Lecture 8

Multivariable and Decentralized Control*

There is a clear trend in modern engineering toward systems of higher and higher complexity. One reason is that demands for efficiency give tighter interconnections between subsystems. For example, to get minimally pollutive emissions in the exhaust gas of a car, it is necessary that engine, carburettor, catalyst, gearbox, etc. all cooperate in an optimal manner. Similarly, the demand for efficient production and distribution of electrical power has led to tighter coupling between production units in different geographical regions and more complex large scale dynamics.

A system with several inputs and several outputs is sometimes called a MIMO (Multiple-Input-Multiple-Output) system. Control theory for such systems is a highly active research area and a review of the available methods is outside the scope of this course. However, many of the ideas that were developed for scalar systems can be easily adapted also to a multivariable setting. This lecture will present a few such items:

- Multivariable performance specifications
- Limitations due to unstable multivariable zeros
- Decentralized control by pairing of signals

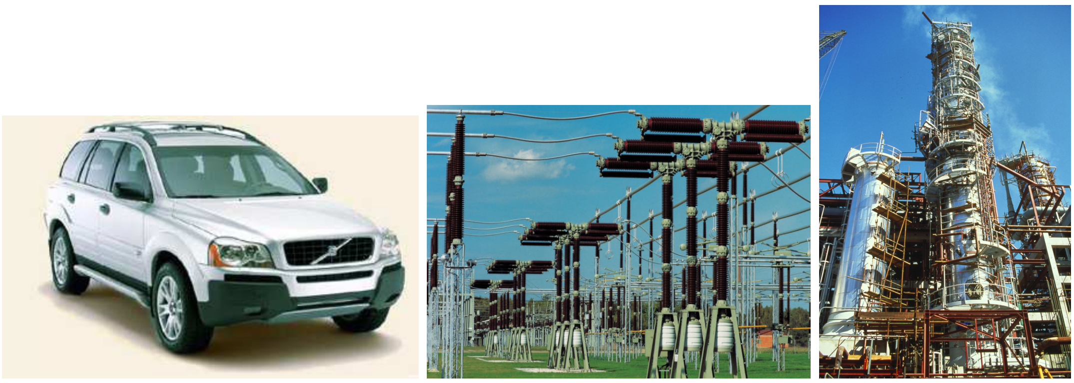

Figure 8.1 A modern car, a power plant and an oil refinery all make extensive use of multivariable control systems

8.1 Multivariable specifications

Consider again the feedback loop illustrated in Figure 8.2 but assume that all signals are vector valued. Then $P(s)$, $F(s)$ and $C(s)$ are matrices, so the closed-loop transfer functions

*Written by A. Rantzer. Some of the examples come from the book by Glad and Ljung, *Control Theory — Multivariable and Nonlinear Methods*, Taylor & Francis, 2000

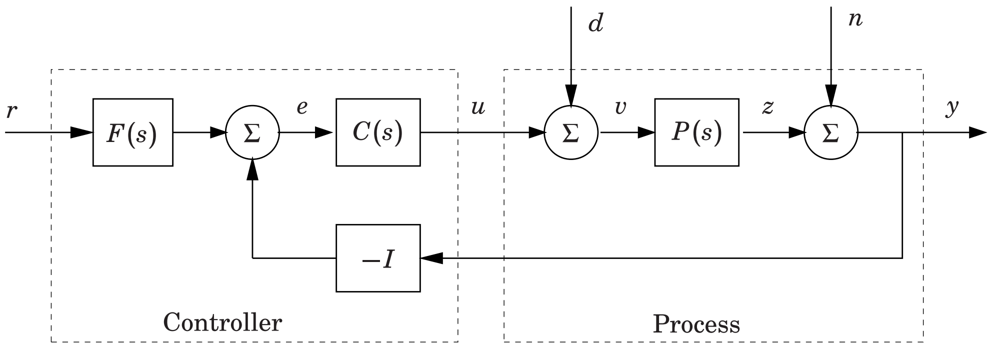

Figure 8.2 A multivariable control loop

need to be derived with some more care:

$$\begin{aligned} Z(s) &= PCF \cdot R(s) + P \cdot D(s) - PC \cdot [N + Z](s) \\ [I + PC]Z(s) &= PCF \cdot R(s) + P \cdot D(s) - PC \cdot N(s) \\ Z(s) &= [I + PC]^{-1}PCF \cdot R(s) + [I + PC]^{-1}P \cdot D(s) - [I + PC]^{-1}PC \cdot N(s) \end{aligned}$$

where a dot denotes matrix-vector multiplication.

Similarly

$$V(s) = [I + CP]^{-1}CF \cdot R(s) - [I + CP]^{-1}C \cdot N(s) + [I + CP]^{-1} \cdot D(s)$$

Notice that $S = [I + PC]^{-1}$ is generally not the same as $[I + CP]^{-1}$. The first is called the (output) sensitivity function and has a matrix size determined by the number of outputs. The second is called *input sensitivity function* and its size corresponds to the number of inputs. The convention is to call $T = [I + PC]^{-1}PC$ the complementary sensitivity function. Notice the following identities:

$$\begin{aligned} [I + PC]^{-1}P &= P[I + CP]^{-1} \\ C[I + PC]^{-1} &= [I + CP]^{-1}C \\ T &= P[I + CP]^{-1}C = PC[I + PC]^{-1} \\ S + T &= I \end{aligned}$$

The first equality follows by multiplication with $I + CP$ from the right and $I + PC$ from the left. The second one is analogous. Using the first two equalities, we immediately get the third. The last one is straight from definitions as well.

Also for multivariable systems, it is common to require S to be small at low frequencies and T to be small at high frequencies. The first specification means that y follows Fr well at small frequencies, while the second means that high frequency measurement noise n does not influence x significantly. Another way to state these requirements is to say that the *loop transfer matrix*

$$P(i\omega)C(i\omega)$$

should have small norm $\|P(i\omega)C(i\omega)\|$ at high frequencies, while at low the frequencies instead $\|[P(i\omega)C(i\omega)]^{-1}\|$ should be small. See Figure 8.3.

8.2 Limitations due to unstable zeros

Just like for scalar systems, there are fundamental limitations on the achievable performance in a multivariable system. For a multivariable system with square tranfer matrix $P(s)$, i.e. the same number of inputs and outputs, the zeros can be defined as the poles of $P(s)^{-1}$. The following theorem captures the influence of an unstable zero:

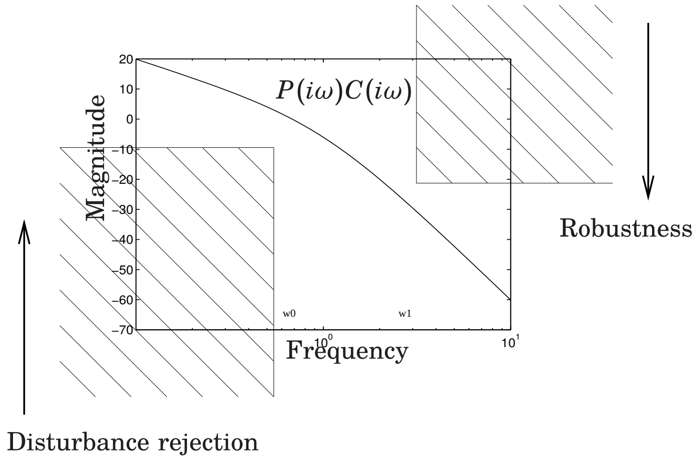

Figure 8.3 Specifications on the singular values of the loop transfer function often have this form. A lower bound on the singular values for low frequencies is needed for disturbance rejection. This means that $\|[P(i\omega)C(i\omega)]^{-1}\|$ should be small. An upper bound on the singular values for high frequencies, making $\|P(i\omega)C(i\omega)\|$ small enough, is needed for robustness to model errors and measurement noise.

THEOREM 8.1
Let $W_S(s)$ be scalar and stable and let $S(s) = [I + P(s)C(s)]^{-1}$ be the sensitivity function of a stable closed-loop system. Then, the specification

$$\|W_S S\|_\infty = \sup_\omega \overline{\sigma}\big(W_S(i\omega)S(i\omega)\big) \leq 1$$

is impossible to satisfy unless $|W_S(z)| \leq 1$ for every unstable zero z of $P(s)$. □

The proof is analogous to the result for scalar systems in Lecture 4. Instead of giving the details, we turn our attention to an example.

Example 1 (Non-minimum phase MIMO System) Consider a feedback system $y = (I+PC)^{-1}r$ with the multivariable process

$$P(s) = \begin{bmatrix} \dfrac{2}{s+1} & \dfrac{3}{s+2} \\ \dfrac{1}{s+1} & \dfrac{1}{s+1} \end{bmatrix}$$

Computing the determinant

$$\det P(s) = \frac{2}{(s+1)^2} - \frac{3}{(s+2)(s+1)} = \frac{-s+1}{(s+1)^2(s+2)}$$

shows that the process has an unstable zero at $s = 1$, which will limit the achievable performance. For further understanding of the limitation, consider the following three different control structures:

Controller 1 The controller

$$C_1(s) = \begin{bmatrix} \dfrac{K_1(s+1)}{s} & -\dfrac{3K_2(s+0.5)}{s(s+2)} \\ -\dfrac{K_1(s+1)}{s} & \dfrac{2K_2(s+0.5)}{s(s+1)} \end{bmatrix}$$

gives the diagonal loop transfer matrix

$$P(s)C_1(s) = \begin{bmatrix} \dfrac{K_1(-s+1)}{s(s+2)} & 0 \\ 0 & \dfrac{K_2(s+0.5)(-s+1)}{s(s+1)(s+2)} \end{bmatrix}$$

Hence the system is decoupled into to scalar loops, each with an unstable zero at $s = 1$ that limits the bandwidth. The closed-loop step responses are shown in Figure 8.4.

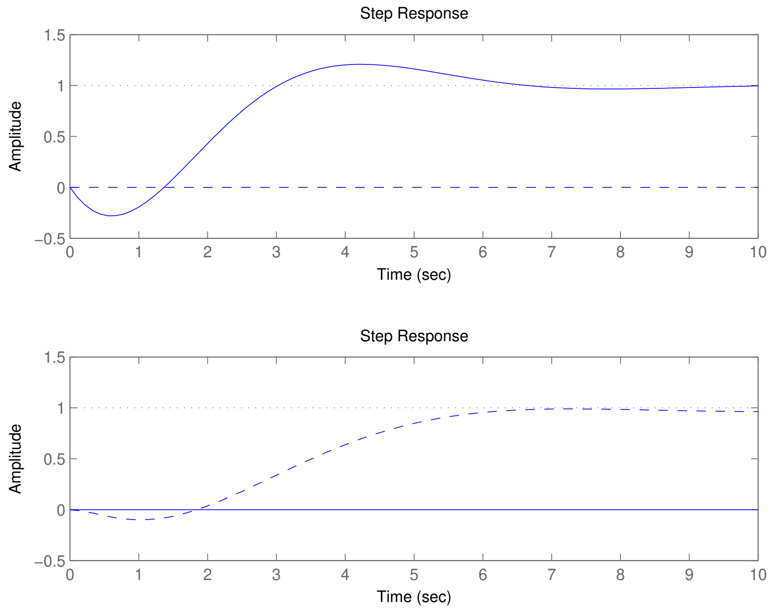

Figure 8.4 Closed-loop step responses with decoupling controller $C_1(s)$ for the two outputs y_1 (solid) and y_2 (dashed). The upper plot is for a reference step for y_1. The lower plot is for a reference step for y_2.

Controller 2 The controller

$$C_2(s) = \begin{bmatrix} \dfrac{K_1(s+1)}{s} & K_2 \\ -\dfrac{K_1(s+1)}{s} & K_2 \end{bmatrix}$$

gives the upper triangular loop transfer matrix

$$P(s)C_2(s) = \begin{bmatrix} \dfrac{K_1(-s+1)}{s(s+2)} & \dfrac{K_2(5s+7)}{(s+2)(s+1)} \\ 0 & \dfrac{2K_2}{s+1} \end{bmatrix}$$

Now the decoupling is only partial: Output y_2 is not affected by r_1. Moreover, there is no unstable zero that limits the rate of response in y_2! The closed-loop step responses for $K_1 = 1$, $K_2 = 10$ are shown in Figure 8.5.

Controller 3 The controller

$$C_3(s) = \begin{bmatrix} K_1 & \dfrac{-3K_2(s+0.5)}{s(s+2)} \\ K_1 & \dfrac{2K_2(s+0.5)}{s(s+1)} \end{bmatrix}$$

gives the lower triangular loop transfer matrix

$$P(s)C_3(s) = \begin{bmatrix} \dfrac{K_1(5s+7)}{(s+1)(s+2)} & 0 \\ \dfrac{2K_1}{s+1} & \dfrac{K_2(-1+s)(s+0.5)}{s(s+1)^2(s+2)} \end{bmatrix}$$

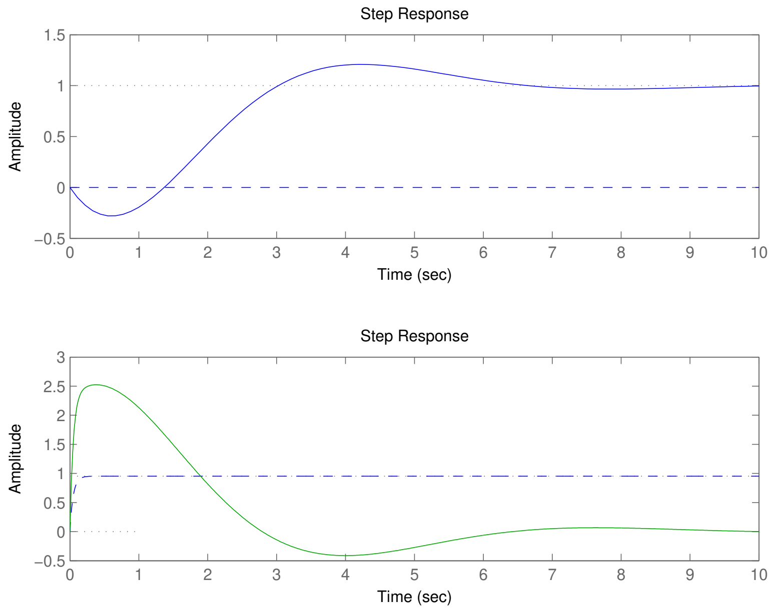

Figure 8.5 Closed-loop step responses with controller $C_2(s)$ for the two outputs y_1 (solid) and y_2 (dashed). The right half plane zero does not prevent a fast y_2-response to r_2 but at the price of a simultaneous undesired response in y_1.

In this case y_1 is decoupled from r_2 and can respond arbitrarily fast for high values of K_1, at the expense of bad behavior in y_2. Step responses for $K_1 = 10$, $K_2 = -1$ are shown in Figure 8.6.

To summarize, the example shows that even though a multivariable unstable zero always gives a performance limitation, it is possible to influence where the effects should show up. □

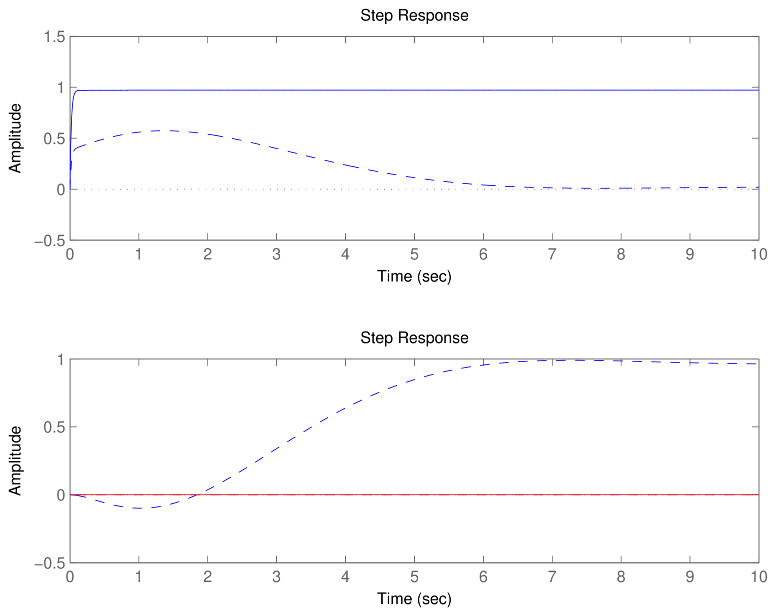

Figure 8.6 Closed-loop step responses with controller $C_3(s)$ for the two outputs y_1 (solid) and y_2 (dashed). The right half plane zero does not prevent a fast y_1-response to r_1 but at the price of a simultaneous undesired response in y_2.

8.3 Pairing of signals

The simplest way to deal with a multivariable control problem is to select an equal number of inputs and outputs and make a pairing, so that each input is responsible for control of one particular output. The corresponding transfer function from input to output of the process is determined for each pair, and the scalar feedback loops are designed ignoring the coupling between the loops inside the plant.

When connecting all the feedback loops simultaneously, the cross-coupling may potentially lead to performance degradation or even instability. The approach therefore works much better if there is a good way to select input and output variables for the scalar loops. One way to do this is to use the Relative Gain Array (RGA) that will be studied next.

For a full rank matrix $A \in \mathbb{C}^{m\times n}$, define the *Relative Gain Array* as

$$\mathrm{RGA}(A) := A\,.\!*(A^{-1})^T$$

where "$.*$" denotes element-by-element multiplication. For non-square matrices, the inverse A^{-1} is replaced by the pseudo-inverse $A^\dagger$ (in Matlab `pinv(A)`). If $m > n$, then $A^\dagger = (A^*A)^{-1}A^*$ and if $m < n$ then $A^\dagger = A^*(AA^*)^{-1}$. The Relative Gain Array has several interesting properties:

- The sum of all elements in a column or row is one.
- Permutations of rows or columns in A give the same permutations in $\mathrm{RGA}(A)$
- $\mathrm{RGA}(A) = \mathrm{RGA}(D_1AD_2)$ if D_1 and D_2 are diagonal, i.e. $\mathrm{RGA}(A)$ is independent of scaling
- If A is triangular, then $\mathrm{RGA}(A)$ is the unit matrix I.

Furthermore, the Relative Gain Array has an interpretation related to control theory: Let $P(s)$ be the transfer matrix from u to y. Then

- The (k, j) element of P is the transfer function $u_j \to y_k$ when $u_i = 0$ for $i \neq j$ (*open loop control*)
- The (j, k) element of P^{-1} is the inverse of the transfer function $u_j \to y_k$ when the other inputs are such that $y_i = 0$ for $i \neq k$ (*closed-loop control*)

Hence, if the (k, j) element of $\mathrm{RGA}(P(s))$ is equal to one, then the value of transfer function from u_j to y_k does not depend on whether the remaining inputs operate in open loop or closed loop. This indicates that the cross-coupling with other loops is weak.

The following rules of thumb use RGA to identify input-output pairings that have small cross-coupling. However, the outcome must always be evaluated using other tools.

1. Find a permutation of inputs and outputs that brings $\mathrm{RGA}(P(i\omega_c))$ as close as possible to the identity matrix. Here ω_c is the closed-loop bandwidth of the system.
2. Avoid pairings that give negative diagonal elements of $\mathrm{RGA}(P(0))$

The second rule is motivated by the desire to have the same *sign* of the static gain in one loop regardless if the other loops are closed or not. We illustrate the rules with an example.

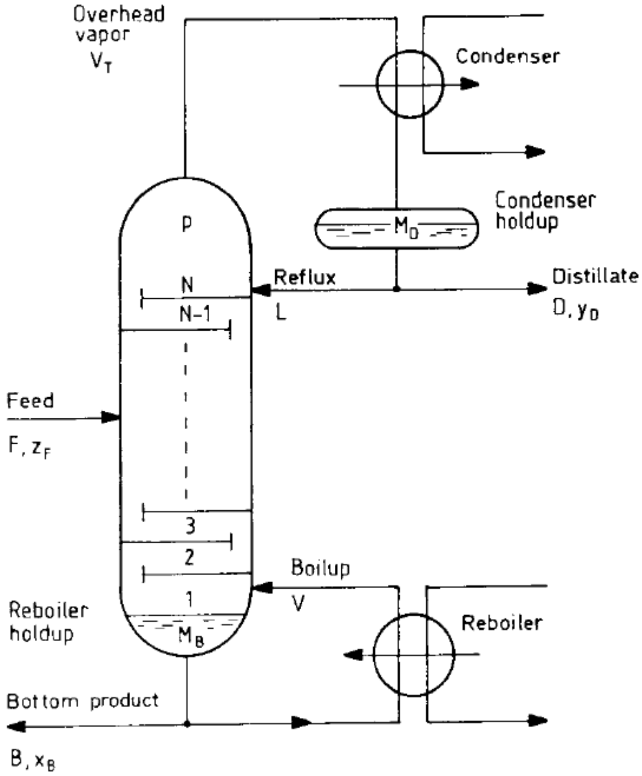

Figure 8.7 Schematic picture of a distillation column

Example: A Distillation Column

A distillation column is used in chemical process industry to separate different components in a chemical product. This example comes from Shell and describes a column where raw oil is separated into various petro-chemical derivatives. Heated raw oil is inserted at the bottom of the column and evaporates. Subcomponents then condense and get extracted at different levels of the column. The outputs are the top draw composition (y_1) and the side draw composition (y_2), and the inputs are the top draw flowrate (u_1), the side draw flowrate (u_2), and the bottom temperature (u_3). A linear model of the plant is given by

$$\begin{bmatrix} Y_1(s) \\ Y_2(s) \end{bmatrix} = \underbrace{\begin{bmatrix} \dfrac{4}{50s+1}e^{-27s} & \dfrac{1.8}{60s+1}e^{-28s} & \dfrac{5.9}{50s+1}e^{-27s} \\ \dfrac{5.4}{50s+1}e^{-18s} & \dfrac{5.7}{60s+1}e^{-14s} & \dfrac{6.9}{40s+1}e^{-15s} \end{bmatrix}}_{P(s)} \begin{bmatrix} U_1(s) \\ U_2(s) \\ U_3(s) \end{bmatrix}$$

Computing the RGA for the distillation column model gives

$$P(0) = \begin{bmatrix} 4.0 & 1.8 & 5.9 \\ 5.4 & 5.7 & 6.9 \end{bmatrix}$$

$$P(i/50) = \begin{bmatrix} 0.6871 - 2.7437i & 0.1548 - 1.1419i & 1.0135 - 4.0469i \\ 1.5758 - 3.4781i & 1.4704 - 3.3397i & 3.0247 - 4.4589i \end{bmatrix}$$

$$\mathrm{RGA}(P(0)) = \begin{bmatrix} 0.2827 & -0.6111 & 1.3285 \\ 0.0134 & 1.5827 & -0.5962 \end{bmatrix}$$

$$\mathrm{RGA}(P(i/50)) = \begin{bmatrix} 0.4355 - 0.3667i & -0.6536 - 0.0171i & 1.2181 + 0.3839i \\ -0.0906 + 0.3667i & 1.5933 + 0.0171i & -0.5027 - 0.3839i \end{bmatrix}$$

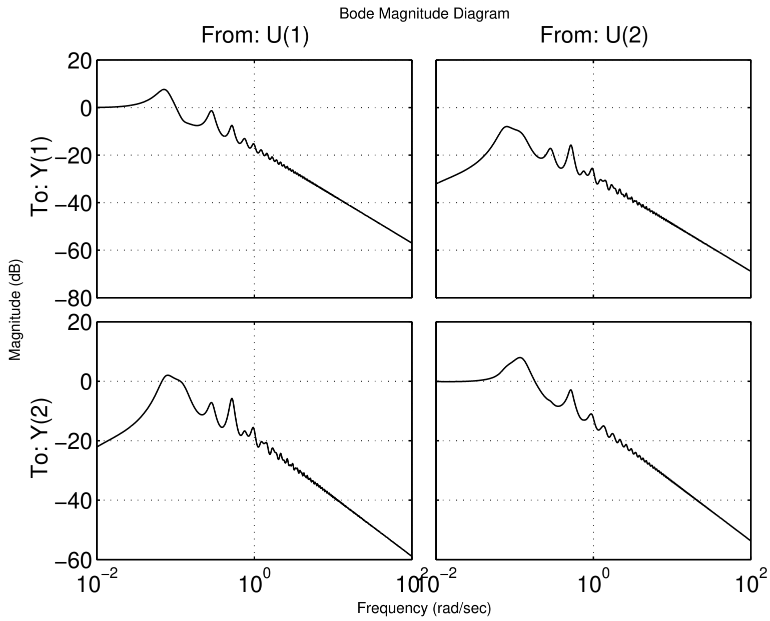

Figure 8.8 Magnitude plots for the closed-loop transfer function of the distillation column

To choose control signal for y_1, we apply the rules of thumb to the top row. This suggests the bottom temperature u_3 for control of the top draw composition y_1, since the third column has values slightly closer to 1 than the first column.

Based on the bottom row, we choose the side draw flowrate u_2 to control the side draw composition y_2. The top draw flow rate u_1 is left unused. The matrix transfer function $\tilde{P}(s)$ from (u_3, u_2) to (y_1, y_2) is now to be controlled by a diagonal controller $C(s)$, say a PI controller:

$$\tilde{P}(s) = \begin{bmatrix} \dfrac{5.9e^{-27s}}{50s+1} & \dfrac{1.8e^{-28s}}{60s+1} \\ \dfrac{6.9e^{-15s}}{40s+1} & \dfrac{5.7e^{-14s}}{60s+1} \end{bmatrix} \qquad C(s) = \begin{bmatrix} \dfrac{60s+1}{50s} & 0 \\ 0 & \dfrac{60s+1}{50s} \end{bmatrix}$$

Without feedforward, the closed-loop transfer matrix from reference to output becomes

$$PC(I + PC)^{-1}$$

and the Bode magnitude plots for the four transfer functions are given in Figure 8.8. As seen in the plots, the resulting cross-coupling is generally small and there is no static error.