a high-pass filter.

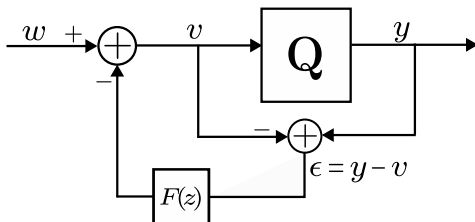


Figure 1: Noise shaping quantiser

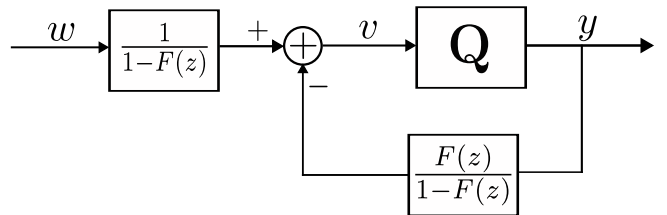


Figure 2: Noise shaping quantiser

In linear analysis, the output is given by

$$Y(z) = \mathbf{STF}.W(z) + \mathbf{NTF}.E(Z) \quad (2)$$

where the signal transfer function  $\mathbf{STF} = 1$ , noise transfer function  $\mathbf{NTF} = (1 - F)$  and  $F = z^{-1}$  for the first-order delta sigma modulator.

### 3 Moving horizon optimal quantiser (MHOQ)

The design criteria for the MHOQ is the minimization of the perceived errors defined as follows:

$$e(t) = H(z)(u(t) - y(t)) \quad (3)$$

where  $H(z)$  is a stable time-invariant linear low-pass filter with the following state-space

$$H(z) = 1 + C(zI - A)^{-1}B \quad (4)$$

The error  $e$  then can be written as the output of the following state-space representation of  $H$

$$\begin{aligned} x(t+1) &= Ax(t) + B(u(t) - y(t)) \\ e(t) &= Cx(t) + u(t) - y(t) \end{aligned} \quad (5)$$

where  $x \in \mathbb{R}^n$  is the state vector. The error  $e$  corresponds to the difference between the filtered quantised signal and the filtered input signal.

For moving horizon implementation, the optimisation problem is defined as the problem of finding  $y \in \mathbb{U}$  that minimises the cost function while satisfying the state equations as follows:

$$y^*(t) = \arg \min_{y(t)} V_N = \sum_{t=k}^{k+N-1} e^2(t) \quad (6)$$

subject to

$$x(t+1) = Ax(t) + B(y(t) - w(t)) \quad (7)$$

$$e(t) = Cx(t) + y(t) - w(t) \quad (8)$$

$$y(t) \in \mathbb{U}. \quad (9)$$

## 4 Frequency response: LPF and NTF

### 4.1 Perception filter

The frequency response and the STF and NTF of the perception filter used in [1] are shown in the figure below,

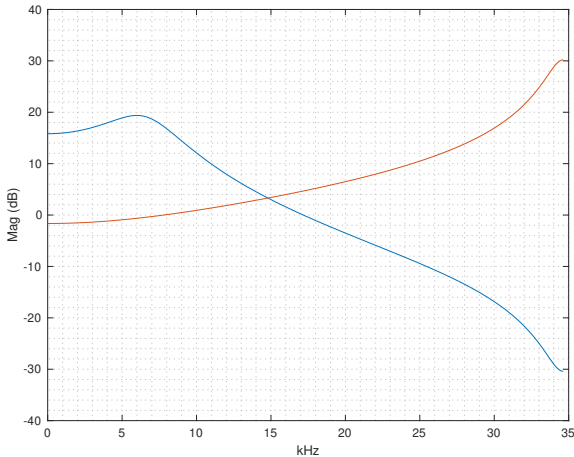


Figure 3: Perception filter Frequency response

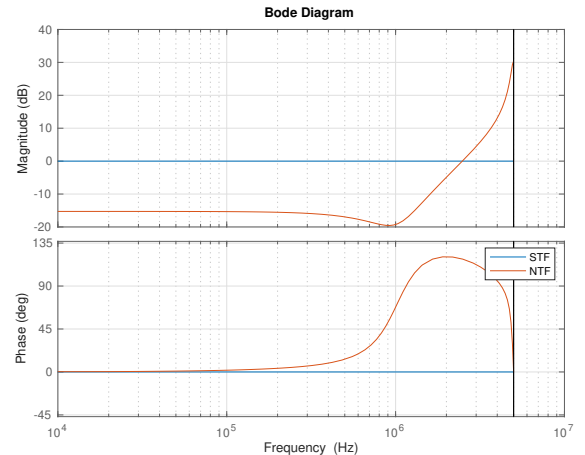


Figure 4: STF and NTF using perception filter

## 5 Noise Transfer Function(NTF)

The frequency response of the noise transfer functions due to butterworth filters at different cutoff frequencies are shown in the figure Fig. 5. In the figure, we can see that the net area under the curve remain the same. In Fig. 6 the frequency reponse of the low pass filter is plotted along with that of the noise transefer function. This observation shows that the better performance can be achieved by increasing the cutoff frequency during MHOQ while keeping the cutoff frequency of the reconstruction as same. The simulation results in the following table confirm this observation.

Table 1: ENOB at different cutoff frequencies. Reconstruction filter: Butterworth LPF with  $n = 2$ ,  $F_c = 100\text{kHz}$  and  $F_s = 1\text{Mhz}$ .

Fc	100 kHz	200 kHz	300 kHz	400 kHz	500 kHz
ENOB	3.981	5.307	7.817	10.481	10.936

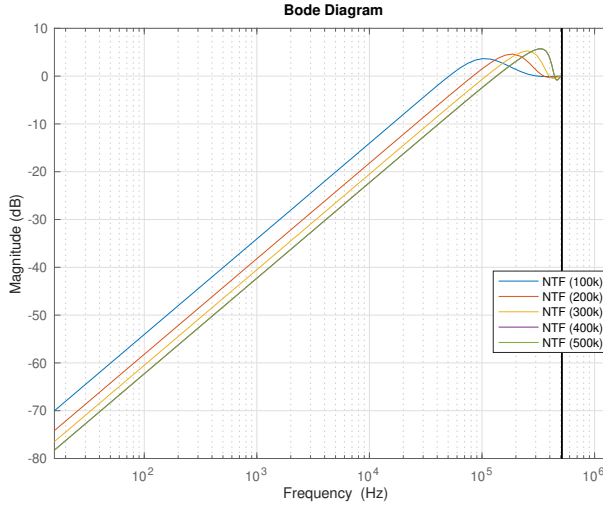


Figure 5: Frequency response of NTF for different cutoff frequency

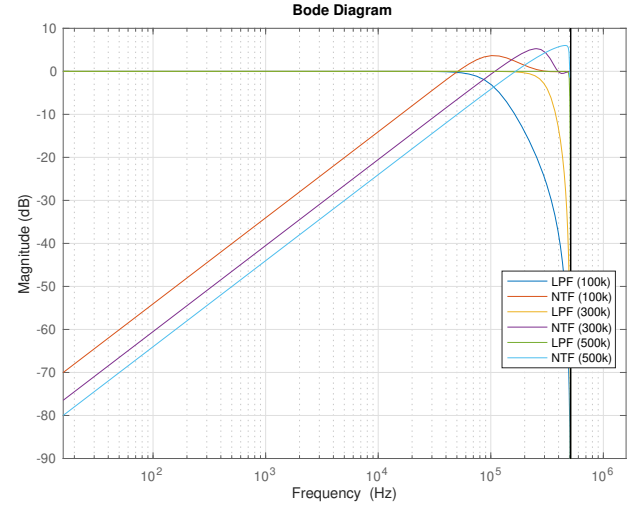


Figure 6: Frequency response of LPF and NTF for different cutoff frequency

## 6 Spectrum of quantisation noise

The sampling frequency effects the spectrum of the quantisation noise and consequently the ENOB as shown in the following figures.

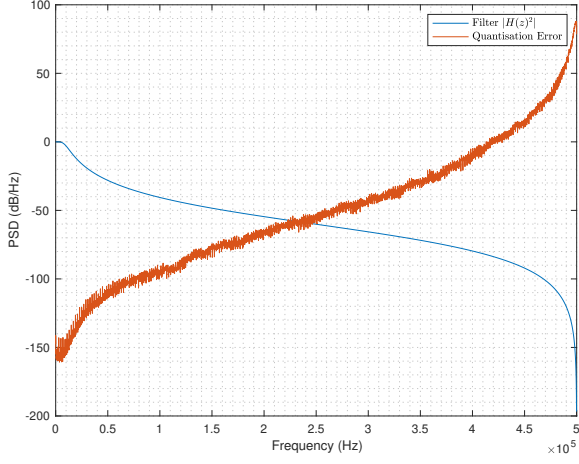


Figure 7: Butterworth Frequency response and frequency spectrum of quantisation noise:  
**Fc = 10 kHz, Fs = 1 MHz, ENOB = 16.58, INL**

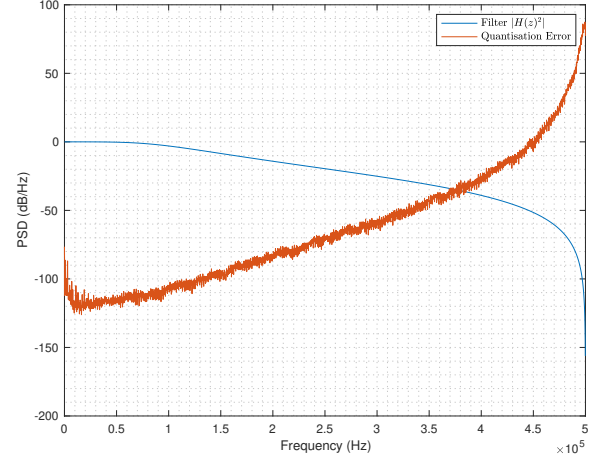


Figure 8: Butterworth Frequency response and frequency spectrum of quantisation noise:  
**Fc = 100 kHz, Fs = 1 MHz, ENOB = 7.43, INL**

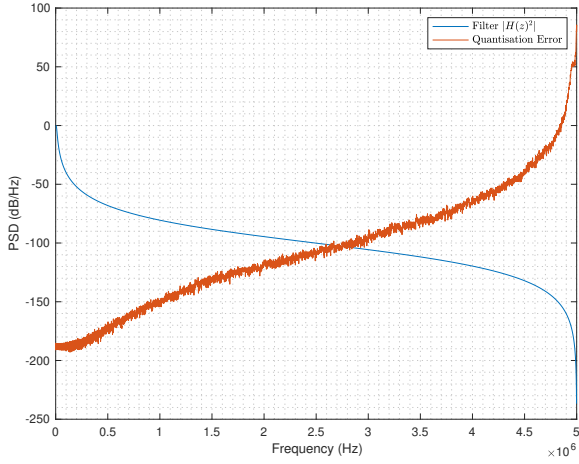


Figure 9: Butterworth Frequency response and frequency spectrum of quantisation noise:  
**Fc = 10 kHz, Fs = 10 MHz, ENOB = 25.12, INL**

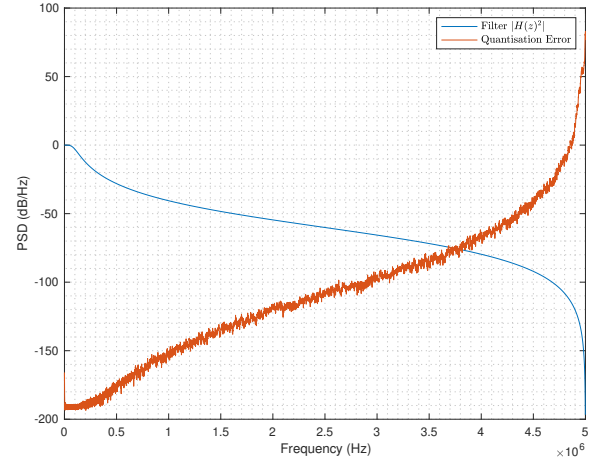


Figure 10: Butterworth Frequency response and frequency spectrum of quantisation noise:  
**Fc = 100 kHz, Fs = 10 MHz, ENOB = 17.03, INL**

## 7 Synthesis of optimal noise-shaping filter

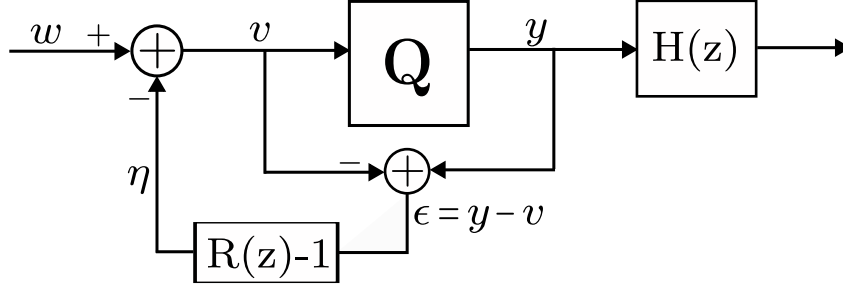


Figure 11: Noise shaping quantiser and a system  $H(z)$ .

In noise shaping quantiser with error-feedback structure as shown in the Figure 11, the input to the quantiser is  $v = w + \eta = w + (R(z) - 1)\epsilon$  and feedback error  $\epsilon$  is

$$\epsilon = y - v = y - w - (R(z) - 1)\epsilon. \quad (10)$$

Then quantisation noise defined as  $q := y - w$  can be expressed as

$$q = y - w = R(z)\epsilon. \quad (11)$$

Then the effect of the quantisation error on the system  $H(z)$  can be expressed as

$$e = H(z)R(z)\epsilon. \quad (12)$$

and it shows that we can reduce the error in the plant output by properly designing the noise shaping filter  $R(z)$  with the knowledge of the plant  $H(z)$ .

The objective is to design stable noise-shaping filter such that it minimises the effect of the quantisation noise in the plant output. A constraint on the error feedback signal should be imposed to prevent the quantiser from overloading and achieve a stable noise-shaping quantiser as

$$\eta = (R(z) - 1)\epsilon.$$

- **Design Problem:** For a fixed pair  $(p, q)$  [2],

$$\begin{aligned} & \min_{R(z) \in \mathbb{RH}_\infty} \|e\|_p \\ & \text{subject to.} \\ & R(\infty) = 1, \\ & \|\eta\|_q < \gamma_\eta \end{aligned} \quad (13)$$

where  $\mathbb{RH}_\infty = \mathbb{R} \cap \mathbb{H}_\infty$  is the set of proper stable rational transfer functions.

### 7.1 Solution 1: Optimal noise shaping filter without constraint

The optimal noise shaping filter without constraint on the error feedback signal  $\eta$  is the scaled inverse of the system  $H(z)$ . If the quantisation error is assumed to be an i.i.d random variable with zero mean, then the variance of the error  $e$  at time  $k$  can be expressed as

$$E\{|e_k|^2\} = \|H(z)R(z)\|_2^2 \sigma_\epsilon^2 \quad (14)$$

where  $E\{\cdot\}$  is the expectation operator and  $\sigma_\epsilon$  is variance of the  $\epsilon$  and  $\|H(z)R(z)\|_2^2$  is the  $\mathbb{H}_2$ -norm.

It is shown in [3], any casual and stable rational sytem function  $H(z)$  can always be stated as the product of the minimum phase system  $H_{min}(z)$  and all pass system  $H_{ap}(z)$ . Then it is shown that the optimal noise shaping filter is given by the scaled inverse of the plant, as follows,

$$R(z) = h_D H_{min}^{-1}(z) \quad (15)$$

where  $h_D$  is the first non-zero entry of the impulse response of the plant  $H_z$ .

## 7.2 Solution 2: Optimal noise shaping filter with constraint

The optimization problem is setup as the minimization of the upper bound of the  $\|e\|_p$  and  $\|\eta_q\|_q$  as follows:

$$\min_{R(z) \in \mathbb{RH}_\infty} \gamma_e \quad (16)$$

subject to  $R(\infty) = 1$  and

$$\|H(z)R(z)\|_{ind.p} < \gamma_e \quad (17)$$

$$\|R(z) - 1\|_{ind.q} < \gamma_\eta \quad (18)$$

where  $\|\cdot\|_{ind.r}$  is the induced norm and  $\gamma_e$  and  $\gamma_\eta$  are the upper bound of  $\|H(z)R(z)\|_{ind.p}$  and  $\|R(z) - 1\|_{ind.q}$ , respectively.

### State space representation

#### 7.2.1 System

Denoting the state-space representation of the  $H(z)$  and  $R(z)$  as  $(A_h, B_h, C_h, D_h)$  and  $(A_r, B_r, C_r, 1)$  respectively, the state space realization of  $H(z)R(z)$  is

$$x_{k+1} = Ax_k + B\epsilon_k \quad (19)$$

$$e_k = Cx_k + D\epsilon_k \quad (20)$$

where

$$A = \begin{bmatrix} A_h & B_h C_r \\ \mathbf{0} & A_r \end{bmatrix} \quad B = \begin{bmatrix} B_h \\ B_r \end{bmatrix} \quad C = [C_h \quad D_h C_r] \quad D = D_h. \quad (21)$$

Similary, the variance of the error  $e$  under the white noise assumption at time  $k$  in state-space form is given by  $\mathbb{H}_2$ -norm as

$$\|H(z)R(z)\|_2^2 = \sum_{k=0}^{\infty} \|CA^k B\|_2^2 + DD^\top. \quad (22)$$

Moreover, if  $A$  is Schur matrix, then there exists a positive semi-definite solution  $P$  of the discrete Lyapunov equation defined as

$$P = A^\top P A + BB^\top \quad (23)$$

and the squared  $\mathbb{H}_2$  norm is given by

$$\|H(z)R(z)\|_2^2 = CPC^\top + DD^\top. \quad (24)$$

#### 7.2.2 Constraint 1

Then  $\|H(z)R(z)\|_2 < \gamma_e$  if and only if there exist a positive definite matrix  $P$  such that

$$\text{(BMI)} \quad \begin{bmatrix} P & PA & PB \\ A^\top & P & \mathbf{0} \\ B^\top & \mathbf{0} & 1 \end{bmatrix} \succ 0 \quad (25)$$

$$\text{(LMI)} \quad \begin{bmatrix} \mu_e & C & D \\ C^\top & P & \mathbf{0} \\ D^\top & \mathbf{0} & 1 \end{bmatrix} \succ 0 \quad (26)$$

$$\mu_e = \gamma_e^2. \quad (27)$$

#### 7.2.3 Constraint 2

The variance of the noise shaping filter is given by

$$E\{|\eta_k|^2\} = \|R(z) - 1\|_2^2 = \sum_{k=1}^{\infty} \|\tilde{C}A^k B\|_2^2, \quad (28)$$

where  $\tilde{C} = \begin{bmatrix} \mathbf{0} & C_r \end{bmatrix}$ . Then  $\|R(z) - 1\|_2^2 < \gamma_\eta$  if and only if there exist a positive definite matrix  $P$  that satisfies

$$\text{(BMI)} \quad \begin{bmatrix} P & PA & PB \\ A^\top & P & \mathbf{0} \\ B^\top & \mathbf{0} & 1 \end{bmatrix} \succ 0 \quad (29)$$

$$\text{(LMI)} \quad \begin{bmatrix} \mu_\eta & \tilde{C} \\ \tilde{C}^\top & P \end{bmatrix} \succ 0 \quad (30)$$

$$\mu_\eta = \gamma_\eta^2. \quad (31)$$

#### 7.2.4 Optimization problem: State Space formulation

BMIs are not convex and NP hard to solve, but they can be converted to convex LMIs and can be solved numerically whereas the LMIs are convex. Thus the optimization problem can be written as follows,

$$\min_{R(z) \in \mathbb{RH}_\infty} \gamma_e \quad (32)$$

subject to  $R(\infty) = 1$  and

$$\begin{bmatrix} P & PA & PB \\ A^\top & P & \mathbf{0} \\ B^\top & \mathbf{0} & 1 \end{bmatrix} \succ 0 \quad (33)$$

$$\begin{bmatrix} \mu_e & C & D \\ C^\top & P & \mathbf{0} \\ D^\top & \mathbf{0} & 1 \end{bmatrix} \succ 0 \quad (34)$$

$$\begin{bmatrix} \mu_\eta & \tilde{C} \\ \tilde{C}^\top & P \end{bmatrix} \succ 0 \quad (35)$$

$$\mu_e = \gamma_e^2, \mu_\eta = \gamma_\eta^2. \quad (36)$$

## References

- [1] Graham C Goodwin, Daniel E Quevedo, and David McGrath. Moving-horizon optimal quantizer for audio signals. *Journal of the Audio Engineering Society*, 51(3):138–149, 2003.
- [2] Shuichi Ohno and M. Rizwan Tariq. Optimization of noise shaping filter for quantizer with error feedback. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 64(4):918–930, 2017.
- [3] Alan V. Oppenheim, Ronald W. Schaffer, and John R. Buck. *Discrete-Time Signal Processing*. Prentice-hall Englewood Cliffs, second edition, 1999.