

Report: Synthesis of optimal filter for MHOQ

June 20, 2024

Contents

1	Quantisation	1
2	Noise shaping quantiser	1
3	Moving horizon optimal quantiser (MHOQ)	2
4	Frequency response: LPF and NTF	2
4.1	Perception filter	2
5	Noise Transfer Function(NTF)	2
6	Spectrum of quantisation noise	3
7	Synthesis of optimal noise-shaping filter	4
7.1	Solution 1: Optimal noise shaping filter without constraint	5
7.2	Solution 2: Optimal noise shaping filter with constraint	5
7.2.1	System	5
7.2.2	Constraint 1	6
7.2.3	Constraint 2	6
7.2.4	Optimization problem: State Space formulation	6
8	LMI Synthesis: Convert BMIs to convex LMIs	7
8.1	Optimization Problem:	8
9	Simulatio: Butterworth LPF	8

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.

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.

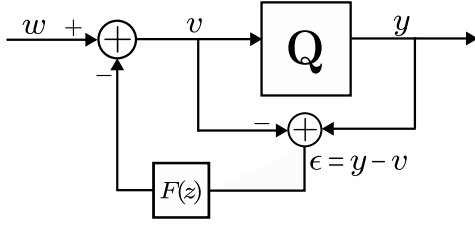


Figure 1: Noise shaping quantiser

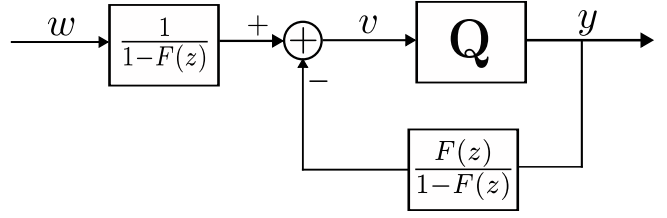


Figure 2: Noise shaping quantiser

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,

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

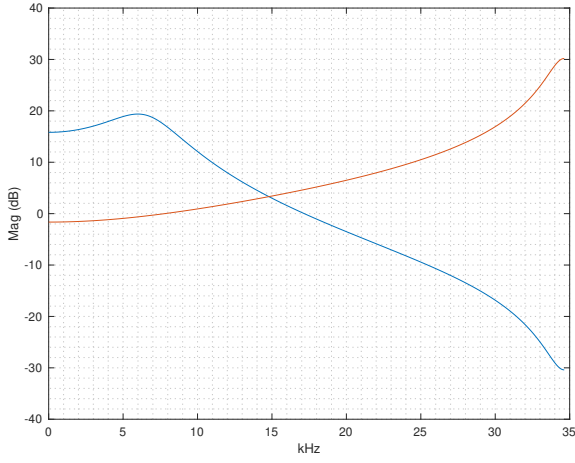


Figure 3: Perception filter Frequency response

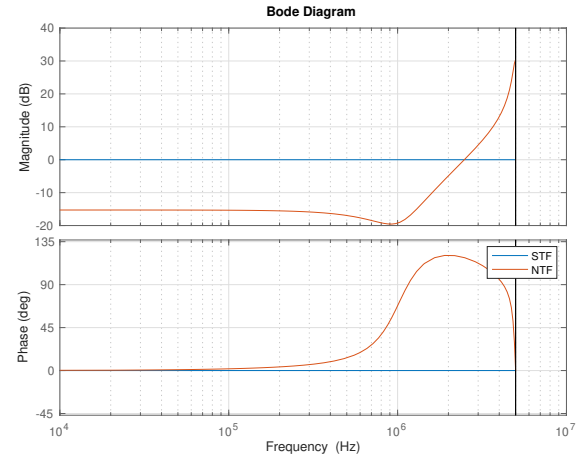


Figure 4: STF and NTF using perception filter

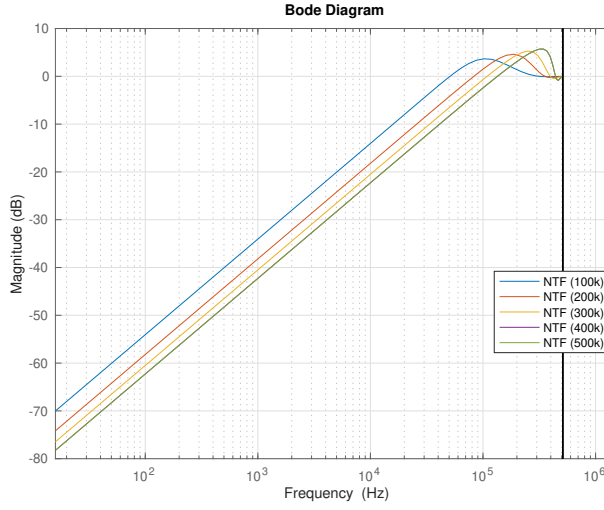


Figure 5: Frequency response of NTF for different cutoff frequency

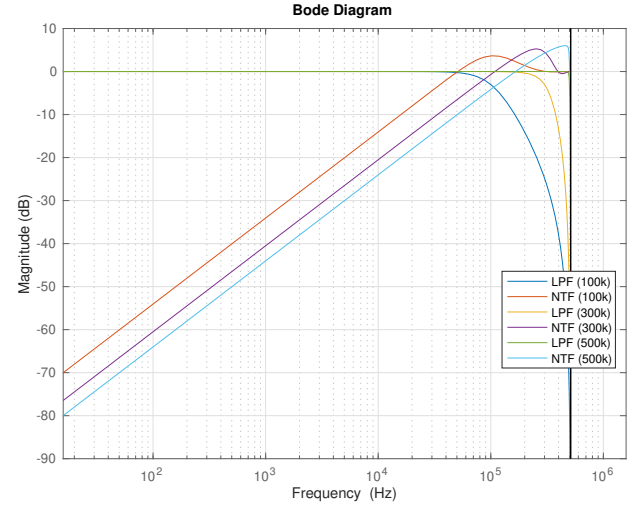


Figure 6: Frequency response of LPF and NTF for different cutoff frequency

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.

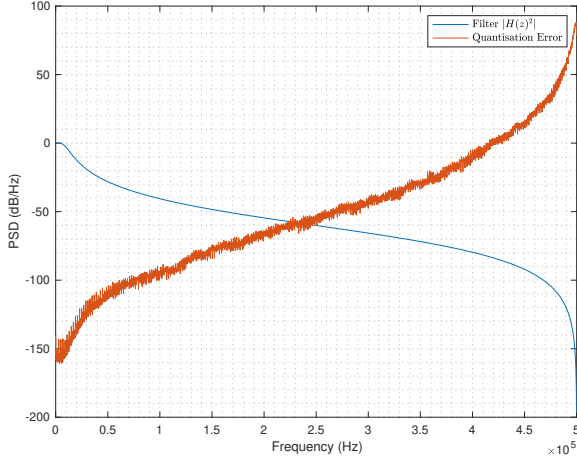


Figure 7: Butterworth Frequency response and frequency spectrum of quantisation noise:
 $F_c = 10$ kHz, $F_s = 1$ MHz, $ENOB = 16.58$, INL

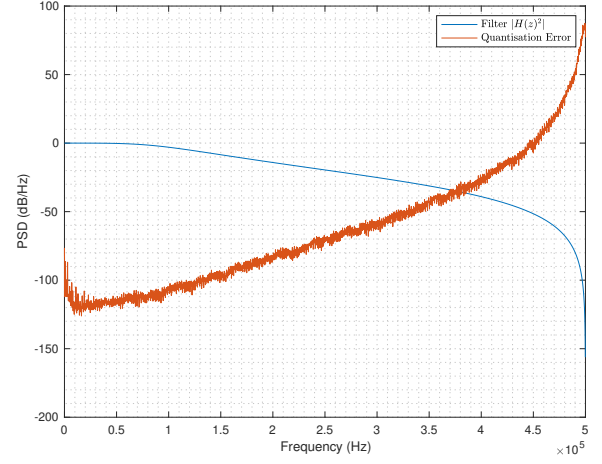


Figure 8: Butterworth Frequency response and frequency spectrum of quantisation noise:
 $F_c = 100$ kHz, $F_s = 1$ MHz, $ENOB = 7.43$, INL

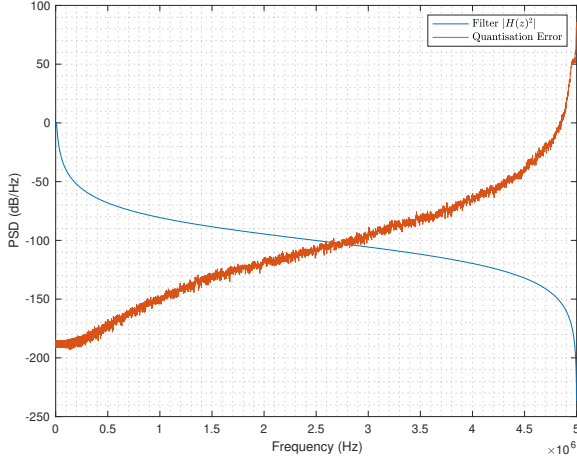


Figure 9: Butterworth Frequency response and frequency spectrum of quantisation noise:
 $F_c = 10$ kHz, $F_s = 10$ MHz, $ENOB = 25.12$, INL

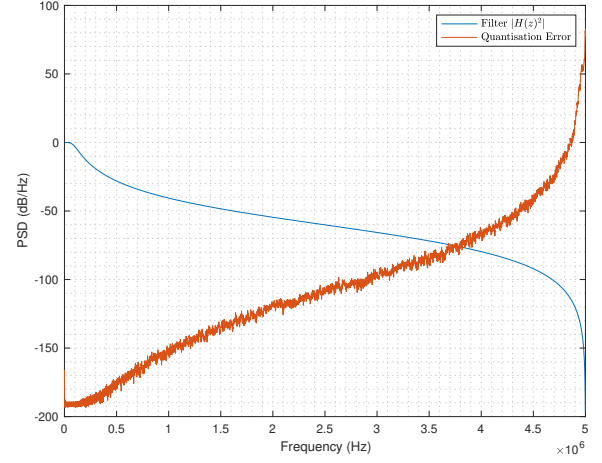


Figure 10: Butterworth Frequency response and frequency spectrum of quantisation noise:
 $F_c = 100$ kHz, $F_s = 10$ MHz, $ENOB = 17.03$, INL

7 Synthesis of optimal noise-shaping filter

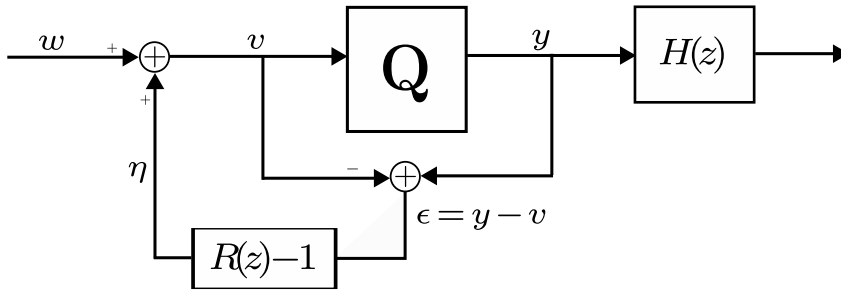


Figure 11: Noise shaping quantiser and a filter $H(z)$.

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 ϵ 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 η 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 σ_ϵ is variance of the ϵ 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 $\|\eta\|_q$ 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 γ_e and γ_η 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 $\tilde{C} = [\mathbf{0} \quad C_r]$. 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

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

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

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

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

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

8 LMI Synthesis: Convert BMIs to convex LMIs

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. The non-convex BMIs can be converted to convex LMIs using change of variables [4].

Change of Variables:

Let the order of $H(z)$ is n and the set of $n \times n$ positive definite matrices is denoted as $\text{PD}(n)$. Denote by \mathcal{P} the set of variables $\mathbf{p} = \{P_f, P_g, W_f, W_g, W_h, L\}$ where $P_f \in \text{PD}(n)$, $P_g \in \text{PD}(n)$, $W_f \in \mathbb{R}^{1 \times n}$, $W_g \in \mathbb{R}^{n \times 1}$, $W_h \in \mathbb{R}$ and $L \in \mathbb{R}^{n \times n}$. Then define the following matrix values function on \mathcal{P} :

$$\begin{aligned} M_A &:= \begin{bmatrix} A_h P_f + B_h W_f & A_h \\ L & P_g A_h \end{bmatrix} \\ M_B &:= \begin{bmatrix} B_h \\ W_g \end{bmatrix} \\ M_C &:= [C_h P_f + D_h W_f \quad C_h] \\ M_P &:= \begin{bmatrix} P_f & I_n \\ I_n & P_g \end{bmatrix}. \end{aligned} \tag{37}$$

Next, define

$$P^{-1} := \begin{bmatrix} P_f & S_f \\ S_f & S_f \end{bmatrix}, \quad U := \begin{bmatrix} P_f & I_n \\ S_f & \mathbf{0} \end{bmatrix} \quad \text{and} \tag{38}$$

$$S_f := P_f - P_g^{-1} (\succ 0) \tag{39}$$

then we have,

$$M_P = U^\top P U. \tag{40}$$

If the matrices (A_r, B_r, C_r) are given by

$$\begin{aligned} A_r &:= [B_h W_f - P_g^{-1}(L - P_g A_h P_f)] S_f^{-1} \\ B_r &:= [B_h - P_g^{-1} W_g] \\ C_r &:= W_f S_f^{-1} \end{aligned} \tag{41}$$

then (A, B, C) satisfy,

$$M_A = U^\top P A U \tag{42}$$

$$M_B = U^\top P B \tag{43}$$

$$M_C = C U \tag{44}$$

$$M_p = U^\top P U. \tag{45}$$

Multiplying with the transformation $\phi = \text{diag}(U, U, 1)$ from the RHS and ϕ^\top from the LHS the BMI condition 33 takes the following form:

$$\begin{bmatrix} M_P & M_A & M_B \\ M_A^\top & M_P & \mathbf{0} \\ M_B^\top & \mathbf{0} & \mathbf{I} \end{bmatrix} \succ 0. \tag{46}$$

Similarly, using the transformation $\text{diag}(1, U, 1)$, the LMI condition 34 is

$$\begin{bmatrix} \mu_e & M_C & D^\top \\ M_C^\top & M_P & \mathbf{0} \\ D & \mathbf{0} & \mathbf{I} \end{bmatrix} \succ 0 \tag{47}$$

and finally the constraint 35 using the transformation is

$$\begin{bmatrix} \gamma_{\tilde{C}}^2 & M_{\tilde{C}} \\ M_{\tilde{C}}^\top & M_P \end{bmatrix} \succ 0 \tag{48}$$

with $M_{\tilde{C}} := \tilde{C}U$. The LMI conditions 46, 47 and 48 are convex and the minimization of γ_e with these constraints is a convex optimization problem.

8.1 Optimization Problem:

With the change of the variable the BMI is converted to the LMI and the since all the LMIs are convex the optimization problem is a convex optimization problem. Also, since the objective is linear and the constraints are LMIs and linearity constraints, the optimization problem is SemiDefinite Program (SDP). Such SDP can be solved numerically using CVX. The optimization problem takes the following form,

$$\min \mu_e = \gamma_e^2 \quad (49)$$

subject to,

$$\begin{bmatrix} M_P & M_A & M_B \\ M_A^\top & M_P & \mathbf{0} \\ M_B^\top & \mathbf{0} & \mathbf{I} \end{bmatrix} \succ 0 \quad (50)$$

$$\begin{bmatrix} \mu_e & M_C & D^\top \\ M_C^\top & M_P & \mathbf{0} \\ D & \mathbf{0} & \mathbf{I} \end{bmatrix} \succ 0. \quad (51)$$

$$\begin{bmatrix} \mu_\eta & M_{\tilde{C}} \\ M_{\tilde{C}}^\top & M_P \end{bmatrix} \succ 0 \quad (52)$$

where μ_e is the variance of the output error and μ_η the variance of the feedback error.

9 Simulatio: Butterworth LPF

Let us consider a second order butterworth low pass filter with $F_c = 100kHz$ and $F_s = 1Mhz$. Also, denote the low pass filter as $H(z)$ then let $F_{opt}(z)$ as the optimal noise transfer function and $F_{mpc}(z)$ as the noise transfer function due to MPC setting i.e. $F_{mpc} = (H(z) - 1)/H(z)$ and the numerator and denominator of noise transfer function $R(z)$ as b_r and a_r , respectively, i.e.,

$$R(z) = \frac{b_r}{a_r}. \quad (53)$$

Then the corresponding noise shaping filter (NSF) as shown in the Figure 11 is

$$F(z) = 1 - R(z) = \frac{a_r - b_r}{a_r}. \quad (54)$$

Following the analogy of MPC formulation the obtained low pass filter due the high- noise shaping filter is

$$H_{mpc}(z) = \frac{1}{1 - F(z)} = \frac{a_r}{b_r}. \quad (55)$$

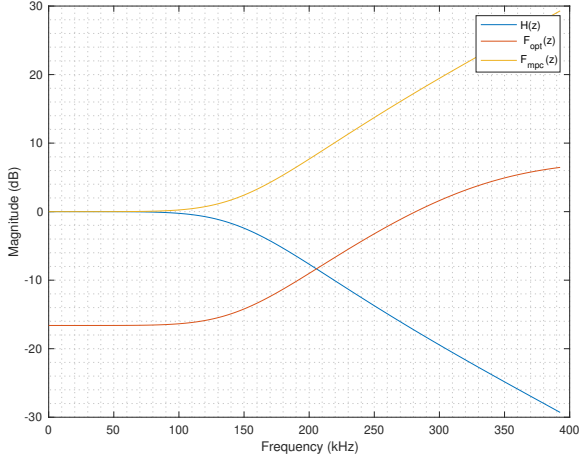


Figure 12: Frequency response: Butterworth filter with $n = 3$ and $F_c = 100kHz$ and corresponding optimal noise shaping filter and low pass filter

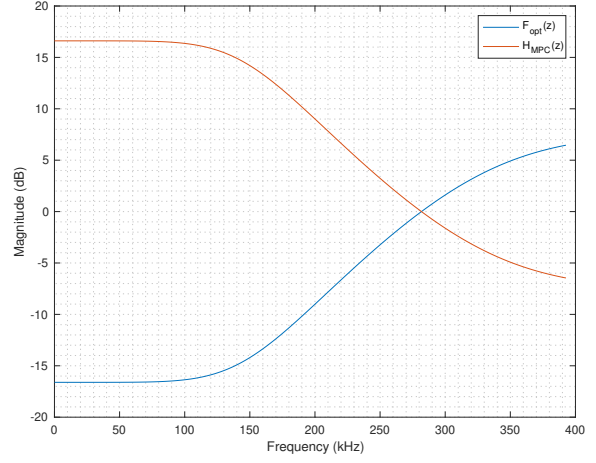


Figure 13: Frequency response: Optimal noise shaping filter and corresponding low pass filter

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.
- [4] Izumi Masubuchi, Atsumi Ohara, and Nobuhide Suda. Lmi-based controller synthesis: A unified formulation and solution. *International Journal of Robust and Nonlinear Control*, 8(8):669–686, 1998.