

2.8 Measurement setups

2.8.1 Zero-order hold setup

In a lot of measurement setups, the input signal is generated digitally with a Zero-order hold (ZOH). This means that the signal is kept constant for a whole sampling period. At the sampling instances, a CT system $G(s)$ preceded by a ZOH can be modeled exactly as a DT system. *instants*

$$G_{\text{ZOH}}(z) = (1 - z^{-1}) \mathcal{Z} \left\{ \mathcal{L}^{-1} \left\{ \frac{G(s)}{s} \right\} \Big|_{t=nT_s} \right\} \quad (2.5)$$

The ZOH measurement setup is shown in figure 2.5. In this case, the actuator $G_{\text{act}}(s)$ is part of the FRF that is measured. The dynamics of the measurement device $G_y(s)$ must be calibrated perfectly ($G_y(s) = 1$) if one wants to measure the FRF from generator to output. Note that, in theory, the ZOH setup is not allowed to contain anti-alias filters.

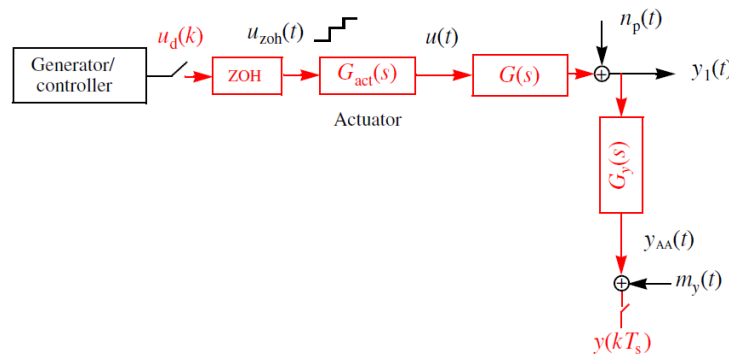


Figure 2.5 – Zero-order hold measurement setup. Red indicates which parts are modeled. Taken from [1].

Thus, if no anti-alias filter is used at the output and assuming that the actuator is perfect $G_{\text{act}}(s) = 1$, when a CT system is excited with a ZOH and is sampled, we are actually measuring the ZOH version of the FRF and not the CT FRF directly. This is bad news if we want to measure the CT FRF $G(s)$. However, this can be circumvented. When $f \ll f_s/2$, there isn't a big difference between the CT FRF $G(s)$ and the ZOH version of $G(s)$. *ll*

$$G(j2\pi f) \approx G_{\text{ZOH}}(e^{j2\pi f T_s}) \quad (2.6)$$

So, when one wants to measure a CT FRF with a ZOH setup, care must be taken to stay within the region where this approximation holds. This can be accomplished by making sure that the highest excited frequency in the input signal is well below the sampling frequency of the measurement setup. *ll*

$$f_{\text{max}} \ll f_s/2$$

Example Consider a second order system.

$$G(s) = \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2} \text{ with } \omega_0 = 2\pi 0.3 [\text{rad/s}] \text{ and } \zeta = 0.01$$

Applying the ZOH transformation (2.5) to this system with $f_s = 2\text{Hz}$ results in the DT system that was used in previous examples (2.4). The magnitude FRF of both the CT system and the ZOH version of it are plotted in figure 2.6. Notice that the DT system has a periodic FRF while the CT system does not. Moreover, at the low frequencies, both FRF overlap. The approximation (2.6) is worse once the frequency gets close to $f_s/2$.

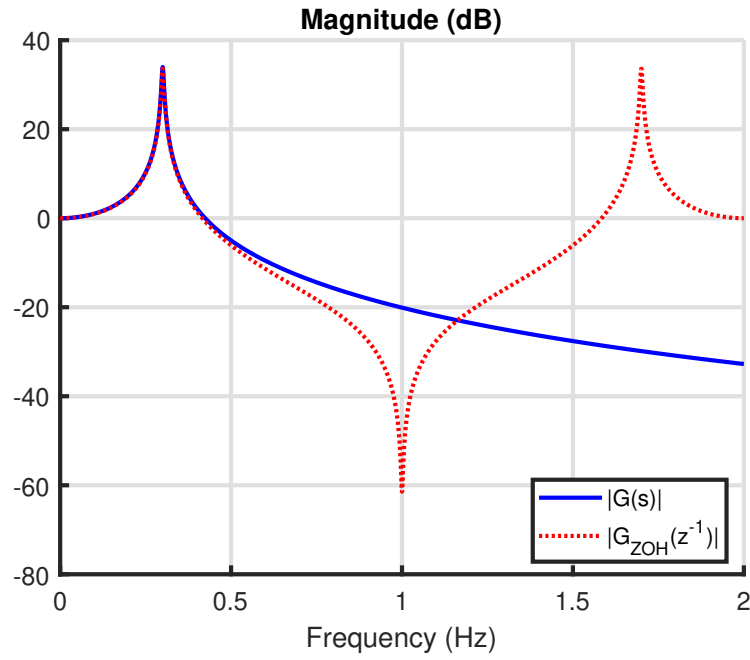


Figure 2.6 – Magnitude FRF of the second-order system and the magnitude FRF of the ZOH version of it with $f_s = 2\text{Hz}$.

2.8.2 Band-limited setup

Even if you use a high sampling frequency f_s , the actuator dynamics will still influence the measurement of the FRF when using a ZOH setup. This is not an issue when a band-limited setup is used. The band-limited setup is shown in figure 2.7. In a band-limited setup, the input to the system can still be generated by a ZOH. But the input to the system must be measured and an anti-alias filter must be used to avoid aliasing. By measuring the input to the system, the dynamics of the actuator won't be part of the FRF estimate. The output of the system must also go through an anti-alias filter before being measured. A relative calibration is needed if one wants to measure the FRF from the input to the output. Concretely this means that the anti-alias filters must be equal to each other.

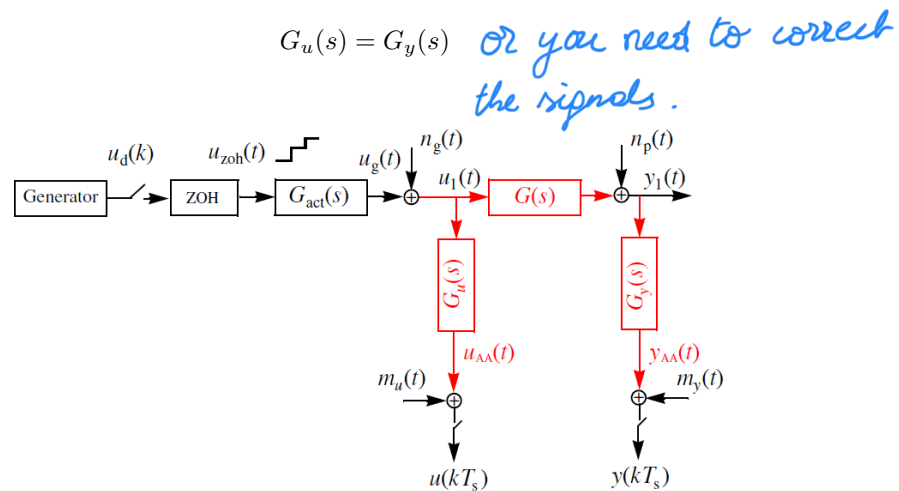


Figure 2.7 – Band-limited measurement setup. Red indicates which parts are modeled. Taken from [1].

2.9 System transients

The first condition of perfect reconstructability (periodicity) implies that the signal has been repeating forever until now and will repeat forever in the future. This is not realistic; when a system is excited, the excitation must have started *at some time* *somewhere* in the past and must end at some point in the future. In the example of the previous section we had briefly mentioned that the measurements were taken in steady state. Concretely, the input was applied to the system for 20 periods and only the last period was used for the analysis, thereby ensuring that the system transients have faded away. *This can be quantified by calculating the RMS of the difference between the output of the p -th period and the output of the last period.* This is plotted in figure 2.8. The RMS of the difference decreases exponentially. If the measurements were noisy, the RMS of the difference would decrease until the transients are below noise level, at which point the transients can be said to have faded away. In this case, the RMS of the difference reaches approximately 10^{-7} by the 20-th period, which is negligible for our purposes.

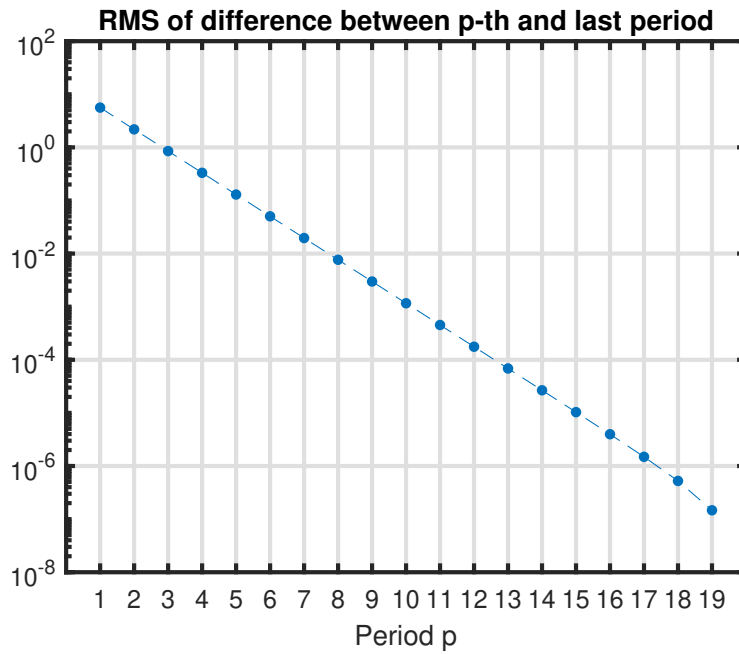


Figure 2.8 – RMS of the difference between the output of every period and the output of the last period.

It turns out that the transient term is just a rational function added on top of the steady state output spectrum.

$$Y(k) = G(e^{-j2\pi k/N})U(k) + T(k) \quad (2.7)$$

with

$$T(k) = \frac{I(e^{-j2\pi k/N})}{A(e^{-j2\pi k/N})}$$

This is proven for DT systems in appendix 2.A. The transient term for CT systems will be discussed afterwards.

If $u(n)$ and $y(n)$ are periodic, then $T(k) = 0$, because $I(e^{-j2\pi k/N})$ only depends on the difference between the in- and output in the current period and the in- and output in the previous period. In this case (2.7) simplifies to what we had before.

$$Y(k) = G(e^{-j2\pi k/N})U(k)$$

A key property of the transient term $T(k)$ is that it is a “smooth” function of the frequency k . This is due to the fact that it is a rational form of $e^{-j2\pi k/N}$. This property can be used to suppress the transient, as will be explained later. Another property of $T(k)$ is that its denominator is the same as the denominator of $G(e^{-j2\pi k/N})$. Therefore, the transient term will “resemble” the shape of the FRF.

CT transients Similar results can be derived for CT systems.

$$Y(k) = G(j\omega)U(k) + T(k) + \delta(k)$$

In this case, the numerator of $T(k)$ depends on the difference in initial conditions at $t = 0$ and $t = T = nT_s$:

$$p < n_a, n_b \quad \left[\frac{d^p y}{dt^p}(T) - \frac{d^p y}{dt^p}(0) \right] \text{ and } \left[\frac{d^p u}{dt^p}(T) - \frac{d^p u}{dt^p}(0) \right]$$

with $p \in \mathbb{N}$. An additional term $\delta(k)$ pops up. This is the alias error and it can be generated when the transient $T(j\omega)$ overextends into the aliasing frequencies $f > f_s/2$. It is present even if the signals have been low-pass filtered. Because the alias error $\delta(k)$ is also smooth, it can be grouped together with the transient term $T(k)$.

Example Again, the same second-order system (2.4) of the previous section is used. This time I'll take a look at the estimation error.

$$\begin{aligned} G_{\text{est}}(e^{-j2\pi k/N}) - G(e^{-j2\pi k/N}) &= \frac{Y(k)}{U(k)} - G(e^{-j2\pi k/N}) \\ &= \frac{G(e^{-j2\pi k/N})U(k) + T(k)}{U(k)} - G(e^{-j2\pi k/N}) \\ &= \frac{T(k)}{U(k)} \end{aligned}$$

As the magnitude spectrum of $u(n)$ is flat, the magnitude of the estimation error will be proportional to the magnitude of the transient term.

The estimation error for the FRF estimated from the first and 20th period is shown in figure 2.9. The error of the FRF calculated from the last period is around -150dB , which is negligible. In other words, there is no transient term as the data is in steady state. The error for the first period is a smooth function of the frequency and “resembles” the shape of the transfer function as expected.

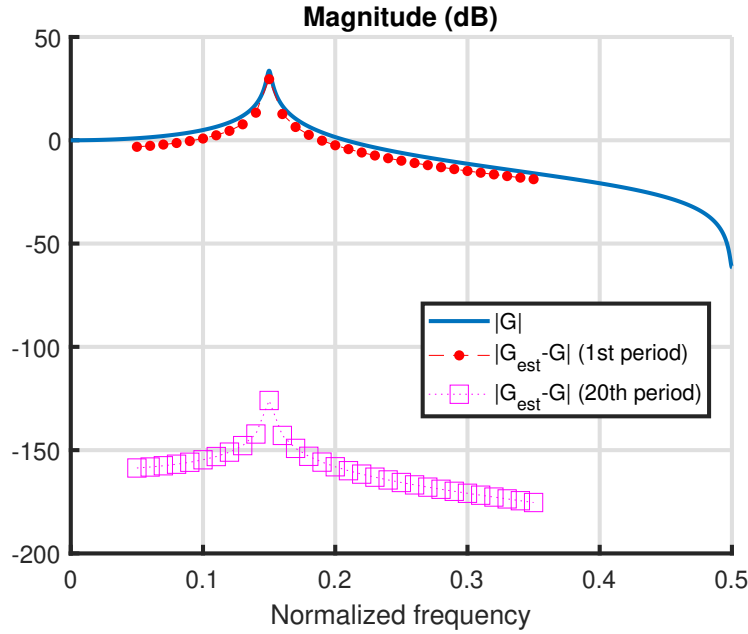


Figure 2.9 – Magnitude FRF of the second-order system and the estimation errors of the FRF at the excited frequencies using (2.3) for the data from the first period and the 20th period.

2.10 Frequency domain methods

Till now we have been treating CT and DT systems as separate cases. However, FD methods can be generalized to both CT and DT systems. There is no reason why a FD method should only work for one and not the other. To generalize these methods it is useful to define Ω to encompass both CT and DT models. Ω can either be s or z^{-1} .

$$\Omega = \begin{cases} s & \text{if working in continuous-time (CT)} \\ z^{-1} & \text{if working in discrete-time (DT)} \end{cases}$$

Thus, denoting an LTI system as $G(\Omega)$ is not specific to CT or DT systems. When the FRF is evaluated in the DFT frequencies $\omega_k = 2\pi k f_s / N$ it, is denoted Ω_k .

$$\Omega_k = \begin{cases} j\omega_k = j2\pi k f_s / N & \text{if working in continuous-time (CT)} \\ e^{-j\omega_k T_s} = e^{-j2\pi k / N} & \text{if working in discrete-time (DT)} \end{cases}$$

2.11 White noise

White noise is a random signal with a flat power spectrum. A zero-mean Gaussian sequence is white noise.

$$v(n) \sim \mathcal{N}(0, \sigma^2)$$

Let's assume that a measurement is perturbed by this white noise.

$$x(n) = x_0(n) + v(n)$$

Due to the linearity of the DFT one obtains

$$X(k) = X_0(k) + V(k)$$

Thus it is interesting to understand the properties of the DFT of this white noise sequence.

$$V(k) = \sum_{n=0}^{N-1} v(n) e^{-j2\pi kn/N}$$

It turns out that (see appendix 2.B)

$$\begin{aligned} \mathbb{E}\{V(k)\} &= 0 \\ \mathbb{E}\{|V(k)|^2\} &= N\sigma^2 \\ \mathbb{E}\{V^2(k)\} &= 0 \text{ if } \text{mod}(2k, N) \neq 0 \end{aligned}$$

2.12 Periodic signals

What happens to the DFT spectrum when we apply a signal periodically? Let's assume that we have a sequence $\tilde{x}(n)$ of length N with corresponding DFT

$$\tilde{X}(k) = \sum_{n=0}^{N-1} \tilde{x}(n) e^{-j2\pi kn/N}$$

Now, the sequence $\tilde{x}(n)$ will be repeated one more time to obtain a sequence of length $2N$.

$$x(n) = \begin{cases} \tilde{x}(n) & \text{if } 0 \leq n < N-1 \\ \tilde{x}(n-N) & \text{if } N \leq n < 2N-1 \end{cases}$$

The DFT of $x(n)$ is then given by

$$\begin{aligned} X(k) &= \sum_{n=0}^{2N-1} \tilde{x}(n) e^{-j2\pi kn/(2N)} = \sum_{n=0}^{N-1} \tilde{x}(n) e^{-j2\pi kn/(2N)} + \sum_{n=N}^{2N-1} \tilde{x}(n-N) e^{-j2\pi kn/(2N)} \\ &= \sum_{n=0}^{N-1} \tilde{x}(n) e^{-j2\pi kn/(2N)} + \sum_{n=0}^{N-1} \tilde{x}(n) e^{-j2\pi kn/(2N)} e^{-j2\pi k/2} \\ &= \sum_{n=0}^{N-1} \tilde{x}(n) e^{-j2\pi kn/(2N)} (1 + e^{-j2\pi k/2}) \end{aligned}$$

Evaluating this in $2k$ and $2k+1$ gives

$$\begin{aligned} X(2k) &= 2\tilde{X}(k) \\ X(2k+1) &= 0 \end{aligned}$$

Thus, the even DFT lines will contain the information of $X(k)$ and the uneven DFT lines will be zero.

This result can be generalized to a signal that is repeated for P times.

$$x(n) = \tilde{x}(\text{mod}(n, P)) \Rightarrow X(kP + r) = \begin{cases} P\tilde{X}(k) & \text{if } r = 0 \\ 0 & \text{if } r = 1, \dots, P-1 \end{cases}$$

Example A random phase multisine ($N = 40$) is created where the 3 lowest DFT bins are excited with an RMS value of 1. This signal is repeated $P = 3$ times. The repeated signal is plotted in the TD and the FD in figure 2.10. It can be seen that in the periodic signal there are $P - 1 = 2$ non-excited lines in between the excited lines.

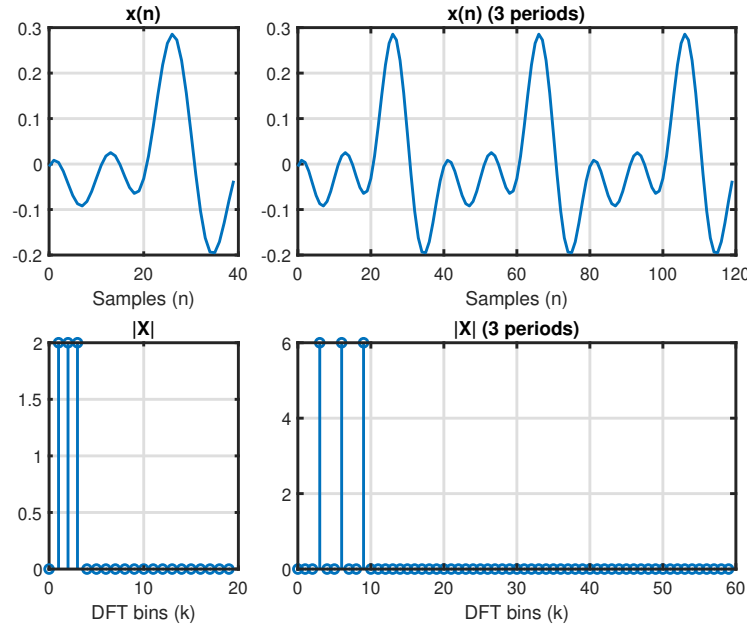


Figure 2.10 – A multisine is repeated 3 times.

2.13 Transient suppression

In section 2.9 it was shown that transients in the output will negatively impact the quality of the FRF estimate. However, there are multiple ways to suppress these transients. The most obvious way is to wait for the system to enter steady state and to only start measuring once the transients have faded away. This is exactly what was done in the example of section 2.9 (see figure 2.9). There are also ways to suppress the transient without throwing away the measurements that are not in steady state. 3 of them will be discussed:

- Windowing
- Parametric estimation
- The local polynomial method

2.13.1 Windowing

There are many possible windows that one can choose from. The Hann window is a popular one.

$$w(n) = \frac{1}{2} \left[1 - \cos\left(\frac{2\pi n}{N}\right) \right], \quad n = 0, \dots, N-1 \quad (2.8)$$

The windowed input and output are then given by

$$u_w(n) = w(n)u(n) \text{ and } y_w(n) = w(n)y(n)$$

A multiplication in the TD becomes a convolution in the FD.

$$U_W(k) = W(k) * U(k) \text{ and } Y_W(k) = W(k) * Y(k)$$

with $W(k) = N \left[\frac{1}{2}\delta(k) - \frac{1}{4}\delta(k+1) - \frac{1}{4}\delta(k-1) \right]$. The windowed estimate of the FRF is then given by

$$G_W(\Omega_k) = \frac{Y_W(k)}{U_W(k)} \quad (2.9)$$

Let's now see why windowing will reduce the effect of the transient.

$$Y_W(k) = W(k) * (G(\Omega_k)U(k) + T(k))$$

For the first term the following approximation can be made:

$$W(k) * [G(\Omega_k)U(k)] \approx G(\omega_k)[W(k) * U(k)] = G(\omega_k)U_W(k)$$

Hereby it is assumed that $G(\Omega_{k-1}) \approx G(\Omega_k) \approx G(\Omega_{k+1})$. In other words, $G(\Omega)$ is flat around Ω_k . (2.9) then becomes

$$G_W(\Omega_k) \approx G(\Omega_k) + \frac{T_W(k)}{U_W(k)}$$

The trick now is to use the property that $T(k)$ is a smooth function of the frequency. Let's assume that $T(k)$ can be captured locally by a second order polynomial.

$$T(k) = ak^2 + bk + c$$

Windowing $T(k)$ gives

$$T_W(k) \propto 2T(k) - T(k+1) - T(k-1) = -2a$$

This is a bit like taking the second derivative, which suppresses the transient.

Example The DT system (2.4) is taken again. The same input as in figure 2.2 is used. The estimation error of the FRF with and without windowing is shown in figure 2.11. Windowing gives better results, except near the resonance frequency. This is because the approximation that $G(\Omega)$ is flat does not hold in this region.

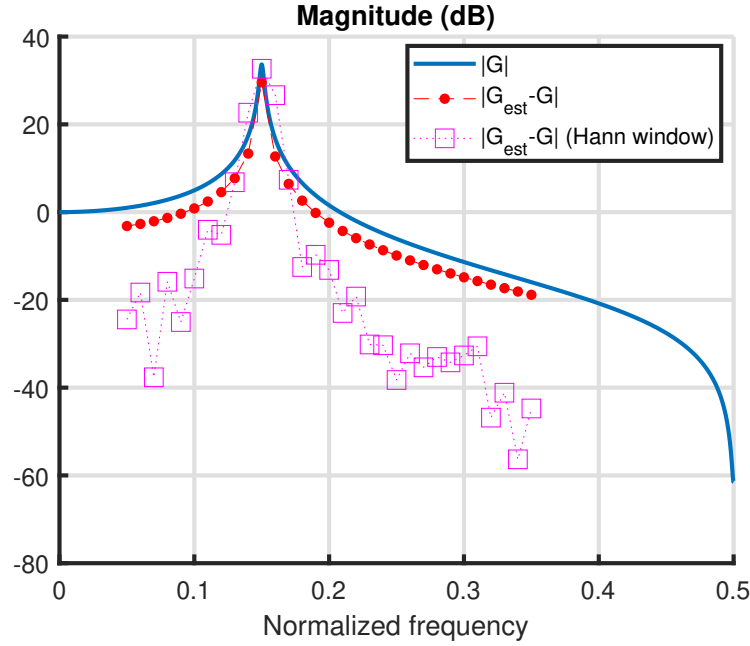


Figure 2.11 – Comparison of FRF estimation error without and with Hann windowing.

2.13.2 Parametric estimation

Until now, only non-parametric estimates of the FRF were discussed. Unlike line-fitting, where one is interested in the slope and offset of the line, the non-parametric estimate does no such thing. The result of the non-parametric estimate is the FRF estimate at every excited frequency. However, in a sense there are still parameters: these are the FRF estimates at every frequency bin $k \in K_{\text{exc}}$.

The TF of the system is however still given by a rational function. The coefficients of this rational function can be estimated. The transient term is also a rational function and the parameters describing it can also be included in the estimation.

$$Y(k) = G(\Omega_k)U(k) + T(k)$$

Using $G(\Omega_k) = \frac{B(\Omega_k)}{A(\Omega_k)}$ and $T(k) = \frac{I(\Omega_k)}{A(\Omega_k)}$ the expression above can be rewritten as

$$A(\Omega_k)Y(k) - B(\Omega_k)U(k) - I(\Omega_k) = 0 \quad (2.10)$$

$A(\Omega_k)$ is a polynomial of order n_a and $B(\Omega_k)$ is a polynomial of order n_b . (2.18) is an explicit formula for $I(\Omega_k)$. It is easy to see that $I(\Omega_k)$ is a polynomial of order n_I with

$$n_I = \max(n_a, n_b) - 1$$

For CT systems, there is also an alias error on top of the transient term. These errors can also be captured well by a polynomial. Grouping the alias error together with the $I(\Omega_k)$ results in [2, Section 6.3.2.3]

$$n_I \geq \max(n_a, n_b)$$

After a bit of calculations, (2.10) can be turned into

$$\begin{bmatrix} \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Y(k) & \dots & Y(k)\Omega_k^{n_a} & U(k) & \dots & U(k)\Omega_k^{n_b} & 1 & \dots & \Omega_k^{n_I} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} a_0 \\ \vdots \\ a_{n_a} \\ -b_0 \\ \vdots \\ -b_{n_b} \\ -i_0 \\ \vdots \\ -i_{n_I} \end{bmatrix} = 0$$

This can be solved by calculating the right null-space of the first matrix. The coefficients i_p will capture the transient. Note that the null space is empty if the measurements are noisy. We won't go into these details in this work, so interested reader is referred to [3].

Example Again, the DT system (2.4) is used with the same input shown in figure 2.2. The parameters are determined as described above. The resulting FRF estimation error is plotted in figure 2.12. Note that as a parametric representation is obtained, the FRF can be calculated at all frequencies. The error is around -300dB , which is MATLAB's precision. This means that the error is negligible and that the transient has been fully suppressed.

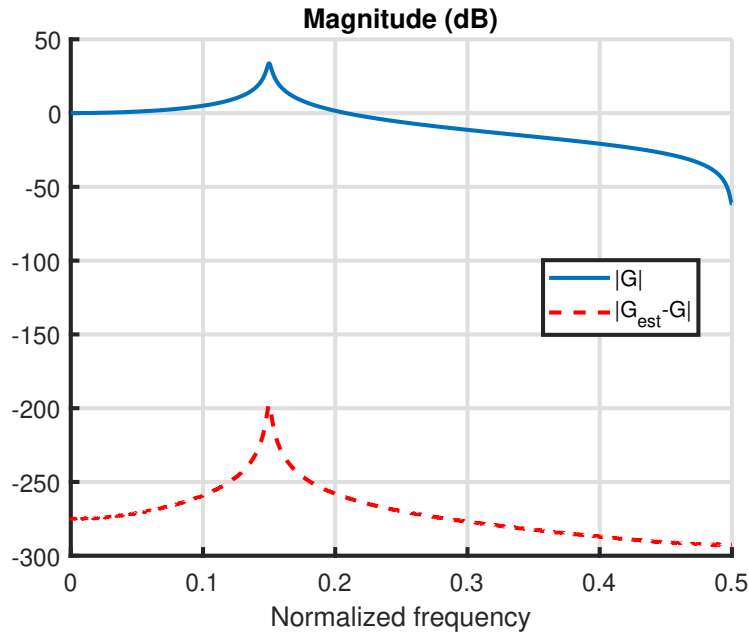


Figure 2.12 – FRF estimation error for the parametric estimation of the TF with transient terms.

2.14 Local polynomial method

The idea of the local polynomial method (LPM) is quite simple. There are multiple variants of the LPM, but here we will focus on the robust LPM for periodic excitations. More information about all the variants of the LPM can be found in [2, Chapter 7].

2.14.1 Response of a system excited by a periodic input

For the robust LPM for periodic signals to work, at least $P = 2$ periods must be measured. Let's assume for the sake of simplicity that the input $\tilde{u}(n)$ is noiseless and in steady state. $\tilde{u}(n)$ is applied P consecutive times to the system.

$$u(n) = \tilde{u}(\text{mod}(n, P)), \quad n = 0, 1, \dots, NP - 1$$

The spectrum of $u(n)$ is

$$U(kP + r) = \begin{cases} P\tilde{U}(k) & \text{if } r = 0 \\ 0 & \text{if } r = 1, \dots, P - 1 \end{cases} \quad (2.11)$$

It will resemble the spectrum shown in figure 2.10. Assuming that the output is perturbed by additive noise, the output spectrum is given by

$$Y(kP + r) = G(kP + r)U(kP + r) + T(kP + r) + N_y(kP + r)$$

$T(kP + r)$ is the transient term that arises from the fact that the system might not be in steady state. $N_y(kP + r)$ is additive noise that is added to the output. It can either be white noise (see section 2.11) or filtered white noise. Using (2.11), we can be more specific about the output spectrum.

$$Y(kP + r) = \begin{cases} G(kP)U(kP) + T(kP) + N_Y(kP) & \text{if } r = 0 \\ T(kP + r) + N_Y(kP + r) & \text{if } r = 1, \dots, P - 1 \end{cases}$$

Example The system (2.4) is excited with the signal shown in figure 2.10. The output is perturbed by Gaussian white noise with a standard deviation of 0.005. The input and output spectra are plotted in figure 2.13. The output spectrum at the DFT lines 3, 6 and 9 are dominated by the $G(kP)U(kP)$ term. There is a peak around the 18-th DFT line that corresponds to the transient term T . The 18-th DFT line corresponds to the normalized frequency $18/(NP) = 18/120 = 0.15$, which corresponds to the normalized resonance frequency of the system (see figure 2.12 for example). This is to be expected as the transient resembles the shape of the transfer function. Finally, after the 30-th DFT bin, the noise terms N_Y dominate the output.

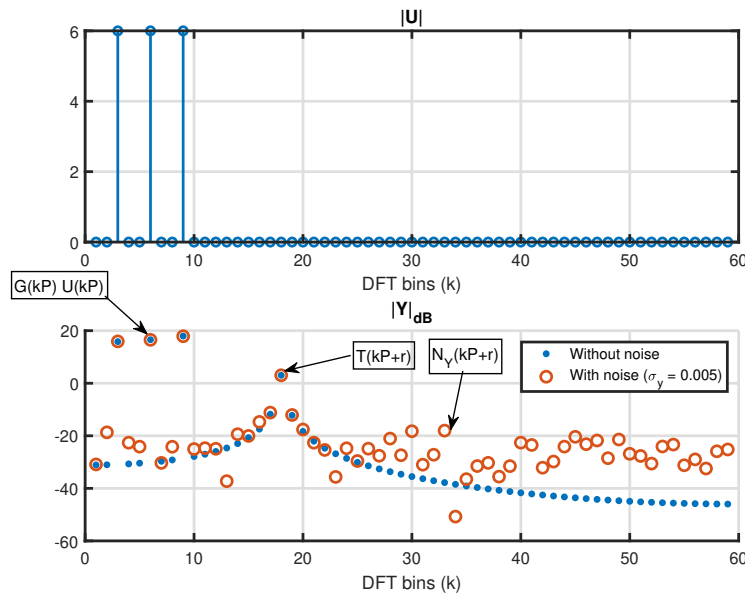


Figure 2.13 – Input and output of the DF system (2.4)

2.14.2 Algorithm

Now we can finally discuss the algorithm of the robust LPM. The idea is to estimate the contribution of the transient term at the excited frequencies by estimating it from the non-excited frequencies. To this end, we will work in a window around every excited DFT bin. In the case of the previous example, the excited bins are $k = 1, 2$ and 3 . Two parameters must be chosen by the user. The first one is the window size $2n$. We must use n unexcited bins before and after kP .

Example $k = 3, P = 3, n = 4$. We are working in a window around $Y(kP) = Y(9)$. We must use 4 unexcited bins before and after 9.

$$Y(9 + r_i) \text{ with } r_i = -5, -4, -2, -1, 1, 2, 4, 5$$

For every one of the unexcited lines the following holds

$$Y(kP + r_i) = T(kP + r_i) + N_Y(kP + r_i)$$

The transient term can be modeled as a polynomial in r . This is because the transient is a smooth function of the frequency. The order of the polynomial R is the second parameter that the user can choose.

$$T(kP + r) \approx T(kP) + \sum_{s=1}^R t_s(k) r^s$$

We can then find the polynomial of best fit through the points $Y(kP + r_i)$

Example continued Putting all the $Y(9 + r_i)$ into a vector and choosing $R = 2$ gives

$$\begin{bmatrix} Y(4) \\ Y(5) \\ Y(7) \\ Y(8) \\ Y(10) \\ Y(11) \\ Y(13) \\ Y(14) \end{bmatrix} = \begin{bmatrix} 1 & (-5) & (-5)^2 \\ 1 & (-4) & (-4)^2 \\ 1 & (-2) & (-2)^2 \\ 1 & (-1) & (-1)^2 \\ 1 & 1 & 1^2 \\ 1 & 2 & 2^2 \\ 1 & 4 & 4^2 \\ 1 & 5 & 5^2 \end{bmatrix} \begin{bmatrix} T(9) \\ t_1(3) \\ t_2(3) \end{bmatrix} + \begin{bmatrix} N_Y(4) \\ N_Y(5) \\ N_Y(7) \\ N_Y(8) \\ N_Y(10) \\ N_Y(11) \\ N_Y(13) \\ N_Y(14) \end{bmatrix} \rightarrow Y_n = K_n \Theta + V_n$$

The least squares solution is given by¹

$$\hat{\Theta} = (K_n^H K_n)^{-1} K_n^H Y_n$$

This results in an estimation of the transient term at the excited DFT bin $\hat{T}(9)$. The other terms $t_1(3)$ and $t_2(3)$ are not important.

Finally, the estimated transient term can be removed from the output spectrum at the excited DFT line.

$$\hat{Y}(kP) = Y(kP) - \hat{T}(kP)$$

Thus, the transient has been suppressed. The FRF can then be calculated simply as

$$\hat{G}(\Omega_k) = \frac{\hat{Y}(kP)}{U(kP)}$$

The entire procedure outlined here must be repeated for all $k \in K_{\text{exc}}$.

¹Note that calculating the solution like this results in an ill-conditioned problem. The \backslash operator in MATLAB solves this problem using QR-factorisation, which is better conditioned. ($\theta = K_n \backslash Y_n$)



2.14.3 Variance estimate

Having a variance estimate of the output spectrum is useful for providing uncertainty bounds. It can also be used as a nonparametric weighting in parametric identification of the system transfer function.

The residual is defined as the difference between the measured output spectrum and the predicted output spectrum.

$$\hat{V}_n = Y_n - K_n \hat{\Theta}$$

Assuming that the variance of the noise is white (flat) in the window $2n$, it can be used to estimate the variance at every $k \in \mathbf{K}_{\text{exc}}$.

$$\hat{\sigma}_Y^2(kP) = \frac{1}{q^{\text{noise}}} V_n^H V_n, \quad \text{with } q^{\text{noise}} = 2n - (R + 1) \text{ (degrees of freedom)}$$

A proof of this is given in appendix 2.C and is based on [2, Appendix 7.B]. The reason why we must divide by the degrees of freedom q^{noise} and not by $2n$ is because $R+1$ parameters are estimated. This is analogous to the reason why the unbiased sample variance is calculated by dividing by the number of observations - 1 when the population mean is also estimated.

As $\hat{G}(\Omega_k) = \hat{Y}(kP)/U(kP)$, the variance of the FRF can be calculated as

$$\hat{\sigma}_G^2(\Omega_k) = \frac{\hat{\sigma}_Y^2(kP)}{|U(kP)|^2} \quad (2.12)$$

2.14.4 Example

The robust LPM is applied to measurements from the system (2.4). This time $N = 16384$ and $P = 2$. The first $F = 5000$ frequencies are excited with a random phase multisine with an RMS value 1. White Gaussian noise is added to the output with $\sigma_y = 0.2$. 100 noise realizations are simulated and the root-mean square (RMS) error is calculated to get an idea of the effectiveness of the estimator.

$$\text{RMS}[|\hat{G} - G|](\Omega_k) = \sqrt{\frac{1}{100} \sum_{i=1}^{100} |\hat{G}^{(i)}(\Omega_k) - G(\Omega_k)|^2}$$

with $\hat{G}^{(i)}(\Omega_k)$ being the nonparametric estimate of $G(\Omega_k)$ for the i -th noise realization. The parameters used for the LPM are $R = 2$, $n = 6$ which results in $q^{\text{noise}} = 9$ degrees of freedom. To make the results more presentable, the data points are taken together in windows of size 50 and are averaged. The results are shown in figure 2.14. Not taking the transient into account is significantly worse around the resonance frequency of the system. The transient resembles the FRF of the system, which is why the error is most pronounced around the resonance frequency. **However, far away from the resonance frequency, not taking the transient into account is approximately 1 dB better than using the Robust LPM. This is because the transient term is below the random noise contribution. LPM uses a noisy estimate of the transient and this estimate is subtracted from the output spectrum, leading to an increased variance.** Finally, the robust LPM seems to be slightly better than windowing in this simulation.

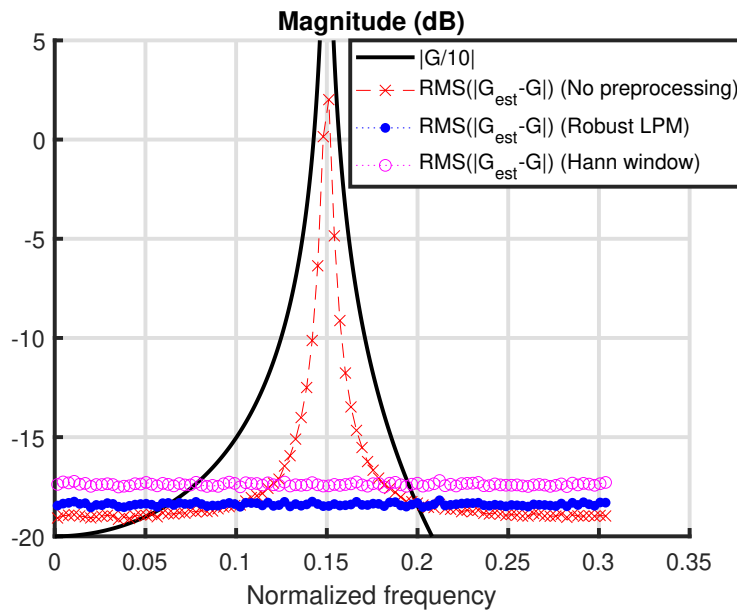


Figure 2.14 – Comparison of the RMS error for the nonparametric estimate acquired without preprocessing, with windowing and with the robust LPM. $R = 2$, $n = 6$, $q^{\text{noise}} = 9$.

2.14.5 Choice of the order and degrees of freedom

Two parameters can be chosen by the user when performing the robust LPM analysis: the order of the polynomial approximation and the degrees of freedom used to estimate the noise variance.

Order A good way to choose the order is to start at $R = 2$ and increment it in steps of two until the estimate of the variance $\hat{\sigma}_G^2$ stops decreasing.

Degrees of freedom Increasing the degrees of freedom q^{noise} will increase the window size $2n$. This gives a better estimation of the variance. There is a trade-off however: it was assumed that the noise variance is white (flat) in the window $2n$. Thus, making the window size too big will result in a loss of frequency resolution. **Additionally, making the window size $2n$ bigger means that the transient is approximated by a polynomial over a larger window. At that point it might be necessary to increase the order R of the polynomial.**

2.15 Generalization

Many things were simplified until now. A few of the assumptions that were made are:

- The input is noiseless.
- White Gaussian noise was simulated in the example, but what if the noise is filtered white noise?
- There is no feedback from output to input.
- The input is a multisine. What about random excitations?

Each of these points will be discussed briefly.

2.15.1 Noisy input

Of course, the measurement of the input can be noisy. Thankfully, the robust LPM is also able to estimate the input noise variance for periodic excitations. Given that the input is not known perfectly, the estimate of the FRF variance (2.12) is not correct anymore. In general, the variance of an FRF estimate $\hat{G}(\Omega_k)$ can be approximated by

$$\hat{\sigma}_{\hat{G}}^2(k) = |\hat{G}(\Omega_k)|^2 \left(\frac{\hat{\sigma}_{\hat{Y}_k}^2}{|\hat{Y}_k|^2} + \frac{\hat{\sigma}_{\hat{U}_k}^2}{|\hat{U}_k|^2} - 2\text{Re} \left(\frac{\hat{\sigma}_{\hat{Y}_k \hat{U}_k}^2}{\hat{Y}_k \hat{U}_k} \right) \right)$$

The equation above is only applicable when the excitation is periodic and when the FRF is estimated by dividing the output spectrum $Y(k)$ by the input spectrum $U(k)$. It is also possible to estimate the variance of the noise for arbitrary excitations, but different formulas must be used. [2, Section 2.6]

2.15.2 Filtered white noise

For simplicity sake, filtered white noise is defined as

$$v(n) = S(z^{-1})e(n), \quad \text{with } e(n) \sim \mathcal{N}(0, \sigma^2)$$

The power spectrum of this noise is not flat. This also means that the noise samples can be correlated over time.

$$\exists n, m \text{ with } n \neq m \text{ such that } \mathbb{E}\{v(n)v(m)\} \neq 0$$

An important consequence of filtered white noise is that there will also be noise transients in the measurements.

$$V(k) = S(\Omega_k)E(k) + T_s(\Omega_k)$$

As the input of the noise filter is random, the noise transients will never fade away. This means that when one waits long enough for the system $G(\Omega)$ to enter steady state, there will still be noise transients in the measurements. This is where the LPM can also be useful.

2.15.3 Feedback

Consider the measurement setup shown in figure 2.15. An LTI system $G(z^{-1})$ is in negative feedback and the output is perturbed by process noise.

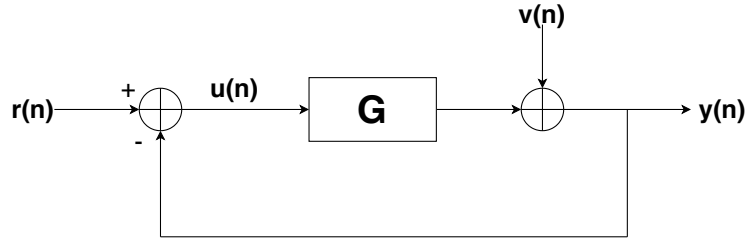


Figure 2.15 – LTI system in negative feedback with additive process noise.

The Z-transform of the output and input are given by

$$\begin{aligned} Y(z) &= \frac{G}{1+G}R(z) + \frac{1}{1+G}V(z) \\ U(z) &= \frac{1}{1+G}R(z) - \frac{1}{1+G}V(z) \end{aligned}$$

Noise that affects the output also affects the input due to feedback. Two possibilities will be considered: $r(n)$ is a multisine and $r(n)$ is a random signal. The statistical properties of the noise are

$$\mathbb{E}\{V(k)\} = 0 \text{ and } \mathbb{E}\{|V(k)|^2\} = \sigma_v^2(k)$$

$r(n)$ is a multisine In this case we can apply multiple periods P to the system. It is assumed that enough time has passed for the transient to fade away and the noise transients will be neglected.

$$\begin{aligned} Y^{(p)}(k) &= \frac{G(\Omega_k)}{1+G(\Omega_k)}R^{(p)}(k) + \frac{1}{1+G(\Omega_k)}V^{(p)}(k) \\ U^{(p)}(k) &= \frac{1}{1+G(\Omega_k)}R^{(p)}(k) - \frac{1}{1+G(\Omega_k)}V^{(p)}(k) \end{aligned}$$

The superscript $^{(p)}$ denotes the period of the measurement. The nonparametric FRF can then be estimated with

$$\hat{G}(\Omega_k) = \frac{\frac{1}{P} \sum_{p=0}^P Y^{(p)}(k)}{\frac{1}{P} \sum_{p=0}^P U^{(p)}(k)} \quad (2.13)$$

Taking the limit for $P \rightarrow \infty$ will allow us to establish whether this estimator is consistent.

$$\lim_{P \rightarrow \infty} \hat{G}(\Omega_k) = \frac{\mathbb{E}\{Y^{(p)}(k)\}}{\mathbb{E}\{U^{(p)}(k)\}} = G(\Omega_k)$$

The first equality applies the law of large numbers for independent experiments. And so this estimator is indeed consistent.

$r(n)$ is an arbitrary signal For arbitrary signals it is not advised to use the estimator (2.13) because a division by a small number is possible, resulting in an estimator that won't converge. **Another thing that must be considered when working with arbitrary signals is that arbitrary signals are aperiodic.** If we want to get a better estimate of the FRF we will need to cut the measurement into M pieces of N samples. If the total number of samples

NM is constant, this will results in a trade-off between the frequency resolution f_s/N and noise suppression. An estimator that is used for arbitrary signals is

$$\hat{G}(\Omega_k) = \frac{\frac{1}{M} \sum_{m=0}^M Y^{(m)}(k) \overline{U^{(m)}(k)}}{\frac{1}{M} \sum_{m=0}^M U^{(m)}(k) \overline{U^{(m)}(k)}} \quad (2.14)$$

Here there is no danger for the denominator to become small. This estimator is consistent when only the output is perturbed by noise. However, in our case the input is also perturbed. Taking the limit as $M \rightarrow \infty$ gives

$$\lim_{M \rightarrow \infty} \hat{G}(\Omega_k) = \frac{\mathbb{E}\{Y^{(m)}(k) \overline{U^{(m)}(k)}\}}{\mathbb{E}\{U^{(m)}(k) \overline{U^{(m)}(k)}\}} = \frac{S_{YU}(k)}{S_{UU}(k)}$$

with S_{YU} being the cross-power spectrum between the output and input and S_{UU} being the auto-power spectrum of the input. Given that

$$\mathbb{E}\{R(k)\} = 0 \text{ and } \mathbb{E}\{|R(k)|^2\} = \sigma_r^2(k)$$

and the fact that $R(k)$ and $V(k)$ are uncorrelated we get the following result

$$\lim_{P \rightarrow \infty} \hat{G}(\Omega_k) = \frac{G(\Omega_k) \sigma_r^2(k) - \sigma_v^2(k)}{\sigma_r^2(k) + \sigma_v^2(k)} \neq G(\Omega_k)$$

Thus, the estimator (2.14) is inconsistent.

Indirect method It is possible to get around the problem of the estimator (2.14) by also using the reference signal $r(n)$. Instead of modeling the transfer function from input (u) to output (y), we will model the transfer function from reference (r) to output (y) and from reference (r) to input (u).

$$\hat{G}(\Omega_k) = \frac{\frac{1}{P} \sum_{m=0}^P Y^{(m)}(k) \overline{R^{(m)}(k)}}{\frac{1}{P} \sum_{m=0}^P U^{(m)}(k) \overline{R^{(m)}(k)}} \quad (2.15)$$

Doing some calculations brings us to

$$\lim_{M \rightarrow \infty} \hat{G}(\Omega_k) = \frac{\mathbb{E}\{Y^{(m)}(k) \overline{R^{(m)}(k)}\}}{\mathbb{E}\{U^{(m)}(k) \overline{R^{(m)}(k)}\}} = \frac{S_{YR}(k)}{S_{UR}(k)} = G(\Omega_k)$$

Thus, by keeping the reference signal and using it, it is possible to get a consistent estimate when using arbitrary excitations.

2.16 Conclusion

Multisines can be used to identify nonparametric estimates of the FRF. The robust LPM allows for the transients to be suppressed. Even if the system is in steady state, the noise transients can still deteriorate the quality of the FRF estimate. This is a reason why the robust LPM can be quite effective.

It is useful to keep the reference signal that is applied to the system. When the system is excited by an arbitrary excitation, this information can be used to get a consistent estimate of the FRF even if both the input and the output are perturbed by noise.

Finally, the authors of [2] offer MATLAB code that can perform the robust LPM for periodic and arbitrary excitations. These MATLAB function can also identify the best linear approximation of a nonlinear system. These functions can also handle MIMO systems.

Appendix

2.A Transient term

Suppose that a DT LTI system is of the following form.

$$a_0 y(n) + a_1 y(n-1) + a_2 y(n-2) = b_0 u(n) + b_1 u(n-1) \quad (2.16)$$

The DFT of a sequence is actually just a windowed version of the Z-transform. To see this, the window is defined as

$$w(n) = \begin{cases} 1, & 0 \leq n < N \\ 0, & \text{otherwise} \end{cases}$$

The windowed Z-transform of a sequence $x(n)$ is then given by

$$\mathcal{Z}\{w(n)x(n)\} = \sum_{n=-\infty}^{+\infty} w(n)x(n)z^{-n} = \sum_{n=0}^{N-1} x(n)z^{-n}$$

Evaluating the expression above in $z = e^{j2\pi k/N}$ is exactly the DFT of $x(n)$.

$$\mathcal{Z}\{w(n)x(n)\}|_{z=e^{j2\pi k/N}} = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N} = \text{DFT}\{x(n)\}$$

Thus, taking the windowed Z-transform of both sides of (2.16) and evaluating it in $z = e^{j2\pi k/N}$ is the same as taking the DFT of both sides.

For simplicity, let's only consider one of the terms.

$$a_p w(n)y(n-p)$$

a_p is just a constant, so that can also be left out in the analysis.

$$\mathcal{Z}\{w(n)y(n-p)\} = \sum_{n=0}^{N-1} y(n-p)z^{-n}$$

After some manipulations:

$$\mathcal{Z}\{w(n)y(n-p)\} = z^{-p} \sum_{n=0}^{N-1} y(n)z^{-n} + \sum_{n=0}^{p-1} [y(n-p) - y(n-p+N)z^{-N}]z^{-n}$$

Evaluating this expression in $z = e^{j2\pi k/N}$ gives

$$\begin{aligned} \mathcal{Z}\{w(n)y(n-p)\}|_{z=e^{j2\pi k/N}} &= (e^{-j2\pi k/N})^p \sum_{n=0}^{N-1} y(n)e^{-j2\pi kn/N} \\ &\quad + \sum_{n=0}^{p-1} [y(n-p) - y(n-p+N)](e^{-j2\pi k/N})^n \end{aligned}$$

Note that z^{-N} disappears because $e^{-j2\pi Nn/N} = 1$. The first term contains the DFT of $y(n)$. The second term is a polynomial in $e^{-j2\pi k/N}$ that depends on $y(n-p) - y(n-p+N)$. In other words, it depends on the difference between samples from the previous period and samples from the current period.

$$\mathcal{Z}\{w(n)y(n-p)\}|_{z=e^{j2\pi k/N}} = (e^{-j2\pi k/N})^p Y(k) + I_{y,p}(e^{-j2\pi k/N})$$



Applying this to (2.16) gives

$$Y(k) \left(\sum_{p=0}^2 a_p (e^{-j2\pi k/N})^p \right) = U(k) \left(\sum_{p=0}^1 b_p (e^{-j2\pi k/N})^p \right) + I(e^{-j2\pi k/N}) \quad (2.17)$$

with

$$I(e^{-j2\pi k/N}) = \sum_{p=0}^1 b_p I_{u,p}(e^{-j2\pi k/N}) - \sum_{p=0}^2 a_p I_{y,p}(e^{-j2\pi k/N}) \quad (2.18)$$

Notice that

$$G(z^{-1})|_{z=e^{j2\pi k/N}} = G(e^{-j2\pi k/N}) = \frac{B(e^{-j2\pi k/N})}{A(e^{-j2\pi k/N})} = \frac{\sum_{p=0}^1 b_p (e^{-j2\pi k/N})^p}{\sum_{p=0}^2 a_p (e^{-j2\pi k/N})^p}$$

Dividing (2.17) by $A(e^{-j2\pi k/N})$ then gives the final form

$$\boxed{Y(k) = G(e^{-j2\pi k/N})U(k) + T(k)} \quad (2.19)$$

with

$$T(k) = \frac{I(e^{-j2\pi k/N})}{A(e^{-j2\pi k/N})}$$

2.B DFT of white noise

A white noise sequence $v(n)$ has the following properties

$$\begin{aligned}\mathbb{E}\{v(n)\} &= 0 \\ \mathbb{E}\{v(n)v(m)\} &= \sigma^2\delta(n-m)\end{aligned}$$

The DFT of this white noise sequence is

$$V(k) = \sum_{n=0}^{N-1} v(n)e^{-j2\pi kn/N}$$

$\mathbb{E}\{V(k)\}$ The expected value of $V(k)$ is

$$\mathbb{E}\{V(k)\} = \sum_{n=0}^{N-1} \mathbb{E}\{v(n)\}e^{-j2\pi kn/N} = \sum_{n=0}^{N-1} 0e^{-j2\pi kn/N} = 0$$

$\mathbb{E}\{V(k)\overline{V(l)}\}$ The expected value of $V(k)\overline{V(l)}$ is

$$\begin{aligned}\mathbb{E}\{V(k)\overline{V(l)}\} &= \mathbb{E}\left\{\sum_{n=0}^{N-1} v(n)e^{-j2\pi kn/N} \sum_{m=0}^{N-1} v(m)e^{j2\pi lm/N}\right\} \\ &= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \mathbb{E}\{v(n)v(m)\}e^{-j2\pi kn/N}e^{j2\pi lm/N} \\ &= \sigma^2 \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \delta(n-m)e^{-j2\pi kn/N}e^{j2\pi lm/N} \\ &= \sigma^2 \sum_{n=0}^{N-1} e^{-j2\pi(k-l)n/N} = \begin{cases} N\sigma^2 & \text{if } \text{mod}(k-l, N) = 0 \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

When $k = l$, this result becomes

$$\mathbb{E}\{|V(k)|^2\} = N\sigma^2$$

$\mathbb{E}\{V(k)V(l)\}$ The expected value of $V(k)V(l)$ is

$$\begin{aligned}\mathbb{E}\{V(k)V(l)\} &= \mathbb{E}\left\{\sum_{n=0}^{N-1} v(n)e^{-j2\pi kn/N} \sum_{m=0}^{N-1} v(m)e^{-j2\pi lm/N}\right\} \\ &= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \mathbb{E}\{v(n)v(m)\}e^{-j2\pi kn/N}e^{-j2\pi lm/N} \\ &= \sigma^2 \sum_{n=0}^{N-1} e^{-j2\pi(k+l)n/N} = \begin{cases} N\sigma^2 & \text{if } \text{mod}(k+l, N) = 0 \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

When $k = l$ and $\text{mod}(2k, N) \neq 0$, this result becomes

$$\mathbb{E}\{V^2(k)\} = 0$$

2.C Covariance estimation

The output spectrum in a window of size $2n$ around an excited frequency line is given by

$$Y_n = K_n \Theta + V_n$$

The least squares solution is given by

$$\hat{\Theta} = (K_n^H K_n)^{-1} K_n^H Y_n$$

A key assumption used to estimate the covariance $C_Y(kP)$ is that V_n is assumed to have a flat power spectrum in the window $2n$.

The residual is the difference between the measured spectrum and the predicted spectrum.

$$\hat{V}_n = Y_n - K_n \hat{\Theta} = (I_{2n} - K_n (K_n^H K_n)^{-1} K_n^H) Y_n = P_n Y_n$$

Using the fact that $P_n K_n = K_n - K_n = 0$ we get

$$\hat{V}_n = P_n V_n$$

Next up, we want to see how $\hat{V}_n^H \hat{V}_n$ relates to $V_n^H V_n$.

$$\hat{V}_n^H \hat{V}_n = V_n^H P_n^H P_n V_n$$

First, it is easy to see that $P_n^H = P_n$. Next, it turns out that P_n is an idempotent matrix.

$$\begin{aligned} P_n P_n &= (I_{2n} - K_n (K_n^H K_n)^{-1} K_n^H) (I_{2n} - K_n (K_n^H K_n)^{-1} K_n^H) \\ &= I_{2n} - 2K_n (K_n^H K_n)^{-1} K_n^H + K_n (K_n^H K_n)^{-1} K_n^H = P_n \end{aligned}$$

Thus, we get

$$\hat{V}_n^H \hat{V}_n = V_n^H P_n V_n = \text{trace}(V_n^H P_n V_n) = \text{trace}(P_n V_n V_n^H) \quad (2.20)$$

In this step we used the fact that the trace of a scalar is a scalar and the fact that matrices in a trace can be circularly permuted ($\text{trace}(ABC) = \text{trace}(BCA)$). It is assumed that V_n has a flat power spectrum in the window $2n$. This means that

$$\mathbb{E}\{V_n V_n^H\} = \sigma_Y^2(kP) I_{2n}$$

Plugging this into (2.20) and taking the expected value gives

$$\mathbb{E}\{\hat{V}_n^H \hat{V}_n\} = \sigma_Y^2(kP) \text{trace}(P_n) \quad (2.21)$$

P_n is an idempotent matrix, meaning that its eigenvalues can only be 0 or 1. Additionally, because P_n and K_n are each other's orthogonal complement, the rank of P_n is related to the rank of K_n .

$$K_n P_n = 0 \Rightarrow \text{rank}(P_n) = 2n - \text{rank}(K_n)$$

If K_n is full column rank and if K_n has more rows than columns, then the rank of K_n is equal to the number of columns in K_n .

$$\text{rank}(K_n) = R + 1$$

The trace of a matrix is the sum of the eigenvalues of a matrix. The sum of the eigenvalues of P_n is exactly equal to the rank of P_n , because the rank of P_n is equal to the number of nonzero eigenvalues and because the eigenvalues can only be 0 or 1.

$$\text{trace}(P_n) = \text{rank}(P_n) = 2n - (R + 1) = q^{\text{noise}}$$

Finally, this explains why $\hat{V}_n^H \hat{V}_n$ must be divided by q^{noise} to get an unbiased estimate of the covariance.

$$\boxed{\sigma_Y^2(kP) = \frac{\mathbb{E}\{\hat{V}_n^H \hat{V}_n\}}{q^{\text{noise}}}}$$

Chapter 3

Model reference control

3.1 Introduction

The first part of this chapter is a summary of the work presented in [4]. From section 3.6 and on, improvements are made to the existing methods.

3.2 Problem statement

The first assumption that is made is that $G(\Omega)$ is stable and minimum-phase. It is also possible to extend this theory to unstable nonminimum-phase systems. This will be done later in section 3.9.

The system is controlled by an unknown controller $K(\Omega, \rho)$ in closed loop (CL).

$$\text{CL}(\Omega) = \frac{K(\Omega, \rho)G(\Omega)}{1 + K(\Omega, \rho)G(\Omega)}$$

$\rho = [\rho_1 \ \dots \ \rho_{n_\rho}]^T$ is a vector containing the optimization parameters. In this work $K(\Omega, \rho)$ is linear in the parameters.

$$K(\Omega, \rho) = \beta(\Omega)\rho \quad (3.1)$$

with $\beta(\Omega)$ being a row vector with n_ρ elements. The idea of model reference control, is to get the closed loop system “as close” as possible to a user-defined reference system $M(\Omega)$. user ✕

$$M(\Omega) \approx \frac{K(\Omega, \rho)G(\Omega)}{1 + K(\Omega, \rho)G(\Omega)}$$

$M(\Omega)$ is assumed to be a stable causal LTI system. This criterion can be quantified by using the 2-norm of a transfer function.

$$J_{mr}(\rho) = \left\| F(\Omega) \left[M(\Omega) - \frac{K(\Omega, \rho)G(\Omega)}{1 + K(\Omega, \rho)G(\Omega)} \right] \right\|_2^2 \quad (3.2)$$

$F(\Omega)$ is a user-defined weighing filter that can be chosen to highlight specific frequencies. The 2-norm of a SISO system is defined differently for CT and DT systems. For CT systems it is

$$\|H(s)\|_2^2 = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |H(j\omega)|^2 d\omega$$

and for DT systems it is

$$\|H(z^{-1})\|_2^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H(e^{j\omega})|^2 d\omega$$