

Control structure design for dynamic systems

Citation for published version (APA):

Wal, van de, M. M. J. (1994). *Control structure design for dynamic systems: a review*. (DCT rapporten; Vol. 1994.085). Technische Universiteit Eindhoven.

Document status and date:

Published: 01/01/1994

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

**Control structure design
for dynamic systems:**
A review

Marc van de Wal

WFW report 94.084

M.M.J. VAN DE WAL
Faculty of Mechanical Engineering
Eindhoven University of Technology
September 1994

Control structure design for dynamic systems:
A review

Marc van de Wal
Faculty of Mechanical Engineering
Eindhoven University of Technology

September 1994

Summary

This report presents a review of the available literature on Control Structure Design (CSD) for dynamic systems. In control system design, CSD is preceding the actual controller design and is defined as the stage in which decisions are made on the number, place and type of actuators (inputs) and sensors (outputs) to be used (Input/Output (IO) selection phase) and on the interconnections between measured and manipulated variables (Control Configuration (CC) selection phase). CSD has only been paid limited attention to, although an appropriate selection and pairing of measured and manipulated variables is as important as controller design itself: a wrong choice for the controller structure may put fundamental limitations on the system's performance, which cannot be overcome by advanced controller design. Moreover, the *complexity* of a control system is largely determined by the underlying control structure. For these reasons, CSD is a very important issue in modern control system design.

The main purpose of this exploratory study is to get an overview of the work that has already been performed in this area. Since CSD is particularly important for large-scale systems, most of the literature on CSD is published in the area of process control. However, applications in aircraft control and control of mechatronic systems have also been encountered. Unfortunately, the literature on CSD is largely restricted to *linear* control systems.

In this report, eleven different approaches for the IO-selection phase and twelve for the CC-selection phase are shortly discussed and compared. A simultaneous solution of both stages in CSD has not been found in literature; IO-selection and CC-selection are always treated successively. The various methods are assessed for some aspects which are practically important, *e.g.*, is the method generally applicable?, is it independent of the controller?, is it efficient/effective?, does it account for robust stability/performance?, is the theory well developed? Only a few of the methods discussed appear to be directly applicable to *nonlinear* control systems. Unfortunately, these methods seem not very effective, so additional selection criteria have to be applied before an appropriate control structure selection is possible.

The main conclusion of this report is that intensive further research has to be done in the area of CSD for both nonlinear *and* linear control systems, since none of the currently available CSD methods seems to be completely satisfactory.

Contents

Summary	i
Notation	iv
1 Introduction	1
2 Preliminaries	6
3 Criteria for selection of inputs and outputs	7
3.1 Control power and speed	7
3.2 Locations of poles and zeros	7
3.3 Controllability and observability	8
3.4 Cause-and-effect graphs	11
3.5 Achievable performance	11
3.6 Accuracy of state estimates	13
3.7 Economics	14
3.8 Morari resiliency index	14
3.9 Condition number	15
3.10 Singular value decomposition	17
3.11 Structured singular value	20
4 Criteria for selection of the control configuration	23
4.1 Stability of fixed eigenvalues	23
4.2 Relative degree	24
4.3 Achievable performance	25
4.4 Relative gain	25
4.5 Relative sensitivity	30
4.6 Closed-loop disturbance gain	31
4.7 Interaction potential	32
4.8 Numerical invertibility	34
4.9 Performance degradation	35
4.10 Nominal stability and closed-loop integrity	37
4.11 Singular value decomposition	39
4.12 Structured singular value	39
5 Applications of control structure design	41
5.1 Applications from literature	41
5.2 Proposal for a vehicle control example	44
6 Comparison	47
7 Recommendations for future research	51

Bibliography

A Tractor-semitrailer model

Notation

Symbols

a	column
a_i	i -th element of a
A	matrix or transfer function
a_{ij}	ij -th element of A
A_{ij}	ij -th block ("subsystem") of A
$\text{diag}[a_{ii}]$	diagonal matrix with elements a_{ii} (diagonal elements of A)
$\text{block diag}[A_{ii}]$	block diagonal matrix with blocks A_{ii} (diagonal blocks of A)
A^{-1}	inverse of A
A^H	complex conjugate transpose of A
a^T, A^T	transpose of a, A
\tilde{A}	perturbed matrix/transfer function A
\tilde{a}_{ij}	perturbed element a_{ij}
\dot{a}	first order time derivative of a
$\bar{\sigma}(A)$	maximum singular value of A
$\underline{\sigma}(A)$	minimum singular value of A
$\ a\ _l$	l -norm of a

Abbreviations

BRG	Block Relative Gain
CC	Control Configuration
CLDG	Closed Loop Disturbance Gain
CSD	Control Structure Design
DBRG	Dynamic Block Relative Gain
DCLI	Decentralized Closed Loop Integrity
DIC	Decentralized Integral Controllability
(D)NBRG	(Dynamic) Nonlinear Block Relative Gain
DOF	Degree(s) Of Freedom
FDLTI	Finite Dimensional Linear Time Invariant
IMC	Internal Model Control
IO	Input/Output
LQ(G)	Linear Quadratic (Gaussian)
MILP	Mixed Integer Linear Program
MIMO	Multi Input Multi Output
MRI	Morari Resiliency Index
NI	Niederlinski Index
NMP	NonMinimum Phase
PID	Proportional Integral Differential
PRGA	Performance Relative Gain Array
RGA	Relative Gain Array

RHP	Right Half Plane
SISO	Single Input Single Output
SSV	Structured Singular Value
SVD	Singular Value Decomposition

Chapter 1

Introduction

This report presents a survey of recent literature in the field of Control Structure Design (CSD) for dynamical systems. The fundamentals of the various methods are explained. This study is performed in the initial stage of a research into CSD for *nonlinear* dynamical systems. The main purpose of this exploratory study is therefore, to get an overview of the work that has already been performed; this survey is certainly not exhaustive, since inevitably some of the relevant work is left out. Unfortunately, the literature studied is largely restricted to *linear* control systems. Therefore, in future research it has to be investigated if it is possible to adapt or generalize some of the concepts for linear systems for use in nonlinear control system design.

Roughly, control system design consists of performing the following steps [8]:

1. definition of the control objectives/specifications
2. modeling of the system to be controlled
3. *control structure design*
4. controller design
5. control system evaluation and tuning (simulations/experiments)
6. controller implementation

Note that it is not always possible, nor desirable, to perform these steps successively, *e.g.*, the controller design may call for a more accurate system model, or closed-loop simulations may indicate the need for a different control structure.

While studying literature, it has been noticed that different definitions are used for the third step, *i.e.*, the "control structure design" phase. In this report, it is specified as follows:

control structure design is the stage of control system design, in which one decides on the number, the place, and the kind of actuators and sensors to be used and on the internal controller structure interconnecting measured and manipulated variables.

The first phase in CSD, which involves choosing measured and manipulated variables to be used for closed-loop feedback control, will be called the Input/Output selection phase (IO-selection phase). It is emphasized, that in this context the term "output" is referring to measured variables and *not* to variables to be controlled: the latter are strongly related to the control objectives and have to be formulated preceding the CSD. The second phase in CSD will be referred to as the Control Configuration selection phase (CC-selection phase), and is preceding the determination of the appropriate control law. This phase is particularly important for *decentralized control systems* (see, *e.g.*, [5, 70]), and refers to the process of specifying how the selected measurements should be fed back to the selected manipulated variables. This process is sometimes referred to as "partitioning" of the inputs and outputs [54, 56].

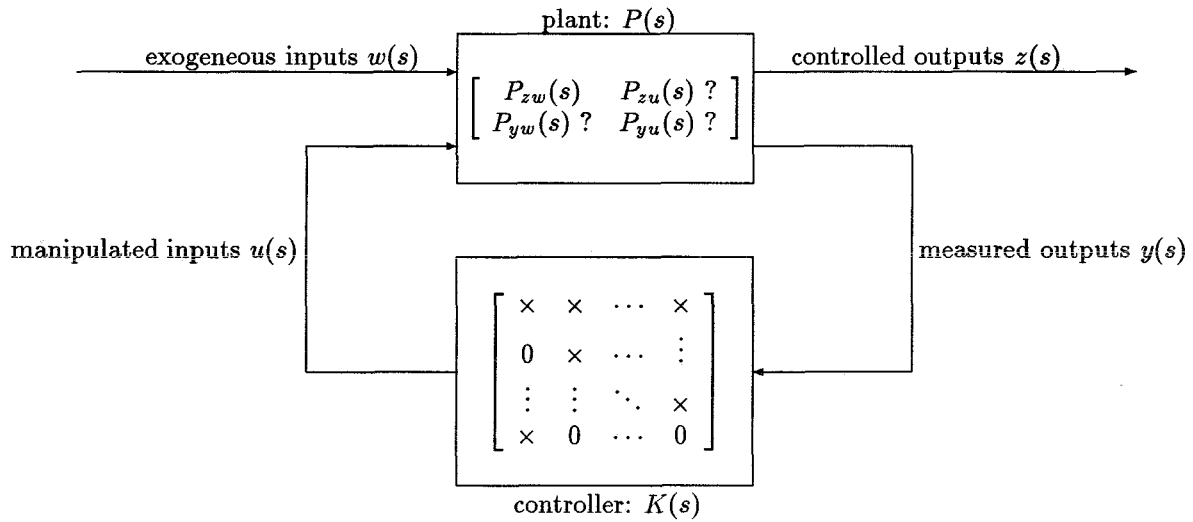


Figure 1.1: General framework for linear control systems

Contrary to centralized control, in a decentralized control system there is only a limited information flow through the controller, *i.e.*, the controller does *not* determine all system inputs from all system outputs. In a decentralized control system one tries to independently control particular subsystems of the full system. The motivation for decentralized control may stem from hardware and design considerations [18, 47]: technically or economically it may not be feasible to apply a fully centralized controller, and moreover the controller design may be simplified, since fewer controller parameters need to be chosen than for the full system.

In [60], it is stated that the problem of control structure selection does not end with the design phase: following control system commissioning, changing process conditions or market demands may alter the dynamics of a plant significantly, by which a redefinition of the control structure is necessary.

The effect of CSD for linear systems under feedback control is illustrated in Fig. 1.1; this control system representation is adopted from [8]. The system is described by the following two relations:

$$\begin{aligned} z(s) &= P_{zw}(s)w(s) + P_{zu}(s)u(s) \\ y(s) &= P_{yw}(s)w(s) + P_{yu}(s)u(s) \end{aligned} \quad (1.1)$$

Note that the transfer function matrices P_{zu} , P_{yw} and P_{yu} are not defined (as indicated by the question marks in Fig. 1.1) until a subset of the candidate measurements and manipulations have been selected for closed-loop control, for the purpose of satisfying the control objectives represented in z . In the CC-selection phase it is decided which of the entries in the controller matrix $K(s)$ should be chosen structurally zero and which not. In [29], it is discussed that after certain modifications in the system description, the controller matrix $K(s)$ can always be represented by a more commonly used block diagonal form.

The motivation to focus on the subject of CSD is, that it is as important as the actual controller design. In [35, 36], it is stated that a wrong choice of actuators and sensors may put fundamental limitations on the system's closed-loop performance, that cannot be overcome by advanced controller design. Moreover, the *complexity* of a control system is largely determined by the underlying control structure [56]. A possible definition of complexity is the sum of the number of inputs and outputs selected and the number of feedback interconnections between them [54, Chapter 5]. In general, the more complex a control system is, the more it costs, the harder it is to maintain and the less reliable it is. For these reasons, CSD is a very important issue in control

system design. Unfortunately, contrary to control law design, CSD has only been paid limited attention to. In [54], it is even stated that by neglecting the CSD phase, modern control theory has set the stage for the design of unnecessarily complex controllers.

In practice, CSD is often carried out in an intuitive ad hoc fashion rather than systematically: engineers use experience, simulation and trial and error to guide actuator and sensor selection and placement. Particularly for large-scale systems, favorable candidate control structures are easily overlooked. In [54, 56], it is shown that the number of alternative control structures grows extremely rapidly as the complexity (as defined above) of the system to be controlled increases. Because of this, ad hoc or other inefficient search techniques for CSD are rendered impractical for large-scale systems by the overwhelming number of candidate control structures. Therefore, in this report emphasis is on systematic and quantitative approaches to CSD, rather than on more qualitative ones based on engineering heuristics.

Certainly the ultimate test of a candidate control structure will be the control system performance once a controller has been designed for the control structure. However, making this the *only* test of a candidate control structure, leads to a CSD-procedure that rapidly becomes infeasible as the system size increases, since for each control structure a controller would have to be designed. Such a computationally intense and time consuming procedure would be intractable for anything but a small group of candidate configurations.

In this report, the quality, *i.e.*, the practical relevance, of the various CSD methods proposed in literature will be addressed. Of course, it is possible to evaluate the CSD methods for a huge set of criteria. Since this is not feasible, only a restricted set of criteria is suggested here, which represents the favorable properties the "ideal" IO-selection or CC-selection procedure must possess.

In [56], it is stated that control system complexity, system uncertainty and accuracy specifications are critical issues in modern control system design. Therefore, a paradigm for control system design is proposed: *Minimize control system complexity subject to the achievement of accuracy specifications in the face of uncertainty*. The problem of CSD can also be formulated in the context of this paradigm. Consequently, some desirable aspects to be accounted for in CSD are:

1. **robust stability:** Robust stability implies that the controlled system will maintain stable operation in the presence of uncertainty.
2. **nominal performance:** It is desirable that a (nominal) performance measure can be specified, which must be achieved with the selected controller structure, *e.g.*, specifications on the closed-loop bandwidth and high-frequency roll-off rate, or offset-free steady-state behavior.
3. **robust performance:** The control system should perform well also in the presence of modeling errors. This aspect implies both robust stability and nominal performance. So, if robust performance is addressed in the CSD method, the first two aspects are redundant. However, these aspects remain very useful criteria for initial screening of a possibly huge set of candidate control structures.
4. **complexity of controller structure:** It should be possible to impose the allowable control system complexity. In [54], it is emphasized that complexity is not a well-defined concept. It is argued that not only the number of inputs, outputs and feedback interconnections is important, but issues such as sensor/actuator costs, reliability, maintainability, and controller design/tuning must be considered as well.

In addition to these aspects to be accounted for during CSD, a CSD method must be:

5. **general:** The CSD method should be applicable to a large class of control systems; generality is often impaired by assumptions on, *e.g.*, the multiplicity of the control loops and by restrictions on the systems considered, *e.g.*, only square systems or stable systems are considered. Moreover, various CSD methods are only applicable at steady state.

6. **applicable to nonlinear control systems:** Desirably, CSD methods developed for linear control systems have a nonlinear counterpart, or it must be possible to generalize them to nonlinear systems.
7. **controller-independent:** The CSD method must be performed with open-loop data only, *i.e.*, the CSD method must be independent of controller data: for initial screening of a large number of candidate control structures, it should be possible to eliminate those candidates for which a controller achieving the desired specifications on the controlled system does not exist, regardless of the controller design method [35] that will be used. If closed-loop data is required, CSD may have to be performed all over again for each controller type or controller tuning.
8. **direct:** Usually, the selection of "viable" control structures is based on testing all possible structures for a set of criteria. In that case, the design is iterative and indirect. Preferably, the CSD method *directly* yields one, or maybe some, favorable control structures given the specified control system requirements.
9. **quantitative:** The CSD method preferably provides quantitative measures in screening candidate control structures, rather than qualitative ones based on, *e.g.*, engineering heuristics.
10. **efficient:** The CSD method must be able to quickly and easily evaluate a possibly very large number of candidate control structures. Efficiency is related to the amount of computational and analytical effort needed in the method. Efficiency may, *e.g.*, call for necessary conditions instead of sufficient conditions during the IO-selection or CC-selection phase, see [54]. This is because sufficient conditions for feasibility of a control structure must address *all* aspects of feasibility. Consequently, they must typically be employed in the context of controller design, which prevents the application of sufficient conditions to a large number of candidate structures, [54, Section 2.4]. Furthermore, if modeling errors are addressed in the CSD method, the use of norm bounds on these errors is more efficient (yet more conservative) than the use of an explicit error model.
11. **effective:** The CSD method must be able to eliminate infeasible candidate control structures and maintain the feasible ones. This implies, that the method must be able to very clearly distinguish between the "promising" control structures and the other ones. An effective method therefore calls for necessary and sufficient conditions for feasibility tests. Generally, necessary conditions are not effective, since there is no guarantee that the control structures maintained are actually feasible.
12. **simple:** Desirably, the theory, implementation and application of the CSD algorithm is not too complex or tedious. The key idea of the CSD method must be clear and selection of candidate control structures must be straightforward.
13. **theoretically well developed:** The theory behind a CSD method is desirably well developed/complete and a successful application should prove the method's practical relevance.

It is emphasized that this list of desirable properties of a CSD method is certainly not complete. Moreover, some of the aspects listed above overlap, *e.g.*, 1 and 3, and 5 and 6, or are closely related, *e.g.*, 9 and 10, and 10 and 12.

As it has already been noted, CSD is particularly important for large-scale systems. Therefore, it is not surprising that the greater part of the literature on CSD stems from process control. In this research area, CSD is related to, *e.g.*, the optimal placement of temperature sensors to measure a temperature profile in a distillation column (see, *e.g.*, [32, 36, 37, 44]), the choice between "material balance control" and "energy balance control" (see, *e.g.*, [72]), and pairing of inputs and outputs to obtain noninteracting decentralized control schemes (see, *e.g.*, [18, 24, 42]). However, CSD is also of great importance in aircraft control [15, 21] and in control of mechatronic systems. With respect to the latter research area, CSD is related to, *e.g.*, the favorable placement of strain

gauges, acceleration sensors [19] and actuators [1] for flexible beams.

This report is set up as follows. First, in Chapter 2 a uniform linear system description is chosen, which will be referred to in the rest of the report. In Chapter 3 and 4 different methods for CSD are presented: Chapter 3 focusses on the IO-selection phase while Chapter 4 does so for the CC-selection phase. An overview of practical applications of CSD from literature is given in Chapter 5, followed by a proposal for an example which could be used to evaluate the various IO- and CC-selection methods discussed in Chapter 3 and 4. In Chapter 6 the different approaches in CSD are compared, based on the aspects discussed above, and the most promising CSD methods are proposed. Finally, Chapter 7 suggests some recommendations for future research in CSD for both linear and nonlinear control systems.

Chapter 2

Preliminaries

The IO-selection and CC-selection methods respectively to be discussed in Chapter 3 and 4 are for the greater part based on linear time-invariant system descriptions. To put the approaches in a general, unambiguous framework, it is decided to describe these systems by the following state equations:

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t) + w(t) \\ y(t) &= Cx(t) + Du(t) + v(t) \\ z(t) &= Ex(t) + Fu(t)\end{aligned}\tag{2.1}$$

with:

$$\begin{aligned}x &\in \mathbb{R}^n && \text{state variables} \\ u &\in \mathbb{R}^m && \text{manipulated variables} \\ y &\in \mathbb{R}^l && \text{measured variables} \\ z &\in \mathbb{R}^r && \text{variables to be controlled} \\ w &\in \mathbb{R}^n && \text{system noise} \\ v &\in \mathbb{R}^l && \text{measurement noise.}\end{aligned}$$

Actually, the control objectives identify the controlled variables z as the primary set of measurements which should be made. However, these theoretically desirable measurements are not always available to monitor the control objectives (*e.g.*, in the case of composition of a distillation product) and they have to be replaced by "secondary measurements" in y , *i.e.*, measurements of other system variables (*e.g.*, pressures or temperatures).

After Laplace transformation, the system (2.1) can be written:

$$\begin{aligned}y(s) &= P(s)u(s) + P_w(s)w(s) + v(s) \\ z(s) &= Q(s)u(s) + Q_w(s)w(s)\end{aligned}\tag{2.2}$$

with:

$$\begin{aligned}P(s) &= P_w(s)B + D; & P_w(s) &= C(sI - A)^{-1} \\ Q(s) &= Q_w(s)B + F; & Q_w(s) &= E(sI - A)^{-1}.\end{aligned}$$

The manipulated variables are often generated by a dynamical feedback of measured variables (see Fig. 1.1) and therefore the relation between these variables can be written as follows:

$$u(s) = K(s)y(s)\tag{2.3}$$

The system descriptions used in the CSD approaches discussed in this report are all based on (2.1)-(2.3), except if it is explicitly remarked.

Chapter 3

Criteria for selection of inputs and outputs

In this chapter, some potential methods for the selection of measured and manipulated variables are outlined.

3.1 Control power and speed

In [58], Chapter 14 is devoted to selection of regulatory "control structures" in process control. However, the term control structure has a broader meaning there: it is not only used to indicate which sensors/actuators are selected and interconnected, but it is also used to indicate the *type* of controller, *e.g.*, feedback and feedforward control, cascade control and decoupling control.

In the approach discussed, it is assumed that the variables to be controlled are measurable or can be replaced by secondary measurements; z is thus assumed to be completely incorporated in y . Furthermore, it is stated that the number of manipulated variables should at least equal the number of variables to be controlled.

The selection of the manipulated variables u is based on control *power* and control *speed* as performance criteria. Unfortunately, the approach is rather qualitative. The first criterion refers to the static influence of the manipulated variable on the variable to be controlled, for instance when the manipulated variable goes from the nominal operating point to fully open (maximum attainable value of the manipulated variable). If this influence is weak, or weak compared to the influences of other controlled variables, the candidate control loop is rejected. The second criterion refers to the speed of reduction of a deviation in the controlled variable. It is desirable that this speed is high and the inputs are therefore chosen to meet this feature. For an objective comparison of regulatory control speeds, it is assumed that in all cases a PID algorithm is used, tuned in the same way.

In [60], two heuristic guidelines for input selection are suggested, which are closely related to the approach discussed above:

- Select inputs that have large effects on the controlled variables, *i.e.*, select inputs associated with a large steady-state gain.
- Select inputs that rapidly affect the controlled variables, *i.e.*, select inputs with small time constants and delays.

3.2 Locations of poles and zeros

It is commonly known (see, *e.g.*, [38, Sections 1.7 and 3.6]), that any zeros or poles incorporated in the plant $P(s)$ which are located in the Right Half Plane (RHP), restrict the range of frequencies

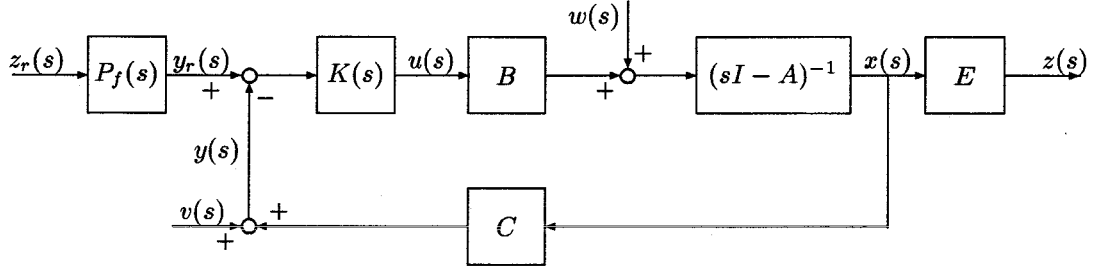


Figure 3.1: General block scheme of linear control systems

over which the use of feedback can be beneficial.

RHP-zeros impose an upper bound on the range of frequencies over which the sensitivity to disturbances acting on the system can be reduced (see Fig. 3.1 in which $z_r(s)$ and $y_r(s)$ represent reference signals and the pre-filter $P_f(s)$ translates the reference signal in terms of controlled variables $z_r(s)$ into terms of measured variables $y_r(s)$). A RHP-(transmission)-zero in the plant limits the achievable bandwidth of the plant, regardless of the type of controller that is used. The reason is, that with a RHP-(transmission)-zero the controller cannot invert the plant and perfect control is impossible. Therefore, plants with RHP-(transmission)-zeros within the desired bandwidth should be avoided [25]. For the SISO case, deterioration in control quality is inversely proportional to the distance of the zero from the origin. In the MIMO case, this is not straightforward, but RHP-zeros with small magnitudes are likely to cause problems in MIMO systems as well [20].

RHP-poles impose a lower bound on the range of frequencies over which measurement noise must be passed without attenuation (the loop bandwidth), *i.e.*, RHP-poles impose a lower bound on the loop bandwidth.

From (2.1)-(2.2) it is seen that the presence of RHP-zeros or RHP-poles is partly determined by the choice of the matrices B , C and D . Therefore, since it is desirable to avoid these zeros and poles, this should be accounted for in the CSD phase, during which the matrices of interest are determined. However, an IO-selection procedure based on the notions of performance limitations due to RHP-zeros and RHP-poles has not been found in literature.

For *nonlinear* systems, RHP-zeros correspond with unstable zero-dynamics, while RHP-poles correspond with unstable manifolds.

3.3 Controllability and observability

A method for IO-selection based on the concepts of controllability and observability is proposed in [46] and summarized in [53]. This idea stems from the notion that each control structure should simultaneously achieve the objectives of bringing the controlled output of the system to the desired one and monitoring all variables (states) which are critical for system performance. It is clarified that the concepts of complete state controllability and observability have some deficiencies for the purpose of IO-selection. Instead, *structural* controllability and observability aspects have to be considered.

The IO-selection procedure was originally developed for regulatory control schemes in process control. The control systems considered are assumed to be represented in a form equivalent to (2.1). Moreover, it is assumed that the controlled variables in z , resulting from the control objectives, are incorporated in y ; if some of the elements in z are not directly measurable, it should be possible to calculate or estimate them from secondary measurements and incorporate them in y . The algorithm to select appropriate measured and manipulated variables is then based on structural controllability and structural observability of the state space description (2.1).

In [46], specifically for PI control schemes the feasibility is investigated. Therefore, the undis-

turbed state space description (2.1) is augmented with variables x^* to include integral action:

$$\begin{bmatrix} \dot{x} \\ \dot{x}^* \end{bmatrix} = \begin{bmatrix} A & 0 \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ x^* \end{bmatrix} + \begin{bmatrix} B \\ D \end{bmatrix} u. \quad (3.1)$$

If the augmented state feedback $u = K[x \ x^*]^T$ is used, proportional and integral (PI) control actions are introduced. The system (2.1) is said to be integral controllable if (3.1) is state controllable, *i.e.*, the pair

$$\left[\begin{pmatrix} A & 0 \\ C & 0 \end{pmatrix}, \begin{pmatrix} B \\ D \end{pmatrix} \right]$$

has to be controllable. If a system does not satisfy this condition, then no feasible PI control system can be found. Before continuing, some definitions have to be made:

Definition 3.1: structural matrix [61]

A structural matrix is a matrix having only two types of entries:

1. fixed zeros which can never take a non-zero value independently of the values of all parameters in the system,
2. arbitrary entries which may take any value (including zero) depending on model parameters.

Definition 3.2: generic rank [16]

The generic rank of a structural matrix M notated by $\rho_g(M)$ is the maximal rank that M achieves as a function of its arbitrary (non-zero) elements.

A structural model, in contrast to a numerical one, is more meaningful for IO-selection based on controllability/observability aspects. This is because a structural model depends on invariant aspects of the system only. A numerical model depends on the values of the model parameters, which are never known precisely, with the exceptions of zeros that are fixed by absence of physical connections between different process units (subsystems). Therefore, an unfortunate choice of some parameters may yield an uncontrollable/unobservable system. So, a numerical model does not provide useful *global* information about the controlled system's behaviour. However, a disadvantage of a structural representation is the impossibility of drawing quantitative conclusions on controllability and observability in the sense of strength and direction of the couplings between input and output variables [31]. In the selection algorithm, the structural controllability of the system (3.1) plays an important role.

Definition 3.3: structural controllability [16]

The structural pair (A, B) is structurally controllable if:

1. every state is *accessible* from at least one input, and
2. the generic rank of $[A \ B]$ is n .

Analogously, (A, C) is structurally observable if (A^T, C^T) is structurally controllable.

Definition 3.4: "extended" structural controllability

The structural pair

$$\left[\begin{pmatrix} A & 0 \\ C & 0 \end{pmatrix}, \begin{pmatrix} B \\ D \end{pmatrix} \right]$$

is structurally controllable if:

1. (A, B) is structurally controllable, and
2. the generic rank of the "structural compound matrix" S_c defined by

$$S_c = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

is $n + l$.

Inaccessibility of an output from a manipulated variable implies that the manipulated variable has no influence on the output [30]. If the accessibility conditions are satisfied for both structural controllability and observability, all unstable modes can be influenced and observed, with the exception of poles in the origin. The generic rank condition serves to detect pure integrators which are not controllable with a given set of manipulated variables. For stable systems, accessibility to the states which are pure integrators only is important [46]. Note that a necessary condition for $\rho_g(S_c) = n + l$ is that $\rho_g(A, B) = n$ and $\rho_g(A^T, C^T) = n$, *i.e.*, the rank conditions for structural controllability and structural observability have to be satisfied, but need not be checked separately. In [16, 30, 61], a lot more can be read about structural aspects of (control) systems and controllability/observability aspects.

The procedure for selecting measured and controlled variables to generate feasible control structures as discussed in [46], is then as follows:

1. Choose y to represent the variables to be controlled by primary or secondary measurements. The selected measurements are then represented by $y = Cx + Du$.
2. Test for "dual accessibility" of the structural pair (C, A) , *i.e.*, test if (A^T, C^T) is accessible with the measurements selected in step 1. If the test is negative, augment y to $y^* = C^*x + D^*u$.

3. Form the structural matrix

$$S_c^* = \begin{bmatrix} A & B \\ C^* & D^* \end{bmatrix}$$

where all the feasible manipulated variables make up the columns of B and D .

4. Delete columns (one at a time) from

$$\begin{bmatrix} B \\ D^* \end{bmatrix},$$

such that the number of remaining manipulated variables \bar{m} , corresponding to the remaining columns in

$$\begin{bmatrix} \bar{B} \\ D^* \end{bmatrix}$$

is equal to the number of observations l^* for the system.

5. Test for accessibility of (A, \bar{B}) . If it is not satisfied, the set of manipulated variables selected is not feasible, it is rejected and a different \bar{B} must be chosen to achieve accessibility. If the test is satisfied, perform step 6.
6. Test if

$$\bar{S}_c^* = \begin{bmatrix} A & \bar{B} \\ C^* & D^* \end{bmatrix}$$

is structurally nonsingular, *i.e.*, test if $\rho_g(\bar{S}_c^*) = n + l^*$. If the test is affirmative, the selected measurements corresponding to the rows of C^* and manipulated variables corresponding to the columns of \bar{B} represent a feasible IO-set. Otherwise, reject the set of selected variables as infeasible.

Applying this algorithm, all structurally controllable IO-sets can be generated, which may still be a very large number. Then, further screening at different levels of sophistication (*e.g.*, by engineering heuristics or dynamic simulations) has to be performed to reduce all possible alternatives. It is emphasized that an IO-set which is structurally controllable need not be numerically controllable [16], *i.e.*, the IO-selection method discussed above may yield infeasible IO-sets, depending on the numerical values of the parameters of the physical system.

In [16], it is stated that the concept of structural controllability as a feasibility criterion can also be used for *nonlinear* systems; by linearizing a nonlinear system description, a linear structural system description can be obtained. Since nonlinear equivalents for complete state controllability and observability have also been defined [51], it is expected that these concepts can be applied in IO-selection for nonlinear control systems as well.

3.4 Cause-and-effect graphs

A systematic procedure for generating alternative feasible IO-sets based on the cause-and-effect representation of the steady-state process is discussed, *e.g.*, in [17] and summarized in [53]. The method is stated to be one of the first non-numerical techniques to solve the problem of synthesizing control structures for process control.

Cause-and-effect relationships between different variables of a system can be represented by a directed graph or digraph (see Fig. 3.2). Nodes in the graph are the system variables (states, inputs, outputs and disturbances) and the edges, *i.e.*, the directed lines, show the relationships between these variables. The edges can carry information about cause-and-effect relationships between the variables, such as steady-state gains, time constants and dead times. So, by tracing paths in the graph, it is possible to find which variables affect a specific process variable or which variables are affected by the given variable. In order for control to be effective, there must be a causal path between the manipulated variables u and the variables to be controlled z .

After the digraph for the complete process is generated, the next step is to determine the "constrained variables": the control objectives define the process variables to be maintained within a certain error around the steady-state value and these variables are called the constrained variables (z). Furthermore, variables which violate production, safety or operational limits may be identified as constrained variables. The next step is to propagate the constraints through the cause-and-effect graph with the goal to locate alternative sets of measured and manipulated variables in the graph to satisfy the objective constraints. As the constraints are propagated, the process variables encountered along the edges are classified as ANDed or ORed by the following rules:

- candidate *input variables* are ANDed if all these variables are required to be controlled in order to control the constrained variable
- candidate *input variables* are ORed if control of either of these variables is sufficient to control the constrained variable
- candidate *measured variables* are ANDed if all these variables have to be measured in order to obtain the value of the constrained variable
- candidate *measured variables* are ORed if a single measured variable is sufficient to obtain the value of the constrained variable.

Thus, by considering the digraph, possible ways of measuring and manipulating the constrained variables are searched for. Any set of IO-variables which makes control and measurement of the constrained variables possible is a candidate IO-set. So, initially a possibly large number of candidate IO-sets is generated. In order to reduce the alternatives to a smaller subset, they are evaluated based on selection heuristics. The resulting subset is in turn further screened, *e.g.*, by performing dynamic simulations.

Since it is also possible to represent *nonlinear* systems by directed graphs, see, *e.g.*, [12], the IO-selection method is expected to be applicable to nonlinear control systems as well.

In fact, the method based on the cause-and-effect graphs uses accessibility as a criterion that a feasible IO-set should satisfy. According to the structural controllability criterion (Definition 3.3) discussed in Section 3.3, accessibility only is not sufficient for feasibility, since the generic rank test may fail for certain IO-sets. In [31], it is stated that the qualitative nature of the selection method based on cause-and-effect graphs is a disadvantage

3.5 Achievable performance

In [8], a method is discussed to determine what performance specifications (of a large but restricted class) can be met using *any* linear controller design method, for a given linear system and IO-set. Given a fixed set of performance (or robustness) specifications on the controlled system, an IO-set is feasible if at least one controller exists that satisfies the specifications. Based on the outcome

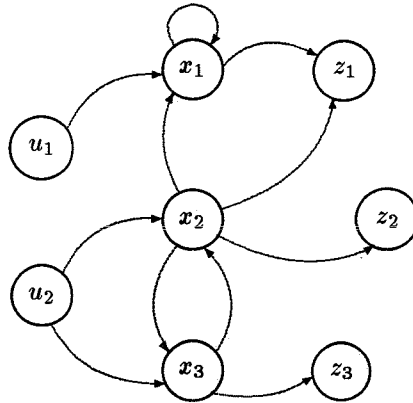


Figure 3.2: Example of a simple digraph for a control system

of the feasibility problem, the designer may, or may not, modify the choice and placement of the sensors and actuators. If the specifications are feasible, the designer might remove actuators and sensors to see if the specifications are still feasible; if the specifications are infeasible, actuators and sensors may be added or relocated until the specifications become achievable. Selecting measured and controlled variables in this way is therefore iterative. A systematic methodology to decide which variables are the best to be removed/added, is however not proposed.

Another possible selection criterion based on performance considerations uses the minimally achievable value of the quadratic (performance) criterion by the combination of an optimal control law and an optimal observer (Kalman filter), see [33]. Consider the time-invariant system (2.1) with $D = 0$ and $F = 0$. In the stationary case, the stochastic linear optimal output (y) feedback regulator minimizes the criterion:

$$\bar{J} = E[z^T(t)W_1z(t) + u^T(t)W_2u(t)] \quad (3.2)$$

with W_1 and W_2 weighting matrices. The minimally achievable value of this criterion can be achieved with a control law $u = -\bar{F}\hat{x}$ (where \hat{x} is the state estimate by the Kalman filter) and can be written as:

$$\bar{J} = \text{tr}[\bar{P}V_w + \bar{Q}\bar{F}^T W_2 \bar{F}] \quad (3.3)$$

in which V_w represents the intensity of the process noise w , \bar{P} and \bar{Q} are solutions of stationary Riccati equations and $\bar{F} = W_2^{-1}B^T\bar{P}$. Since $\bar{P} = \bar{P}(A, B, E, W_1, W_2)$ and $\bar{Q} = \bar{Q}(A, C, V_w, V_v)$ (V_v represents the intensity of the measurement noise), the minimum of the performance criterion \bar{J} depends, among others, on the matrices B and C , which are determined in the IO-selection phase during CSD.

From (3.2), it is obvious that the minimally achievable \bar{J} is also dependent on the scaling of the variables to be controlled z and the candidate manipulated variables u . For example, if a manipulated variable in u with unit [N] is replaced by one with unit [mN], a much larger \bar{J} will result. Therefore, it is important that z and u are scaled in such a way that their values are representative of their *relative* importance, so that the scaled variables can be compared numerically to each other. Scaling must be accounted for by proper choice of the weighting matrices W_1 and W_2 .

From the point of view of good performance under Linear Quadratic Gaussian (LQG) control, a small value of \bar{J} is desired. Provided that z and u are properly scaled, it is recommended to choose B and C , *i.e.*, to select the IO-set, such that the smallest \bar{J} is achieved. This method can also be used for time-dependent state space descriptions in the instationary case. However, computation

of the performance index J then implies the solution of two differential Riccati equations, requiring high computational effort.

In [31], a number of candidate input sets is evaluated by computing \bar{J} , which is the minimal value of the criterion $\int_{\tau=0}^{\infty} (z^T(\tau)z(\tau) + u^T(\tau)u(\tau))d\tau$; the input set which achieves the smallest \bar{J} is the most promising for control. The same procedure is possible for determining a proper set of measured variables.

Optimal control of *nonlinear systems* is, *e.g.*, discussed in [50]. The optimal control problem in this case consists of finding a control u that minimizes, *e.g.*, the criterion $\int_{\tau=0}^t f(x, u, \tau)d\tau$. Maybe, this offers a potential tool for IO-selection in nonlinear control systems.

3.6 Accuracy of state estimates

In [32], the optimal location of temperature and concentration measurements along the length of a tubular reactor is considered. Since the objective of measurements is to gain information on the system, it is stated that a sensible criterion for optimal measurement selection is, that the best possible estimates of the system state can be made on the basis of the selected measurements. The optimal sensor location problem is then posed as that of selecting a given number of points out of the prespecified locations, the so-called "collocation points". Furthermore, it is stated that the obtainable quality of state estimates is often more dependent on the *location* of the sensors than on the *number* of sensors.

The selection problem is approached from a stochastic point of view, acknowledging the uncertainties inherent in system parameters and measurements. A conventional way of accounting for such model uncertainties, is to introduce random disturbance terms into the system equations (w and v in (2.1)) of which the statistical properties reflect in some manner the expected degree of modeling error. For the purpose of optimal measurement location, the nonlinear system equations are linearized around steady-state:

$$\begin{aligned} \dot{x} &= Ax + w & \text{with } w : \text{zero-mean white Gaussian process noise} \\ y &= Cx + v & \text{with } v : \text{zero-mean white Gaussian measurement noise.} \end{aligned} \quad (3.4)$$

It is assumed that the measurement noise at different locations is not correlated. The goal is to obtain the optimal estimate of the state x , given the measurements $y(\tau)$, $0 \leq \tau \leq t$. For a given system and measurement set, the estimate \hat{x} for which the estimation error covariance matrix $P(t) = E[\{x(t) - \hat{x}(t)\}\{x(t) - \hat{x}(t)\}^T]$ is "minimal", is called the optimal estimate. The corresponding optimal covariance matrix $P(t)$ is calculated from a Riccati equation, which depends on the system matrix A , the measurement matrix C and the process and measurement noise intensity matrices.

To account for the number and location of measurements performed, a vector a is introduced with $a_i = 1$ if a measurement is used and $a_i = 0$ if it is not. So, if the total number of sensors is specified to be m , only m of the candidate measurements in y are performed and only m of the elements in a take the value 1; these m measurements have to be chosen from the N collocation points. The covariance matrix $P(t)$ which indicates the accuracy of the state estimate becomes now dependent on a . An optimal measurement policy is then defined as the one that minimizes an appropriate scalar measure of $P(t)$. Since $P(t)$ is a nonnegative definite matrix, its trace is a measure of its magnitude and therefore the "optimality index for measurement location" is chosen as:

$$J = \alpha \text{tr}[P(t_f)] + \beta \int_{t_0}^{t_f} \text{tr}[P(t)]dt \quad (3.5)$$

with α and β positive constants specifying the relative weights of the two terms. The optimal sensor selection problem is then solved by minimizing J with respect to the parameters a_i . In [32], an iterative algorithm is proposed to solve this problem; the set of a_i 's determined by the procedure designates the optimal sensor locations.

3.7 Economics

In [48], a systematic method is outlined, that can be used to select the economically optimal measured and manipulated variables in process control, without designing the controller, while maintaining good controllability characteristics. It is stated that different IO-selections lead to different controller performance, as well as to different capital and maintenance costs; it is with the trade-off between instrumentation costs and operating benefits that the paper is concerned.

The scope of the problem is restricted to linear(ized) system descriptions for processes whose operation is dominated by steady-state aspects. Furthermore, the *measured* variables are assumed to be perfectly regulated ($y = 0$) and only square systems are considered, *i.e.*, $l = m$. Operating constraints and disturbances are accounted for; the effects of disturbances on the plant are reflected in the variation of plant variables which are not selected as measured variables.

The method for selection of IO-sets is based on varying a permutation matrix Ω ; the permutation matrix can be used to select each of the candidate combinations of measured and manipulated variables (compare with the method discussed in Section 3.6, where the vector a contains information on the measurements to be performed). The integers in Ω define the IO-set. Varying Ω , it is possible to assess the influence on the economics for all the candidate IO-sets and choose the optimum directly. This method is only feasible for problems with a small number of candidate manipulated and measured variables. For this reason, a Mixed Integer Linear Programming (MILP) method is used to evaluate the candidate IO-sets. However, as the problem is combinatorial in nature, this method requires too high computational effort if the number of IO-variables is large. Therefore, initial screening is desirable, in order to eliminate infeasible IO-sets without first evaluating all of them. This may be achieved by the introduction of "structural connectivity constraints", *i.e.*, the integer search space is reduced by using structural information about the relationships between the inputs and outputs (input/output connectivity).

Ideally, the transfer function matrix $P(0)$ between the active manipulated variables and the perfectly controlled measurements is structurally nonsingular. However, a weaker set of structural connectivity constraints is used: for every selected measured variable to be perfectly controlled, at least one manipulated variable which affects the measured variable must be active. Similarly, for each active manipulated variable selected, at least one measured variable that it affects must be perfectly controlled. By introducing these extra constraints, not all possible IO-sets have to be considered and computational effort to solve the IO-selection problem is reduced.

In [48], it is emphasized that the MILP should only be used as a *screening tool* for prediction of economically sound IO-sets, for different reasons:

- The analysis does not examine the controllability of the process.
- The MILP analysis only calculates an estimate of the "dynamic economics", *i.e.*, the economics under influence of the disturbances.
- The linearization only has a limited accuracy.

Because of these limitations, the following IO-selection procedure is suggested. A number of IO-sets which are to be examined in detail are proposed. For these candidate IO-sets, the MILP algorithm is solved, by which conclusions can be drawn on the IO-sets yielding the best dynamic economics. These IO-sets are then all subjected to controllability analyses to test the validity of the perfect control assumption. Minimum condition number plots (desirably small) and RHP transmission zeros (desirably none) are used as controllability indicators. Moreover, the selected IO-sets can be used for a *nonlinear* dynamic economic analysis, which is unfortunately a complex problem. The results of these analyses should then be used in conjunction to select the best IO-set.

3.8 Morari resiliency index

In [73], a method is discussed for CSD for multiloop SISO controllers in a multivariable process environment. No attention is paid to the measurement selection problem. Instead, it is only stated

that the controlled variables should be directly measured or should be computed from other directly measured variables (secondary measurements). The selection of the controlled variables z is primarily based on engineering judgement and good understanding of the process. Considerations of economics, safety, constraints, availability, and reliability of sensors must be factored into this decision.

The input selection problem is treated in a more quantitative way; it is based on the "resiliency" of the plant. In [45], the term resiliency is used to describe the ability of the plant to move fast and smoothly from one operating condition to another (including start-up and shut-down) and to deal efficiently with disturbances and model-plant mismatches. Based on the work in [45], the authors of [73] propose the so-called Morari Resiliency Index (MRI) to guide the selection of manipulated variables:

$$\text{MRI} = \underline{\sigma}[P(j\omega)]. \quad (3.6)$$

The MRI is the minimal singular value $\underline{\sigma}$ of the plant transfer function matrix $P(j\omega)$ (or $Q(j\omega)$, since y and z are assumed to be equivalent). The set of manipulated variables that gives the *largest* minimum singular value over the frequency range of interest is the best, *i.e.*, the corresponding IO-set yields the most resilient system. Unfortunately, the MRI is expected to be not effective, since it fails to satisfactorily address all aspects of resiliency as mentioned above.

The selection of IO-sets based on the MRI is independent of the control configuration and controller design. However, the procedure is scaling dependent. This problem can be circumvented by expressing the gains of all the plant transfer functions in dimensionless form, or by otherwise properly scaling of the system description.

3.9 Condition number

In [21, 54, 56], an IO-selection procedure is presented that is based on the condition number of the plant; the proposed algorithm has been implemented in the MATLAB Control Configuration Design Toolbox [54, 57]. The theory proposed provides quantitative, efficient, and necessary conditions for *viability*: a control structure is termed viable (feasible) if it allows accuracy specifications to be achieved in the face of uncertainty. The conditions are quantitative because they incorporate quantitative expressions of the control performance requirements (accuracy and uncertainty) and the conditions are efficient because they can be applied to the open-loop system prior to control law design.

Robust stability is one fundamental issue for viability; other issues are, *e.g.*, nominal stability, nominal performance, robust performance and closed-loop integrity. The criterion for selecting measured and manipulated variables is based on a necessary and sufficient condition for robust stability, arising from the small gain theorem in robust control theory, see, *e.g.*, [38]. An additive unstructured uncertainty description is used to represent the uncertainties in the transfer function matrix of the plant. It is very important to note that in the criterion discussed below the "plant" P_s corresponds to various selected subsets of the candidate measured and manipulated variables, *i.e.*, P_s is a *subsystem* of P in (2.2):

Suppose P_s is a Finite-Dimensional Linear Time-Invariant (FDLTI) nominal plant. Suppose also that K is a FDLTI controller which stabilizes P_s . Under these conditions, K stabilizes all $\tilde{P}_s = P_s + \Delta_a$ with the same number of RHP-poles as P_s and $\bar{\sigma}(\Delta_a) \leq \delta_a$ if and only if:

$$\bar{\sigma}[K(I + P_s K)^{-1}] < 1/\delta_a \quad \forall \omega. \quad (3.7)$$

This necessary and sufficient condition is then considerably weakened [54] to a *necessary* condition for robust stability, which is independent of the controller $K(s)$ and is more appropriate for screening of candidate IO-sets:

Suppose P_s is a square, FDLTI nominal plant. Under these conditions, there exists a FDLTI controller K which

1. stabilizes all $\tilde{P}_s = P_s + \Delta_a$ with
 - (a) the same number of RHP-poles as P_s and
 - (b) $\bar{\sigma}(\Delta_a)/\bar{\sigma}(P_s) \leq \delta_{ra}$, and
2. achieves $\bar{\sigma}(S) \leq \sigma_S$, $\sigma_S < 1 \forall \omega \leq \omega_S$

only if

$$\kappa(P_s) < \frac{1}{\delta_{ra}} \left(\frac{1}{1 - \sigma_S} \right) \forall \omega \leq \omega_S \quad (3.8)$$

where:

- $S = (I + P_s K)^{-1}$ is the nominal output sensitivity function of the closed-loop system,
- $\kappa = \bar{\sigma}(P_s)/\underline{\sigma}(P_s)$ is the Euclidean condition number of the plant,
- δ_{ra} is the specified, possibly frequency-dependent, relative-additive uncertainty margin, and
- σ_S and ω_S specify the closed-loop bandwidth of the system in terms of S .

Qualitatively, the selection criterion implies that a selected subset with a "large" condition number can only tolerate "small" amounts of unstructured uncertainty without sacrificing robust stability. Moreover, the criterion is only meaningful for systems where tracking and disturbance rejection are important, so that a bandwidth ω_S is specified. A selected subsystem/subset which fails to satisfy (3.8) is not considered a viable IO-set. The condition number criterion has a *quantitative* nature in the sense that uncertainty and performance specifications enter explicitly through δ_{ra} , σ_S and ω_S in provisions 1 and 2 of the criterion above. Furthermore, the condition is *efficient* in that it is easy to compute and does not require prior design of a control law. However, the criterion is *not necessarily effective*, since it checks a necessary condition only, *i.e.*, infeasible candidate IO-sets may pass (3.8).

In [54, 56], it is stated that satisfying the control objectives requires that the selected measurements y be "strongly related" to the performance variables z : since the performance variables may not always be measurable, one attempts to control z by controlling y , using knowledge of the performance variables z as a function of the selected measurements y . Thus performance specifications expressed in terms of z must always be translated into performance specifications on y . Selection of an appropriate IO-set is therefore crucial to satisfactorily control the performance variables. So, the IO-sets which pass the proposed selection criterion are practically useful only, if it is possible to relate z with the measured variables selected. Since this is not explicitly stated in the selection criterion, it must have been assumed that this is always possible, no matter which candidate measurements are selected.

Unfortunately, the condition number of the plant is scaling dependent, *i.e.*, it depends on the choice of the units for u and y , while the uncertainty margin δ_{ra} and the closed-loop bandwidth ω_S are specified under the assumption that the plant is properly scaled. However, scaling each subset individually would make the IO-selection procedure burdensome and less efficient. This problem can be avoided by replacing the condition number $\kappa(P_s)$ by the minimal condition number $\kappa^*(P_s)$, where $\kappa^*(P_s) \leq \kappa(P_s)$. Since computation of $\kappa^*(P_s)$ is still an open problem, lower bounds on $\kappa^*(P_s)$ can be established, leading to the following result:

The IO-selection criterion remains valid when either of the following is substituted for $\kappa(P_s)$ in (3.8):

$$2\max\{\|\Lambda(P_s)\|_1, \|\Lambda(P_s)\|_\infty\} - 1 \text{ or } \bar{\sigma}(\Lambda(P_s)) \quad (3.9)$$

where $\Lambda(P_s) = P_s * P_s^{-T}$ is the Relative Gain Array (RGA) of the plant P_s and "*" denotes elementwise multiplication known as the Schur (or Hadamard) product.

The RGA will be discussed in more detail in Section 4.4. This modification of (3.8) weakens the necessary condition for robust stability in the sense that the selection procedure becomes less severe and a larger number of candidate IO-sets will pass the criterion. Efficiency is very important in the initial screening of a large number of candidate IO-sets. Once a smaller "pool" of IO-sets is left, the stronger, but less efficient, scaling-dependent criterion (3.8) can be used.

Finally, it is remarked that the selection criterion could also be used for testing closed-loop integrity, *i.e.*, testing if the closed-loop system remains robustly stable if one or more actuator/sensor failures occur.

3.10 Singular value decomposition

The Singular Value Decomposition (SVD, see, *e.g.*, [47, Chapter 10]) is frequently encountered in literature as a tool for IO-selection. In this section, some approaches will be outlined.

In [65] and [47, Chapter 13], a method is proposed for selecting input variables based on the effectiveness of disturbance suppression, which is strongly dependent on the *disturbance direction*, *i.e.*, the direction of the system output vector z resulting from a specific disturbance. Disturbance rejection is often the main objective of process control. For multivariable systems, usually each disturbance affects all the outputs; a well designed control system should be able to reject these disturbances at steady-state. The linear control system considered is described by the following equation:

$$z(s) = Q(s)u(s) + Q_d(s)w(s) = Q(s)u(s) + d(s) \quad (3.10)$$

where $d(s)$ represents the *effect* on the controlled outputs of the physical disturbances $w(s)$. The transfer function matrix $Q(s)$ is assumed to be square ($r \times r$).

The input selection procedure uses the SVD of the complex matrix Q :

$$Q = W\Sigma V^H \quad (3.11)$$

where W and V are unitary matrices and Σ is a diagonal matrix containing the real nonnegative singular values in descending order:

$$\Sigma = \text{diag}[\sigma_i]; \quad \bar{\sigma} = \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r = \underline{\sigma} \geq 0. \quad (3.12)$$

Matrix W consists of the so-called left singular vectors and matrix V of the right singular vectors. For the singular vectors associated with the largest and smallest singular value, it can be written:

$$\begin{aligned} Qv_{max} &= \bar{\sigma}(Q)w_{max} \\ Qv_{min} &= \underline{\sigma}(Q)w_{min}. \end{aligned} \quad (3.13)$$

Vector v_{max} therefore corresponds to the direction of the input which undergoes the largest amplification and v_{min} to the direction with the smallest amplification.

Consider the system (3.10). As it is stated above, it is of interest to investigate the magnitude of the manipulated variables necessary to compensate for the effect of a disturbance. In this context, it is reasonable to use the Euclidean norm (2-norm) as a measure of magnitude of u , because it "sums up" the deviations of all manipulated variables. Consider a particular disturbance d . For complete rejection of this disturbance ($y = 0$) at steady-state, u should satisfy $u = Q^{-1}d$. The quantity

$$\|u\|_2/\|d\|_2 = \|Q^{-1}d\|_2/\|d\|_2 \quad (3.14)$$

depends only on the *direction* of the disturbance, but not on its magnitude. It measures the magnitude of u needed to reject a disturbance d of unit magnitude which enters in a particular direction expressed by $d/\|d\|_2$. The "best" disturbance direction in the sense that it requires the least action by the manipulated variables, is that of the left singular vector $w_{max}(Q)$ associated with the largest singular value of Q . In this case (see [65]):

$$\|Q^{-1}d\|_2/\|d\|_2 = 1/\bar{\sigma}(Q). \quad (3.15)$$

By normalizing (3.14) with this "best" disturbance, the following measure, the so-called disturbance condition number κ_d of the plant Q , is obtained:

$$\kappa_d = \frac{\|Q^{-1}d\|_2}{\|d\|_2} \bar{\sigma}(Q). \quad (3.16)$$

It expresses the magnitude of the *manipulated variables* needed to reject a disturbance in the direction d relative to rejecting a disturbance with the same magnitude, but in the "best" direction. For nonsquare plants, Q^{-1} should be replaced by the pseudo-inverse [47]. The "worst" disturbance direction is $d = w_{min}(Q)$. In this case $\kappa_d(Q)_{max} = \bar{\sigma}(Q)/\underline{\sigma}(Q) = \kappa(Q)$ with $\kappa(Q)$ the condition number of the plant. It follows that

$$1 \leq \kappa_d(Q) \leq \kappa(Q) \quad (3.17)$$

and $\kappa_d(Q)$ may be viewed as a generalization of the condition number of the plant, which also takes into account the direction of the disturbances. Just like $\kappa(Q)$, $\kappa_d(Q)$ is scaling dependent.

The disturbance condition number may be used to select manipulated inputs: input sets generating low values of $\kappa_d(Q)$ are preferred. An advantage of the disturbance condition number is, that it is independent of the controller. However, a possible disadvantage is its scaling dependence.

In [31], the problem of input selection for linear multivariable systems is addressed. It is stated that the SVD method facilitates the *quantification* of the system's excitation by the various inputs, and thus the selection of an appropriate set of inputs. It is put forward that the importance of reducing the number of inputs not only stems from economic reasons (less actuators amount to lower hardware costs), but also from robustness reasons: a control system design with modeling errors and too many inputs may lead to overparametrized controllers, which may exhibit reduced robustness characteristics. The general idea of the more inputs there are, the better control of the process is possible, is only true if control is based on a perfect model. The goal of the input selection method proposed in [31], is to reduce the number of inputs, while still preserving the controllability of the system. Ideally, the retained inputs are "strong" (see Section 3.1) and such that each of them excites the modes in a way orthogonal to the others.

Consider the following, not necessarily square, system:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ z &= Ex. \end{aligned} \quad (3.18)$$

This system is assumed to be both controllable and observable. Moreover, a primordial assumption is, that the system has been scaled in such a way, that the numerical values of x , u and y are representative of their relative importance and can be compared numerically to each other. The system in (3.18) can also be described by:

$$z(s) = Q(s)u(s) \quad \text{with: } Q(s) = E(sI - A)^{-1}B \quad (3.19)$$

or in modal form (see, e.g., [69]):

$$Q(s) = \hat{E}(sI - \Lambda)^{-1}\hat{B} \quad (3.20)$$

with $\Lambda = H^{-1}AH$, $\hat{E} = EH$ and $\hat{B} = H^{-1}B$. Matrix Λ is diagonal with the eigenvalues of A along its diagonal and H contains the n linearly independent eigenvectors of A . Equation (3.20) clearly shows that the transmission of information from the inputs to the outputs can be considered as a series of three mappings:

1. mode excitation: \hat{B} maps the input space \mathbb{R}^m onto \mathbb{C}^n
2. mode dynamics: $(sI - \Lambda)^{-1}$ maps \mathbb{C}^n onto \mathbb{C}^n
3. mode composition: \hat{E} maps \mathbb{C}^n onto the output space \mathbb{R}^r .

The matrix \hat{B} indicates both the magnitude and the direction in modal space of the excitation from the various inputs. However, the analysis of the inputs by considering the matrix \hat{B} alone requires the dynamics and the composition to be similar for all the modes, *i.e.*, the modal basis must be scaled. In order to find a set of input variables which contains the strongest inputs in terms of a norm of $Q(s)$, a scaling procedure is proposed in [31]. After scaling, $Q(s)$ can be written as $Q(s) = \bar{E}(sI - \Lambda)^{-1}(-\Lambda)\bar{B}$ in which the matrix \bar{B} describes the excitations of the normalized modes. The modes are called normalized, because they all exhibit the same maximal contribution to the dynamics of the system, and the modal space is properly scaled. Consequently, \bar{B} contains information about the *magnitude* and the *direction* of the excitation of the normalized modes by the inputs.

The SVD of \bar{B} is used for input selection:

$$\bar{B} = W\Sigma V^H = W \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^H \\ V_2^H \end{bmatrix} \quad (3.21)$$

where Σ_1 contains the dominant singular values of \bar{B} . The column vectors of V provide an orthonormal co-ordinate system for viewing the inputs; V_1 spans the subspace of the input space which is mapped onto the normalized modes via the dominant singular values. In [31], a method is developed to select the dominant inputs by using V_1 .

Papers [6, 44] discuss the use of SVD in the selection of measured variables for distillation column control.

In [6], a method is presented to select appropriate (temperature) measurement locations, which is based on a compromise between the measurements' sensitivity to manipulated variables and "inferential error", *i.e.*, the error in the unmeasured variables to be controlled, under perfect control of the measured variables. The unmeasurable variables to be regulated z (top and bottom composition errors) are considered to be functions of the measurements y (tray temperatures). So, to keep z at the desired level, secondary measurements y are used by the controller; this is called "inferential control", see, *e.g.*, [68]. The SVD method is used to study the sensitivity of y to the inputs, which should preferably be high, while the sensitivity of z to load disturbances should be low.

The following dynamic system model is considered:

$$\begin{aligned} \dot{x} &= Ax + Bu + Hw \\ y &= Cx \\ z &= Ex \end{aligned} \quad (3.22)$$

with w representing load disturbances. At *steady-state* and under perfect control of the measured variables ($y = 0$):

$$y = Cx = CG_s u + CG_{ls} w = 0 \quad (3.23)$$

with:

$$\begin{aligned} G_s &= (-A)^{-1}B \quad \text{the "state-input gain matrix"} \\ G_{ls} &= (-A)^{-1}H \quad \text{the "state-load gain matrix"}. \end{aligned}$$

The required control to maintain $y = 0$ is:

$$u = -(CG_s)^{-1}CG_{ls}w. \quad (3.24)$$

The error in z when y is perfectly controlled, *i.e.*, the inferential error, is then:

$$z = G_{le}w \quad (3.25)$$

with the "inferential error" matrix:

$$G_{le} = E[-G_s(CG_s)^{-1}CG_{ls} + G_{ls}].$$

For any load disturbance w , only a small deviation in z is permitted, *i.e.*, $\|z\|/\|w\|$ should be kept small, which is equivalent to keeping $\bar{\sigma}(G_{I_e})$ small. The SVD of G_{I_e} can be written as in (3.11). The first column of the left singular matrix, w_1 , indicates the strongest direction of composition error z , while the first column of the right singular matrix, v_1 , indicates the strongest disturbance direction; a load of norm 1 in the v_1 direction yields an inferential error of norm $\bar{\sigma}$ in the w_1 direction (see also Equation (3.13)).

The SVD of G_s and G_{I_e} can be used to trade-off the sensitivity of y to manipulated variables (SVD of G_s) with the magnitude of inferential error (SVD of G_{I_e}). A possible measurement selection rule is then: assuming u , w and y are properly scaled, the elements in the first column of the left singular matrix of G_s with the largest magnitude are chosen to represent the measurements, at the same time avoiding a large value of $\bar{\sigma}$ in the corresponding inferential error matrix G_{I_e} .

In [44], a measurement (location) selection procedure is described, which is a practical compromise between sensitivity of the measurements with respect to the manipulated variables and independence in the sense of a low degree of interaction between the measurements. The paper presents four systematic sensor selection procedures, all based on a SVD analysis of a, not necessarily square, process gain matrix $P(s)$ at steady-state. Again, the elements of $P(0)$ should be scaled properly, see, *e.g.*, [54, Chapter 5].

The main method discussed, is the "principal component analysis". This method bases the sensor selection on the location of the principal components in the left singular matrix W . Each column in W is an orthonormal vector whose co-ordinate directions are described by each one of the candidate process sensors. The left singular vector w_1 points in the direction of the most sensitive combination of sensors; w_2 points in the second most sensitive direction, which is perpendicular to the major direction of sensitivity, etc. Therefore, it stands to reason that the location of the principal component of each vector indicates good choices for sensors which are relatively sensitive and also relatively (but not completely!) independent. The measurement selection is thus determined by performing a SVD analysis on $P(0)$ and then choose the measurements (locations) which correspond to the largest absolute value in each of the left singular vectors. Note that this approach is different from the one presented in [6], where the largest absolute values in w_1 only are used to select a measurement set; by doing this, independence of the measurements is not accounted for!

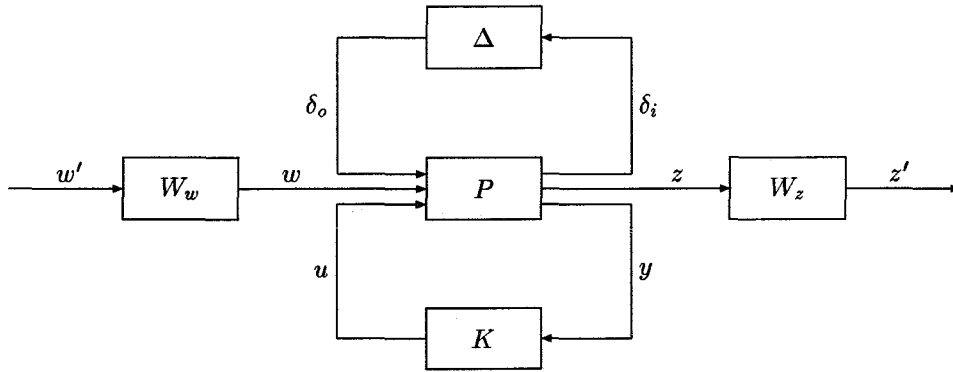
The second and third method discussed in [44] are strongly related with the principal component method. The second method provides modifications in the case sensitivity should be sacrificed for the sake of reduced sensor interaction, while the third method is based on a comparison between the overall (all candidate sensor locations) SVD analysis and a partial (selected sensor locations only) SVD analysis. If the plant condition number $\kappa(P(0)) = \bar{\sigma}(P(0))/\underline{\sigma}(P(0))$ is about the same for both methods, the sensor locations are probably good choices.

The fourth method is a global one, in which a partial SVD analysis is performed for each combination of sensors. This procedure is the most definitive, but also much more time-consuming. Moreover, it is stated that the method in most cases does not yield results which are significantly better than the principal component method.

3.11 Structured singular value

In [35, 36, 37], a measurement selection method is discussed in the context of the Structured Singular Value (SSV) theory. The proposed method is used for selecting secondary measurements in inferential control systems, *i.e.*, if it is difficult or impossible to directly measure the variables to be controlled ("primary measurements"), other ("secondary") measurements have to be performed. The primary measurements could then be estimated on the basis of a system model.

The method discussed is restricted to linear, time-invariant, stable plants. The assumption of stability is introduced for simplicity only. The measurement selection procedure is based on the fundamental idea that a control system must achieve *robust performance*. Candidate measurement sets for which a controller satisfying robust performance cannot be designed are undesirable and



- z' : weighted controlled variables, including manipulated variables u
 w' : weighted external inputs, *i.e.*, disturbances, measurement noise and reference signals
 δ_i : input to uncertainty block
 δ_o : output of uncertainty block

Figure 3.3: General framework for uncertain linear control systems

are systematically eliminated.

The SSV theory is used to develop a number of measurement screening tools that addresses the issue of modeling errors. Some tools are independent of the controller design methods, while others are tied to a specific controller design method. For initial screening of a large number of candidate measured variables, it is proposed to apply *general* screening tools, *i.e.*, initially those candidates are eliminated for which a linear time-invariant controller achieving robust performance does not exist, no matter what controller design method is used. If the number of candidates left is reduced to a sufficiently low level, *design-specific* screening tools can be used, *i.e.*, those candidates for which the particular design approach under consideration cannot yield a controller achieving the desired level of robust performance are eliminated. Although only measurement selection is considered, it is stated that the same norm-bound method can be applied to select an appropriate set of actuators among the available candidates.

The general framework for the systems considered is depicted in Fig. 3.3 (compare with Fig. 1.1). The model uncertainty Δ is described as a set of norm-bounded (more specifically, bounded maximum singular value) perturbations to the nominal frequency response matrix P at each frequency. It seems to be almost impossible to obtain a practically useful uncertainty description, *i.e.*, a rigorous yet nonconservative uncertainty description, that encompasses *all* system/model mismatches, including the effects of nonlinearities. Therefore, only the uncertainty that is believed to always exist and is important for closed-loop stability and performance is modeled; in [36], *e.g.*, only structured multiplicative uncertainty on the inputs and outputs is modeled. For measurement selection, such a "parsimonious" uncertainty modeling is justified, since the procedure involves eliminating undesirable candidates, for which a controller achieving robust performance cannot be found for the given uncertainty structure and level. An overly conservative uncertainty description will either leave *no* viable candidate or eliminate some of the viable candidates.

So, the concept of robust performance is used as a screening tool in the selection of measured variables. Consider Fig. 3.3 and define M_Δ to be the closed-loop transfer function matrix relating the normalized inputs w' to the weighted outputs z' , *i.e.*, $z' = M_\Delta w'$. Robust performance is achieved if:

$$\|M_\Delta\|_\infty < 1 \text{ for all perturbations considered.} \quad (3.26)$$

This condition can be tested through a measure called the Structured Singular Value (see, *e.g.*, [7]), which is referred to as μ . The development of screening tools based on the fundamental idea

represented by (3.26) will not be elaborated here, since this requires thorough knowledge of, *e.g.*, SSV theory and Youla parametrization, which is beyond the scope of this literature review.

In [35], four *necessary* conditions for testing robust performance for a given set of measurements are derived. These necessary conditions can be checked efficiently and are proposed as *general screening* tools. The first condition is a test for the existence of a causal controller achieving *nominal* performance, which is a necessary condition for robust performance. By also allowing for *noncausal* controllers, which are physically not realizable, three more necessary conditions for the existence of a controller achieving robust performance can be developed.

Some *design-specific* screening tools are also proposed in [35]. Necessary conditions for robust performance for systems under LQG control and Internal Model Control (IMC) are developed, as well as a necessary condition for controllers with integral action, regardless of their tuning.

In [36, 37] and [35, Appendix B], design-specific screening tools are developed in the context of *inferential* loop-shaping, which is an extension of the multivariable loop-shaping design technique to systems with secondary measurements. Norm-bounds on particular transfer function matrices that parametrize the controller K , *e.g.*, those functions that play similar roles as the sensitivity and complementary sensitivity functions in the standard loop-shaping problems, are derived and used as screening tools. In [37], development of tight, *sufficient* conditions to locate measurement sets for which a controller achieving robust performance exists is the focus, thereby eliminating a larger number of candidates than in the case necessary conditions are applied.

Moreover, in [37] it is stated that design-dependent screening tools for loop-shaping can also be developed for cases where the controllers are restricted to diagonal or block-diagonal configurations. This would mean that the same screening tools could be used for control configuration selection! Unfortunately, this issue is not further explained.

In [60], a method for control structure design is proposed, which combines heuristic knowledge (such as discussed in Section 3.1), simple analytical criteria and more elaborate computational routines (SSV theory). Control structure design is defined here as selection and pairing of manipulated variables, and *controlled* variables to be used for feedback. The IO-selection phase will be discussed in this section, while variable pairing, *i.e.*, the CC-selection phase, will be discussed in Section 4.12. The method is restricted to steady-state linear control system descriptions, with emphasis on applications in process control.

The first step in the IO-selection procedure, is to apply simple numerical tests to the candidate IO-sets, by which the alternatives are classified on a best-to-worst scale. An example of such a computationally fast "rank" test, is the so-called intersivity index $\underline{\sigma}(P)/\kappa^*(P)$, see also [44]. This measure trades off the minimum multivariable process gain, determined by the minimum singular value $\underline{\sigma}(P)$ (desirably large, see Section 3.8) and the sensitivity to model uncertainty, expressed by the minimum condition number $\kappa^*(P)$ (desirably small, see also Section 3.9).

After classifying all candidate IO-sets in this way, the IO-selection method proceeds to screen the alternatives. In this second step, the following practical issues are addressed:

1. **constraint satisfaction:** It is required that the magnitudes of the manipulated variables and particular other variables resulting from the control action, remain within specified bounds.
2. **robust stability:** It is required that the control system remains closed-loop stable in the presence of modeling errors.

All candidate IO-sets are individually tested for both issues and additionally for a combination of 1 and 2. This combined test addresses the ability of the control system to simultaneously handle constraints and maintain closed-loop stability for all possible modeling errors. The basis for the tests is the SSV for particular transfer function matrices at steady-state. The SSV theory allows both robustness and performance aspects of control problems to be captured simultaneously. The method is independent of a specific controller design technique; maintenance of integral control in the presence of modeling errors, is the only requirement. The development of the tests is discussed in [59].

Chapter 4

Criteria for selection of the control configuration

In this chapter, some methods to decide on the interconnections between the measured and manipulated variables are discussed. In all approaches, decentralized control is a basic assumption. As is it discussed in the Introduction, decentralized control may considerably reduce the complexity of both the controller design and the controller itself.

4.1 Stability of fixed eigenvalues

In [22, 29], the issue of fixed eigenvalues under decentralized control is discussed. Consider the following state space description of a linear constant control system:

$$\begin{aligned} \text{system:} \quad & \dot{x} = Ax + Bu \\ & y = Cx \\ \text{controller:} \quad & u = Ky. \end{aligned} \tag{4.1}$$

It is stated, that if this system is both controllable and observable, the poles of the controlled system, *i.e.*, the eigenvalues of $(A + BKC)$, can be placed completely arbitrarily (with complex poles in complex conjugate pairs) by centralized (dynamical) output feedback. The uncontrollable and unobservable poles always remain the same in case of output feedback, independent of the controller parameters and the order of the controller; the controllers even may be time-dependent and nonlinear.

Definition 4.1: fixed eigenvalues [29]

Fixed eigenvalues are defined as the eigenvalues of the system that cannot be shifted by a controller, independent of the controller parameters and the controller type, provided the controller is based only on inputs u and outputs y of the system to be controlled.

Therefore, the uncontrollable and unobservable eigenvalues are called *centrally* fixed eigenvalues. So, these eigenvalues are invariant for all possible controllers for the system. In the case *decentralized* controllers are applied, *i.e.*, $K = \text{block diag}[K_{ii}]$, additional fixed eigenvalues may appear; these are called *decentrally* fixed eigenvalues and are caused by, and dependent of, the structure of K , *i.e.*, the control configuration. Suppose there are no centrally fixed eigenvalues. The simplest way to show that an eigenvalue of A is decentrally fixed, is to calculate the eigenvalues of $(A + BKC)$ for an arbitrary feedback matrix K with the specific configuration under consideration; the eigenvalues which coincide for the uncontrolled system and controlled system correspond to the decentrally fixed eigenvalues with probability one [29].

Even if the original system is controllable and observable, the introduction of a decentralized (dynamic) feedback controller may prevent particular poles to be placed or to be shifted. This may

be highly undesirable, *e.g.*, in the case of unstable poles or stable poles close to the imaginary axis. Therefore, during the CC-selection phase, the eventual appearance of decentrally fixed eigenvalues caused by a particular control configuration has to be paid strict attention to and desirably has to be avoided. In [29], a necessary and sufficient condition is discussed to test the presence of decentrally fixed eigenvalues for a particular control configuration. The mathematical formulation of this criterion will not be elaborated here. It is expected that the condition can serve as a useful tool for screening alternative control configurations. However, a practical application has not been encountered in literature.

For *nonlinear* control systems, fixed RHP-poles correspond with unstable feedback-invariant manifolds.

4.2 Relative degree

In [12, 13], the relative degree of an output with respect to an input is used to obtain a characterization of the dynamic interactions among the input and output variables. Intuitively, this concept is based on "closeness" between controlled and manipulated variables, which is mentioned in [45] as an important heuristic commonly used in CSD. In [12], the relative degree is used to evaluate and select a control configuration among a set of *feasible* ones; concepts related to controllability, such as the system's invertibility, could be used as a criterion for feasibility, but this is not elaborated in this report. The proposed method for CC-selection is applicable to both linear and *nonlinear* control systems.

Consider square MIMO affine nonlinear systems in state space form:

$$\begin{aligned} \dot{x} &= f(x) + \sum_{j=1}^m g_j(x)u_j \\ y_i &= h_i(x) \quad i = 1, \dots, m \end{aligned} \quad (4.2)$$

with f and g_j smooth vector fields on \mathbb{R}^n and h_i smooth functions on \mathbb{R}^n . For this system, the relative degree r_{ij} of the output y_i with respect to an input u_j is defined as the smallest integer for which:

$$L_{g_j} L_f^{r_{ij}-1} h_i(x) \neq 0 \quad (4.3)$$

or $r_{ij} = \infty$ if such an integer does not exist [12]. Whenever r_{ij} is finite, it can be verified that $r_{ij} \leq n$. In (4.3), $L_\alpha^\beta \gamma$ denotes the repeated Lie derivative, see, *e.g.*, [67, Section 6.2].

The relative degree captures the dynamic effect of the manipulated variables on the measurable variables to be controlled, *i.e.*, $y = z$. Its calculation requires only structural information about the system, *i.e.*, information about the dependence of f and h_i on the state variables in x , as well as about the zero and nonzero elements in g_j . If the digraph in Fig. 3.2 is considered, $r_{ij} + 1$ is equal to the length of the shortest path connecting u_j and y_i [12]. The relative degree is thus related to the number of state variables involved in the particular IO-interaction. This is consistent with the interpretation of the relative degree as the number of integrations the input has to go through before it affects the output. In this sense, it is a meaningful measure of dynamic interaction. Furthermore, in [12] it is shown, that r_{ij} can also be interpreted as a measure of how responsive the output y_i is for step changes at the input u_j : the smaller r_{ij} , the faster the initial response of the output is.

The contribution of the concept of relative degree in CC-selection is now explained as follows. For a nonlinear system in the form of (4.2), the relative degree *matrix* is defined as:

$$R_u = \begin{bmatrix} r_{11} & \dots & r_{1m} \\ \vdots & \ddots & \vdots \\ r_{m1} & \dots & r_{mm} \end{bmatrix} \quad (4.4)$$

where the r_{ij} 's are the individual relative degrees as defined above. Next, the outputs are rearranged by exchanging rows, so that the minimum relative degree r_i in each row of R_u appears

in the diagonal, *i.e.*, R_u has the form:

$$R_u = \begin{bmatrix} r_1 & r_{12} & \cdots & r_{1n} \\ r_{21} & r_2 & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_m \end{bmatrix} \quad (4.5)$$

where $r_{ij} \geq r_i$. Under the assumption of feasibility (*i.e.*, "invertibility") of all candidate configurations, such a rearrangement is always possible [12]. The output rearrangement that leads to this form of R_u indicates the IO-pairings with the dominant interactions. Off-diagonal relative degrees in a row indicate the interaction between a specific output and the other inputs, and they are necessarily larger than or equal to the diagonal relative degree. Off-diagonal relative degrees in a column indicate the interaction between a specific input and the other outputs, and they are not necessarily larger than or equal to the diagonal relative degree. Large deviations of the off-diagonal terms both of the same row and column from the diagonal ones, imply weak dynamic structural coupling of the rest of the inputs and outputs from the particular input/output pair; therefore, from a structural point of view, employment of multiloop SISO control for the particular IO-pairing appears favorable. It is also possible to identify *groups* of inputs and outputs that are characterized by weak structural coupling from the other ones, suggesting thus favorable candidates for *block* decentralized control configurations.

It is emphasized in [12], that the guidelines for CC-selection based on the relative degree only have a *preliminary* and *qualitative* character; other analytical tools should be employed too, towards a more quantitative assessment of the merits or shortcomings of each configuration before the final selection.

4.3 Achievable performance

In Section 3.5, it is discussed that the minimum achievable value of an integral quadratic performance index can be used as a criterion to select an appropriate IO-set. Analogously, such a criterion can be applied during the CC-selection phase.

In [28], a design method for optimal decentralized controllers is discussed. This method is based on the optimization of controller parameters: the parameters in a control law based on *static* decentralized output feedback ($u = Ky$; $y = Cx$) are numerically optimized for the standard quadratic criterion:

$$\min_{K \in \mathcal{K}} \left\{ J = E \int_{t=0}^{\infty} (x^T Q x + u^T R u) dt \right\} \quad (4.6)$$

with \mathcal{K} the class of decentralized controllers with a certain structure. The choice for the decentralized structure of K , *i.e.*, the control configuration, is arbitrary, provided the system is stabilizable. Consequently, K must be chosen such, that no unstable decentrally fixed eigenvalues occur.

With respect to decentralized controllers, the minimization problem can be written:

$$\min_{K \in \mathcal{K}} \{ J = \text{tr} [P \cdot E[x(0)x(0)^T]] \} \quad (4.7)$$

with P the solution of a particular matrix Lyapunov equation: $P = P(A, B, C, K, Q, R)$ and thus $J = J(A, B, C, K, Q, R, x(0))$. So, the minimum value of J , *i.e.*, the achievable performance, depends on the control configuration represented in K . The solution K cannot be obtained analytically as for centralized optimal state feedback controllers, but it must be obtained numerically by an iterative algorithm [28]. A possible criterion for CC-selection is then, to choose the control configuration that minimizes J and therefore maximizes performance.

4.4 Relative gain

A commonly known tool for attacking the CC-selection problem for single-loop controllers, *i.e.*, fully decentralized or diagonal controllers (often called "multiloop SISO" controllers), is the so-

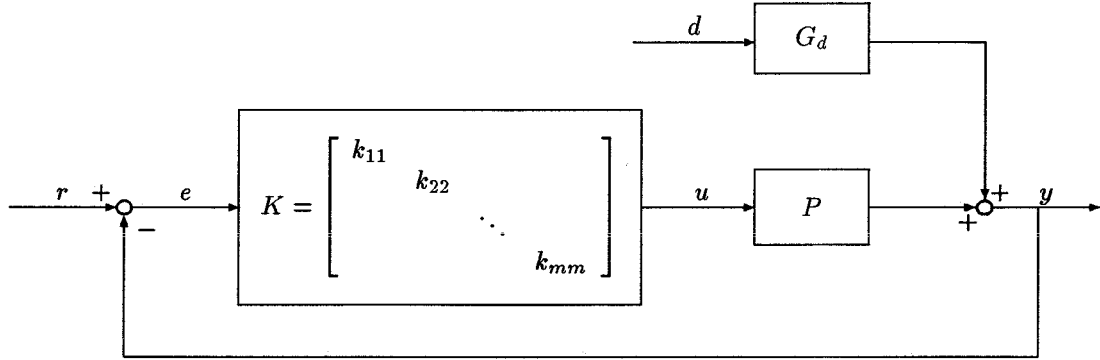


Figure 4.1: Framework of a decentralized control system with a diagonal controller

called Relative Gain Array (RGA). It has found widespread use as a measure of interaction in square linear decentralized control systems, and may be generalized to block diagonal controllers by introducing the Block Relative Gain (BRG). In this section, both the RGA and the BRG and their applications in the CC-selection procedure will be shortly discussed.

The RGA was first introduced by Bristol in 1966 [9] to answer the following question: *How is the transfer function between a given manipulated variable u_j and a given controlled variable z_i affected by complete control of all other controlled variables?*

Originally, the RGA was defined and applied at steady state, but it may easily be extended to higher frequencies, see, e.g., [23, 24, 25, 63, 64, 71] in which also applications in process control are discussed. For a further discussion of the RGA, the control system in Fig. 4.1 is considered, assuming all variables to be controlled z are represented in the measured variables y , i.e., y is to be controlled.

Consider the square $m \times m$ plant $P(s)$:

$$y(s) = P(s)u(s) \quad (4.8)$$

When all other outputs are uncontrolled, i.e., all other loops are open, the gain from input u_j to output y_i is $p_{ij}(s)$. Furthermore, writing (4.8) as $u(s) = P^{-1}(s)y(s)$, it can be seen that the gain from u_j to y_i with all the other elements in y perfectly controlled ($y_j = 0 \forall j \neq i$) is $1/[P^{-1}(s)]_{ji}$. The relative gain is now defined as the ratio of these "open-loop" and "closed-loop" gains. If this ratio is close to 1, one might infer that control of the other outputs would not have a significant impact on the control of the IO-pair under consideration. A matrix of relative gains, the Relative Gain Array, can be computed at each frequency ($s = j\omega$), using the formula:

$$\Lambda(s) = P(s) * (P^{-1}(s))^T \quad (4.9)$$

where "*" denotes element-by-element multiplication (Hadamard or Schur product). The assumption that all the other loops are perfectly controlled is in practice only justified if it is restricted to a specific range of frequencies. This is the main reason that the use of RGA has for a long time been restricted to steady state problems ($s = 0$), where perfect control may be achieved by integral action. However, the RGA proves to have a number of useful properties at high frequencies as well [24, 63]. Some of them will be discussed in this section.

The RGA as defined above has some interesting algebraic properties [38]:

1. It is independent of input and output scaling.
2. The sum of all elements in one row or column equals one, i.e., $\sum_{j=1}^m \lambda_{ij} = \sum_{i=1}^m \lambda_{ij} = 1$, with λ_{ij} the ij -th element of Λ .

3. Any permutations of rows or columns in $P(s)$ result in the same permutations in the RGA.
4. If $P(s)$ is triangular (and hence also if it is diagonal), $\Lambda(s) = I$.

The RGA is a measure of interaction and is frequently used to decide on the pairing of input and output variables, *i.e.*, which manipulated variable must be determined by which measured variable. Therefore, the RGA is a potential tool for CC-selection. In [9, 24, 25, 38, 63], some useful pairing rules are proposed, which will shortly be discussed:

1. Avoid pairings y_i/u_j (control configurations) with large RGA elements $|\lambda_{ij}|$, in particular at frequencies near cross-over [25, 63].

Large entries in the RGA indicate strong couplings and a poorly conditioned plant, which is clear from the relation $\bar{\sigma}(\Lambda(P)) \leq \kappa^*(P)$ [54], indicating that large relative gains imply a large optimally scaled, *i.e.*, minimal, condition number. The implication of large λ_{ij} 's in perturbed plants is explained in [24, 63, 66]:

- **Diagonal input uncertainty.** Consider the input disturbed plant $\tilde{P} = P(I + \Delta)$. The diagonal matrix Δ consists of the relative uncertainty in the gain of each input channel. It is shown, that if the nominal plant P has large RGA elements and an inverse-based controller is used (" $u = P^{-1}y$ "), trying to reduce interactions, the closed-loop system's performance (and even its stability) is likely to be very sensitive to diagonal input uncertainty. Therefore, an inverse-based controller should never be used for plants with large RGA elements
- **Individual element uncertainty.** Small parametric uncertainties may have large consequences for control performance. In [24, 63, 66], this issue is discussed in relation to the RGA. The following theorem is proposed and proven:

The complex matrix P becomes singular if a relative change $-1/\lambda_{ij}$ in its ij -th element occurs, *i.e.*, if a single element in P is perturbed from p_{ij} to $\tilde{p}_{ij} = p_{ij}(1 - 1/\lambda_{ij})$.

This provides a necessary and sufficient condition for singularity of a system matrix with element uncertainty, and has an important control implication. Consider a plant with transfer function matrix $P(s)$. If the relative uncertainty in an element at a given frequency is larger than $|1/\lambda_{ij}|$, the plant may have $j\omega$ -axis zeros and RHP-zeros at this frequency. As it has been discussed before (Section 3.2), RHP-zeros are highly undesirable, since they limit the achievable bandwidth for any type of controller. A large value of λ_{ij} means that only small relative errors on the corresponding plant element p_{ij} are tolerated.

Since modeling errors are always present in practice, a criterion for CC-selection that states that input-output pairings causing large RGA elements have to be avoided, seems justified.

2. Avoid pairings y_i/u_j with negative values of the steady-state RGA elements $\lambda_{ij}(0)$ for control systems using diagonal controllers with integral action.

This is because such a pairing will give the closed-loop system at least one of the following, undesirable, properties [38, 63]:

- The closed-loop system is unstable.
- The loop with the negative relative gain is unstable by itself.
- The closed-loop system is unstable if the loop with the negative relative gain is removed.

3. Prefer pairings with RGA values close to 1, especially for ω in the crossover-region.

It is emphasized that this "conventional" pairing rule is just an engineering rule of thumb, and is not based on any proof. As it is shown by an example in [24], pairing in accordance with this rule may result in an unstable system, even if the individual loops are tuned to be

stable. Desirably, $\Lambda \approx I$, since this indicates only weakly coupled control loops and therefore independent control of each loop is easier to achieve. However, it is not always possible to achieve this by row and column changes. If $\Lambda = I \forall \omega$, stability of the individual loop implies stability of the overall system [24]. Unfortunately, it is not known which deviations from $\Lambda = I$ can be tolerated without impairing stability.

4. Avoid pairings with different signs of $\lambda_{ij}(0)$ and $\lambda_{ij}(j\infty)$.

In [24, 63], it is proven that if the relative gains $\lambda_{ij}(0)$ and $\lambda_{ij}(j\infty)$ for any pairing y_i/u_j have different signs, then at least one of the following is true:

- $p_{ij}(s)$ has a RHP-zero.
- $P(s)$ has a RHP transmission zero.
- $P^{ij}(s)$, *i.e.*, the subsystem with u_j and y_i removed, has a RHP transmission zero.

However, it is emphasized that this condition is *not* a necessary one and there may be RHP-zeros present, even if the RGA elements do not change sign. As it is discussed in Section 3.2, any such zeros may be detrimental for control of the system.

One inadequacy of the RGA is, that it only measures two-way interactions, *e.g.*, $\Lambda = I$ for a triangular plant. Therefore, it may indicate that interactions are not a problem, even if significant one-way coupling exists. This is the motivation for the introduction of the so-called Performance Relative Gain Array (PRGA), see, *e.g.*, [24, 25, 26, 63]. For the discussion of the PRGA, consider the undisturbed ($d = 0$) system in Fig. 4.1. Assume that, after the variable pairings have been determined, the order of the elements in y and u has been arranged such that the diagonal elements of the plant transfer function matrix $P(s)$ correspond to the paired variables. Let $y(s)$ denote the output response for the overall system when all loops are closed ($u = Ke$) and let $e(s) = y(s) - r(s)$ denote the output error. The closed-loop response becomes:

$$e(s) = -S(s)r(s) \quad (4.10)$$

with $S = (I + PK)^{-1}$ the sensitivity function for the overall system. At low frequencies ($\omega < \omega_b$, with ω_b the frequency where $\bar{\sigma}[PK(j\omega)]$ crosses one, *i.e.*, the cross-over region), the controller gains are usually higher and $S \approx (PK)^{-1}$. Equation (4.10) can then be rewritten [25]:

$$e \approx -(PK)^{-1}r = -(\bar{P}K)^{-1}\bar{P}P^{-1}r \approx -\bar{S}\Gamma r \quad \text{for } \omega < \omega_b \quad (4.11)$$

with $\bar{S} = (I + \bar{P}K)^{-1}$, $\bar{P} = \text{diag}[p_{ii}]$ and $\Gamma = \bar{P}P^{-1}$. The matrix Γ is termed the Performance Relative Gain Array (PRGA), with elements denoted by γ_{ij} . It is noted, that the diagonal elements of Γ and Λ are the same and that the PRGA depends only on output scaling. When the effect of a setpoint change r_j on the error e_i is considered if all loops are closed, this gives:

$$e_i \approx \frac{-\gamma_{ij}}{p_{ii}k_{ii}}r_j \quad \text{for } \omega < \omega_b. \quad (4.12)$$

So, the ratio $\gamma_{ij}/(p_{ii}k_{ii})$ gives the magnitude of the error in output i to a setpoint change for output j . In fact, γ_{ij} is a measure of performance degradation due to the closing of other loops than loop i . The ratio in (4.12) is preferably small, *i.e.*, on a Bode plot, the curve for $|\gamma_{ij}|$ should lie below $|p_{ii}k_{ii}|$ at frequencies where small errors e_i are desired. This gives rise to the following pairing rules (CC-selection rules) [25]:

1. Prefer pairings y_i/u_j where p_{ij} puts minimal restrictions on the achievable bandwidth for this loop. That is, avoid pairings with RHP-zeros in $p_{ij}(s)$ and avoid pairings where $p_{ij}(s)$ is small.
2. Avoid pairings with large values of the PRGA in the crossover region, particularly if the achievable bandwidth for the corresponding loop i is restricted because of $p_{ii}(s)$, see rule 1.

As it has already been noted, the RGA may be generalized from multiloop SISO controllers to block diagonal controllers by introducing the Block Relative Gain (BRG). This interaction measure is introduced in [42]; in [49] the BRG is treated in a more mathematical context. The (left) BRG associated with the i -th square subsystem is defined as [42]:

$$\text{BRG}_{ii} = P_{ii}[P^{-1}]_{ii} \quad (4.13)$$

The diagonal blocks P_{ii} constitute the subsystems that are under decentralized control. For a multiloop SISO control system, the BRG's reduce to the diagonal elements of the RGA. Analogously to the interpretation of the RGA, the BRG of a plant P is the "ratio" of two transfer matrices of the plant, evaluated at a certain frequency; the first transfer matrix is evaluated with the remaining plant outputs uncontrolled (open loop) and the second is evaluated with the remaining plant outputs perfectly controlled. The resulting BRG is a measure for performance deterioration of each subsystem due to interactions with the other subsystems. If the plant P is block triangular, all BRG's are identity matrices.

In [42], it is stated, that an acceptable control configuration is provided, if all the BRG's corresponding to the diagonal blocks of the plant P , are close to the identity matrix. If for a particular subsystem the BRG is exactly the identity matrix, the closed-loop performance of this subsystem is as if it were isolated from the rest of the plant and operating under the influence of only its own control law. In [42], a procedure is proposed for CC-selection. The essence of the procedure is to consider candidate configurations for a decreasing degree of decentralization, and to screen out those configurations with diagonal elements and eigenvalues of the BRG's which are not "close" to 1. The BRG's associated with different configurations are evaluated both at steady state and at higher frequencies.

The concepts of RGA and BRG can be extended to *nonsquare* systems. In [10], a modified RGA is proposed for systems with $l \geq m$, *i.e.*, systems which may have a larger number of outputs than inputs. If $l > m$ it is not possible to perfectly control all outputs. Therefore, the nonsquare RGA is developed under the assumption of least-square perfect control [10] and is defined as:

$$\Lambda^N = P * [P^+]^T \quad (4.14)$$

with P^+ the pseudo-inverse of P , see, *e.g.*, [69]. It is stated that the nonsquare RGA can be used as a criterion to square the system down, *i.e.*, to select the same number of outputs as inputs, if a square control system is preferred. The outputs associated with a small row sum are eliminated to minimize the error of the uncontrolled outputs when the square subsystem is under perfect control. For properties of the nonsquare RGA, the reader is referred to [10].

In [55], the BRG for nonsquare (sub)systems is derived, also with the concept of the pseudo-inverse. This will not be further discussed here.

The major criticism related to the use of the BRG (and the RGA) as a closed-loop interaction measure is, that it is developed under the assumption of perfect control in the complementary subsystems, *i.e.*, the other subsystems under decentralized control, which will only hold for a certain frequency range. In the notion of the Dynamic Block Relative Gain (DBRG), discussed in, *e.g.*, [2, 3], perfect control is not assumed. In [3], the relation between the DBRG and stability and performance of decentralized control systems, is discussed. Unfortunately, the DBRG of a particular subsystem depends on the controllers associated with the complementary subsystems, and is therefore less appropriate for CC-selection purposes. In [55], the DBRG is extended to nonsquare systems.

In [39, 41], the (Dynamic) Nonlinear Block Relative Gain ((D)NBRG) is introduced as an interaction measure for decentralized *nonlinear* control systems.

Consider a nonlinear plant P connected with a nonlinear block diagonal feedback controller K . Except for K_1 , all controller blocks are incorporated into a single block K_2 , see Fig. 4.2. The

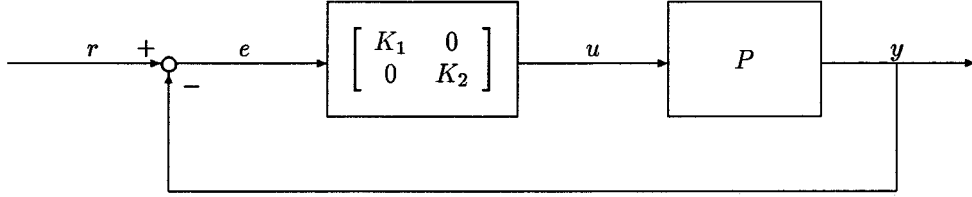


Figure 4.2: Framework of a nonlinear decentralized control system with a block diagonal controller

following relations hold:

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} P_1(u_1, u_2) \\ P_2(u_1, u_2) \end{bmatrix} = Pu \quad (4.15)$$

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = P^{-1}y \quad (\text{assuming } P^{-1} \text{ can be calculated}) \quad (4.16)$$

$$u = Ke \quad (4.17)$$

with P and K operators that describe a nonlinear system's input-output dynamic behaviour, see, *e.g.*, [52]. Following the same approach as for linear systems, the NBRG and DNBRG are developed. Again, the DNBRG appears to be controller-dependent and is therefore not very useful for CC-selection. Under the assumption of perfect control of y_2 , *i.e.*, $y_2(t) = 0 \forall t \geq 0$, the (left) NBRG associated with subsystem 1 is defined as the operator:

$$\text{NBRG}_{11} = P_{11}(P^{-1})_{11}. \quad (4.18)$$

with P_{11} the operator "relating" u_1 to y_1 if $u_2 = 0$, *i.e.*, $y_1 = P_1(u_1, u_2 = 0) = P_{11}u_1$. Contrary to the DNBRG, the NBRG depends only on plant information, and can be used for screening out those control configurations for which the NBRG is "far" from the identity operator. For a more detailed discussion and a practical 2×2 example, the reader is referred to [39, 41]. A variant of the NBRG for nonsquare (sub)systems has not been encountered in literature.

4.5 Relative sensitivity

To circumvent the disadvantages related to the BRG (assumption of perfect control) and the DBRG (dependent on controller type and *tuning*), in [4] the relative sensitivity is introduced as a closed-loop interaction measure for performance. Contrary to the (D)BRG, the relative sensitivity accounts also for one-way interaction.

Unfortunately, this measure is developed for a specific controller type, *i.e.*, decentralized Internal Model Control (IMC), see Fig. 4.3, with K_{IMC} a block diagonal controller and $\bar{P} = \text{block diag}[P_{ii}]$ consisting of the square subsystems under decentralized control. When each subsystem i is *treated in isolation*, its setpoint response, *i.e.*, its response to a change in r_i , is given by (note that r_i is a vector for subsystems with a dimension larger than 1):

$$y_i = T_{ii}^* r_i \quad (4.19)$$

with $T_{ii}^* = P_{ii}K_{IMC,ii}$ the "achievable" performance of the i -th subsystem in isolation. The setpoint response of the system in Fig. 4.3 is denoted $y = Tr$, where:

$$T = PK_{IMC}(I + (P - \bar{P})K_{IMC})^{-1} = P(T^{*-1}\bar{P} + (P - \bar{P}))^{-1} \quad (4.20)$$

with $T^* = \text{block diag}[T_{ii}^*]$. The relative sensitivity between any two subsystems is defined as:

$$v_{ji} = \left(\frac{\partial y_j}{\partial r_i} \right) \left(\frac{\partial y_i}{\partial r_i} \right)^{-1} \quad (4.21)$$

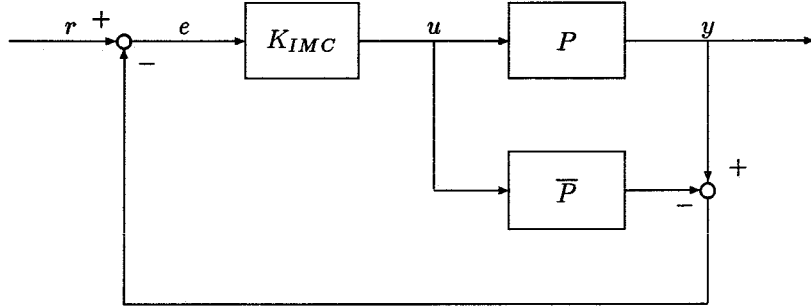


Figure 4.3: Decentralized Internal Model Control

and expresses how much the j -th subsystem is excited relative to the response of the i -th subsystem, when a setpoint change r_i in the i -th subsystem occurs. Desirably, this effect is small. All the relative sensitivities together make up the relative sensitivity matrix Υ , which becomes for k diagonal blocks:

$$\Upsilon = \begin{bmatrix} I & v_{12} & \cdots & v_{1k} \\ v_{21} & I & \cdots & v_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ v_{k1} & v_{k2} & \cdots & I \end{bmatrix}. \quad (4.22)$$

For a totally noninteracting system, Υ is identity, while for a one-way interacting system it is either upper or lower triangular, depending on the direction of the interactions.

In [4], the magnitudes of the interactions are denoted by $\bar{\sigma}(\Upsilon_i T_{ii})$, where Υ_i is composed of the off-diagonal blocks in the i -th column of Υ . These magnitudes should be kept as small as possible. Therefore, $\bar{\sigma}(\Upsilon_i)$ should be small (compared to 1) within the same bandwidth for which $\bar{\sigma}(T_{ii}(\omega)) \approx 1$. The larger this bandwidth is, the better the dynamic interactions are rejected.

If $T_{ii}^* = I$ at $\omega = 0$, it is concluded from equation (4.20) that $T(\omega = 0) = I$ also, and therefore $\Upsilon(\omega = 0) = I$. In [4], the following is derived for high frequencies:

$$\bar{\Upsilon} = \lim_{\omega \rightarrow \infty} \Upsilon = \begin{bmatrix} I & P_{12}P_{22}^{-1} & \cdots & P_{1k}P_{kk}^{-1} \\ P_{21}P_{11}^{-1} & I & \cdots & P_{2k}P_{kk}^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ P_{k1}P_{11}^{-1} & P_{k2}P_{22}^{-1} & \cdots & I \end{bmatrix}. \quad (4.23)$$

So, the values of these high-frequency asymptotes do not depend on controller tuning (remind that Υ is developed for IMC, so the controller *type* is fixed), and are therefore a potential tool for CC-selection. The following guideline is proposed in order to screen out some undesirable configurations [4]:

- Avoid configurations with asymptotes that persistently exceed 1, since they will decrease the bandwidth within which $\bar{\sigma}(\Upsilon_i)$, and therefore interactions, can be made small. Specifically, configurations must be chosen that minimize the magnitudes of the high-frequency asymptotes $\bar{\sigma}(\bar{\Upsilon}_i)$, with $\bar{\Upsilon}_i$ composed of the off-diagonal blocks in the i -th column of $\bar{\Upsilon}$.

The relative sensitivity for systems with more inputs than outputs is discussed in [55].

4.6 Closed-loop disturbance gain

The development of the frequency-dependent closed-loop disturbance gain is closely related to that of the PRGA, see, *e.g.*, [23, 24, 25, 26, 63]. Consider the control system in Fig. 4.1 in which

d denotes disturbances. Assume the controller is diagonal and suppose $r = 0$. Furthermore, G_d may be nonsquare. In analogy to (4.10), the closed-loop response to disturbances becomes:

$$e(s) = S(s)G_d(s)d(s). \quad (4.24)$$

Again, for frequencies below the cross-over frequency ω_b , it is assumed $S \approx (PK)^{-1}$. Equation (4.24) can then be written in the following form [25]:

$$e \approx \bar{S}\Delta d \quad \text{for } \omega < \omega_b \quad (4.25)$$

with $\bar{S} = (I + \bar{P}K)^{-1}$, $\bar{P} = \text{diag}[p_{ii}]$ and $\Delta = \bar{P}P^{-1}G_d$. The matrix Δ is termed the Closed Loop Disturbance Gain (CLDG), with elements denoted by δ_{ij} . It is scaling dependent, as it depends on the magnitude of disturbances and outputs. The approximate effect of a disturbance d_j on the error e_i is given by (compare with (4.12)):

$$e_i \approx \frac{\delta_{ij}}{p_{ii}k_{ii}}d_j \quad \text{for } \omega < \omega_b. \quad (4.26)$$

So, the ratio $\delta_{ij}/(p_{ii}k_{ii})$ gives the magnitude of the error in output i to a disturbance d_j , which is preferably small. That is, the curve for the CLDG element $|\delta_{ij}|$ should lie below $|p_{ii}k_{ii}|$ at frequencies where large errors e_i have to be avoided.

To get a better physical interpretation of the CLDG, consider the response to a disturbance d_j when all the other loops are open:

$$e_i \approx \frac{g_{d_{ij}}}{p_{ii}k_{ii}}d_j \quad \text{for } \omega < \omega_b. \quad (4.27)$$

Comparing (4.26) and (4.27), $g_{d_{ij}}$ is replaced by δ_{ij} , which explains why the name closed loop disturbance gain is chosen for δ_{ij} .

For process control, disturbance rejection is usually more important than setpoint tracking, *i.e.*, CC-selection for process control systems should be directed by the CLDG rather than by the PRGA discussed in Section 4.4. Of course, it is also possible to use a combination of both measures for CC-selection. From the point of view of disturbance rejection, one possible pairing rule is [25]:

- Avoid control configurations (pairings) with large CLDG values in the crossover region, particularly if the achievable bandwidth ω_b for the corresponding loop is restricted due to p_{ii} , *e.g.*, due to RHP-zeros (see also Section 4.4).

4.7 Interaction potential

In [27], an approach is proposed for selecting IO-pairings in diagonal controllers, which consists of two steps. The first is to eliminate as many of the infeasible pairings as possible by applying a selection criterion based on the steady-state RGA; the second step is to determine *preferred* IO-pairings by using the "interaction potential", which is a measure that provides information on the *possible* interaction on a particular loop by all the other loops.

Consider the undisturbed ($d = 0$) multiloop SISO control system in Fig. 4.1. The output y can be written:

$$y = PK(I + PK)^{-1}r = Tr \quad (4.28)$$

with T the complementary sensitivity function matrix of this system. For an *independent* single loop, the associated complementary sensitivity function is defined as:

$$\bar{T}_{ii} = \frac{p_{ii}k_{ii}}{1 + p_{ii}k_{ii}}. \quad (4.29)$$

In [27], it is demonstrated, that the controller itself has profound effects on the loop interactions, see also the remarks on the DBRG in Section 4.4. Therefore, CC-design is coupled with controller

design. In [27], it is stated that under the assumption of Internal Model Control, it is possible to formulate the system's performance with the following factorization:

$$\bar{T}_{ii} = \chi_{ii} f \quad (4.30)$$

where χ_{ii} contains all of the NonMinimum Phase (NMP) parts of p_{ii} and f designates the *desired* dynamics of an independent loop, and is assumed to have the following form:

$$f_{ii} = f = \frac{1}{\tau s + 1}. \quad (4.31)$$

The motivation for the introduction of (4.31) is not quite clear from [27]. Moreover, the approach may not always be justified since a controller k_{ii} which fulfills (4.30) may not always be realizable.

Each loop in the system may be affected by the outputs of all other loops [27]. If the i -th output and j -th input are to be configured as a SISO loop, the strength of these "disturbances", *i.e.*, interactions, by other loops is represented by a frequency dependent vector $h^{ij}(j\omega)$. For details on the development of h^{ij} , the reader is referred to [27]. The vector h^{ij} thus contains information on interactions and depends, among others, on the functions f and χ_{ij} , which is defined as a factor that contains all the NMP parts of p_{ij} . The integral of the frequency-scaled 2-norm of h^{ij} is designated ϕ_{ij} :

$$\phi_{ij}(\tau) = \frac{1}{\pi} \int_0^\infty \|h^{ij}(j\omega)/\omega\|_2^2 d\omega. \quad (4.32)$$

The average of ϕ_{ij} over a particular interval of τ is termed the *interaction potential* associated with p_{ij} , *i.e.*:

$$\bar{\phi}_{ij} = \frac{1}{\tau^u - \tau^l} \int_{\tau^l}^{\tau^u} \phi_{ij}(\tau) d\tau \quad (4.33)$$

where $[\tau^l \ \tau^u]$ is a region of ϵ chosen for averaging ϕ_{ij} . Finally, the interaction potential matrix can be written:

$$\bar{\Phi} = \begin{bmatrix} \bar{\phi}_{11} & \cdots & \bar{\phi}_{1m} \\ \vdots & \ddots & \vdots \\ \bar{\phi}_{m1} & \cdots & \bar{\phi}_{mm} \end{bmatrix}. \quad (4.34)$$

Calculation of $\bar{\Phi}$ only requires process information in terms of $P(s)$, and a region for the parameter τ ; $\bar{\Phi}$ is independent of the control configuration.

The interaction potential $\bar{\phi}_{ij}$ is a lumped sum of the strengths of possible interactions from all of the other loops to the loop configured by the pairing y_i/u_j . Thus, possibly with some weightings, the interaction potentials of each loop in a multiloop system can be summed to a *total* interaction potential for the control system. As a result, the most preferable control configuration (IO-pairings), is the one with the least total interaction potential.

The procedure for developing preferable IO-pairings as discussed in [27], is as follows:

1. *Eliminate infeasible pairs:*

Calculate the steady-state RGA and eliminate candidate pairings with $\lambda_{ij} < 0$ to avoid stability problems (see Section 4.4). Hopefully, the number of candidate pairings is significantly reduced in this way, by which the investigation of dynamic interaction in the subsequent steps becomes less time-consuming.

2. *Find a preliminary preferable pairing:*

Assign an appropriate region $[\tau^l \ \tau^u]$. The choice of τ^u is related to the choice of τ^l and is dependent on the modeling errors of interest. Start with a small τ^l and compute $\bar{\Phi}$. Choose m numbers from $\bar{\Phi}$, all from different rows and columns, such that the *sum* of the numbers is minimal, *i.e.*, choose the configuration with the least interaction. For each number $\bar{\phi}_{ij}$ selected, its position in the matrix describes the IO-pairing for this *preliminary* configuration.

3. *Find the most preferable pairing by iteration:*

This is a checking procedure, since the current values of τ^l and τ^u may not be "feasible" for the temporarily chosen pairing of step 2: the stability of the multiloop control system indicates the feasibility of the current τ^l . Since the factorization of the performance of the control system (4.30) corresponds to one commonly used in IMC design with SISO controllers, τ^l is chosen based on its ability to stabilize the closed-loop system. If the value of τ^l is feasible, the current configuration is the most preferable one. If it is not, the values of τ^l and τ^u are increased and steps 2 and 3 are repeated until an appropriate τ^l and the most preferable configuration are found.

4.8 Numerical invertibility

In [43], a quantitative measure of the best pairing of IO-variables in a multiloop SISO control system is developed, based on the proposition that the best pairing is the one for which the system most closely resembles a set of independent single-loop systems, *i.e.*, the one for which the system can be represented best by the diagonal terms of its steady state gain matrix. The criterion is developed by an analysis of an iterative technique for obtaining the solution to a set of linear equations, or the inverse of a matrix. It is shown that the system interactions can be assessed by analyzing the effect of the off-diagonal elements of $P(0)$ on the difficulty of obtaining the desired solution or matrix inverse by the iterative procedure. The rate of convergence of the iterative process for the different variable pairings provides the basis for the pairing criterion. Although the development of the criterion is based on purely algebraic properties, its practical validity is shown to be evident from the relationship between the pairing criterion and stability characteristics of the system.

Consider the $m \times m$ system $y(s) = P(s)u(s)$. The changes dy from one *steady state* to another resulting from changes in the manipulated variables du are given by:

$$dy = P(0)du \quad (4.35)$$

in which the elements $p_{ij}(0)$ of $P(0)$ are commonly called the "process gains". In the rest of this section $P(0)$ will be denoted P . If the process gains are constant over the range of operation for all choices of du of interest, (4.35) may be integrated to give:

$$\Delta y = P\Delta u. \quad (4.36)$$

Provided the inverse of P exists, (4.36) may be solved for Δu to give:

$$\Delta u = P^{-1}\Delta y. \quad (4.37)$$

In general, the ease of obtaining P^{-1} by iterative techniques increases with an increase in the *dominance* of the *diagonal* elements p_{ii} . Since the best pairing of variables is believed to be the one that makes the system most closely resemble a set of independent single-loop systems, it is also the one whose off-diagonal terms make the smallest contribution to the inverse of the gain matrix P . Therefore, the pairing criterion developed in [43] is based on an analysis of the difficulty of finding P^{-1} , *i.e.*, it is based on a measure of the contribution of the interaction elements to P^{-1} .

In [43] it is shown, that after N iterations with $\bar{P}^{-1} = (\text{diag}[p_{ii}])^{-1}$ as an initial guess for P^{-1} , the following approximation $(P^{-1})_N$ of P^{-1} is obtained:

$$(P^{-1})_N = (I + J + J^2 + \dots + J^N)\bar{P}^{-1} \quad (4.38)$$

with N the number of iterations and J the Jacobi iteration matrix defined by:

$$J = - \begin{bmatrix} 0 & j_{12} & j_{13} & \cdots & j_{1m} \\ j_{21} & 0 & j_{23} & \cdots & j_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ j_{m1} & j_{m2} & j_{m3} & \cdots & 0 \end{bmatrix} \quad (4.39)$$

with $j_{ij} = p_{ij}/p_{ii}$. The necessary and sufficient condition for the matrix series (4.38) to converge to P^{-1} is that the largest eigenvalue of J in modulus is less than unity (see, *e.g.*, [69, Chapter 7]). The convergence rate is defined as follows:

$$R(J) = -\log \rho(J) \quad (4.40)$$

with $\rho(J)$ the spectral radius of J , *i.e.*, the largest modulus of the eigenvalues of J , see, *e.g.*, [69, Chapter 7]. Thus, a quantification for the pairing that makes the system most closely resemble a set of single-loop systems has been obtained, and a CC-selection criterion is stated:

- Of all possible pairings (control configurations) choose that pairing whose corresponding Jacobi iteration matrix J has the smallest spectral radius.

In [43], it is shown that the eigenvalues of J not only indicate the desired pairings, but they also indicate whether the configuration would be stable with a perfect diagonal controller, defined as a controller for which $K(0) = \bar{P}^{-1}(0)$, or with a diagonal controller containing integral action. Some examples are presented to illustrate the usefulness of the selection criterion, and criteria for CC-selection based on the RGA (Section 4.4), the SVD (Section 4.11) and the criterion proposed above are compared.

4.9 Performance degradation

In [21, 54, 56], the CC-selection is based on a performance degradation criterion for decentralized control systems. The proposed algorithm has been implemented in a MATLAB toolbox [57].

Consider the controller $u(s) = K(s)y(s)$. It is assumed that the selected measurements in y represent the variables to be controlled. Partitioning of feedback interconnections among the manipulated and controlled variables, *i.e.*, CC-selection, leads then to a decentralized control scheme. When a system is not partitioned, each manipulated variable action is determined by feedback from all the measured variables; full information exchange takes place and the control system is fully connected or centralized. This means that a decentralized control system will generally be less complex than a centralized one. However, closed-loop system performance may suffer, since the amount of information exchange is reduced. That is, the effect of manipulated variable actions on measured variables which were not used to compute those actions, the so-called cross-feed, may degrade the overall system performance significantly. Cross-feed can degrade the system performance in a number of ways [54]. Not only the *nominal performance* of the system can be endangered, but the *nominal stability* of the closed-loop system as well. Furthermore, the *robust stability and performance* characteristics of the closed-loop system can degrade as a result of cross-feed. Finally, the cross-feed affects the stability of the system when blocks of the controller are tuned individually or taken out of service. A necessary condition for low cross-feed performance degradation is therefore a potential CC-selection criterion, by only allowing configurations which pass the criterion; these configurations are termed viable.

To determine an appropriate definition of cross-feed performance degradation, the closed-loop performance of the partitioned system $P(s)$ with the decentralized controller $K(s)$ is contrasted with that of the associated block diagonal system $\bar{P}(s)$ with the decentralized controller $K(s)$ (see Fig. 4.1 with the diagonal controller replaced by a block diagonal one). That is, if the off-diagonal blocks of a system were not present and a decentralized controller with the same block structure were employed, cross-feed would not take place. In this case, each of the subsystems P_{ii} are controlled independently. Thus the performance of the block diagonal system, composed of *independent* subsystems, would be considered the "ideal", which the decentralized system's performance should approximate.

Closed-loop performance is studied by using the complementary sensitivity function of the partitioned system $T = PK(I + PK)^{-1}$ and that of the associated block diagonal system $\bar{T} = \bar{P}K(I + \bar{P}K)^{-1}$. Here performance includes not just nominal performance issues such as setpoint tracking, but also the stability and robustness properties of the closed-loop system [54, 56]. That

is, T and \bar{T} specify completely all aspects of respectively the closed-loop performance of the partitioned and block diagonal systems. The extent to which T and \bar{T} differ characterizes the impact of the cross-feed on the closed-loop performance. Therefore, one measure of cross-feed performance degradation is the "difference" between T and \bar{T} , relative to the "ideal" performance represented by \bar{T} :

$$\bar{\sigma}[(T - \bar{T})\bar{T}^{-1}] \quad (4.41)$$

In [18], a performance degradation measure closely related to this one is discussed. The proposed CC-selection criterion is based on a necessary and sufficient condition for low cross-feed performance degradation [54, Appendix B]:

$$\bar{\sigma}[(T - \bar{T})\bar{T}^{-1}] \leq d_T \quad (4.42)$$

with d_T the specified maximum allowable cross-feed performance degradation margin at each frequency. This condition is then weakened to yield a *necessary* condition for low cross-feed performance degradation [54, Appendix B]:

Suppose P is a square FDLTI plant with its measurements and manipulations partitioned such that:

$$P = [P_{ij}] = \begin{bmatrix} P_{11} & \dots & P_{1k} \\ \vdots & \ddots & \vdots \\ P_{k1} & \dots & P_{kk} \end{bmatrix}$$

Under these conditions, there exists a FDLTI controller $K = \text{block diag}[K_1, \dots, K_k]$ which achieves

1. $\bar{\sigma}[(T - \bar{T})\bar{T}^{-1}] \leq d_T$, and
2. $\bar{\sigma}(\bar{T}) \leq \sigma_{\bar{T}}$, $\sigma_{\bar{T}} < 1 \forall \omega \geq \omega_{\bar{T}}$

only if:

$$\frac{(1 - \sigma_{\bar{T}})\bar{\sigma}(V)}{1 + (1 - \sigma_{\bar{T}})\bar{\sigma}(V)} \leq d_T \forall \omega \geq \omega_{\bar{T}} \quad (4.43)$$

where:

- $V = (P - \bar{P})P^{-1}$ for $\bar{P} = \text{block diag}[P_{11}, \dots, P_{kk}]$
- $\sigma_{\bar{T}}$ and $\omega_{\bar{T}}$ specify the closed-loop bandwidth of the block diagonal system.

Note that this condition is *quantitative*, since performance specifications are stated explicitly in provisions 1 and 2. As it is discussed in Section 3.9, these performance specifications on the measurements y must be derived from specifications on the performance variables z . By shaping \bar{T} , *i.e.*, choosing $\sigma_{\bar{T}}$ and $\omega_{\bar{T}}$, the closed-loop performance of the "ideal" system is completely specified. In the CC-selection procedure, by the function d_T it can be specified how much the partitioned system may vary from the ideal one as a result of cross-feed. The design parameter d_T serves as an upper bound on the maximum cross-feed performance degradation margin permitted at each frequency. For example, if offset-free steady-state behavior is desired, d_T is zero at $\omega = 0$.

The condition (4.43) is also *efficient*, since the candidate control configurations can be evaluated prior to control law design. However, it could be made significantly more efficient if the scaling dependence via $\bar{\sigma}(V)$ could be relaxed. In [54, 56], it is shown that this can be achieved by introducing the "partial row sums" Ψ_i of the elements lying in the diagonal blocks of the RGA $\Lambda(P)$:

$$\Psi_i = \sum_{j=\alpha_k}^{\beta_k} \Lambda(P)_{ij} \quad (4.44)$$

where i denotes the i -th row and k denotes the k -th diagonal block of $\Lambda(P)$, and α_k and β_k represent respectively the beginning and ending column indices of the k -th diagonal block of $\Lambda(P)$. The "complementary partial row sum" is defined by:

$$\bar{\Psi}_i = 1 - \Psi_i. \quad (4.45)$$

So, $\bar{\Psi}_i$ corresponds to the sum of the RGA elements on the i -th row lying outside the diagonal block. In [54, 56], it is shown that these definitions can be employed as follows:

The CC-selection procedure remains valid if the following inequality is substituted for (4.43):

$$\frac{(1 - \sigma_{\bar{T}})|\bar{\Psi}|_{max}}{1 + (1 - \sigma_{\bar{T}})|\bar{\Psi}|_{max}} \leq d_T \quad \forall \omega \geq \omega_{\bar{T}} \quad (4.46)$$

where $|\bar{\Psi}|_{max} = \max_i |\bar{\Psi}_i|$ is the maximum absolute value complementary partial row sum of $\Lambda(P)$, for the configuration corresponding to \bar{P} .

This modification is a viability criterion based on a *weaker* necessary condition for low cross-feed performance degradation. Therefore, more candidate configurations will pass this criterion. Note that in (4.46) the RGA must be computed at and above the specified closed-loop bandwidth, while traditionally the RGA is evaluated only at steady state, see also the discussion in Section 4.4. The CC-selection method discussed in this section may *not* be *effective*, since it is only based on necessary conditions. This implies that configurations passing (4.43) or (4.46) need not be feasible in the sense that low performance degradation is guaranteed.

4.10 Nominal stability and closed-loop integrity

In [73], the so-called Niederlinski Index (NI) is proposed as a criterion to select input-output pairings for multiloop SISO controllers (see Fig. 4.1). Consider square $m \times m$ open-loop stable plants P . The Niederlinski index is defined as:

$$NI = \frac{\det[P(0)]}{\prod_{i=1}^m p_{ii}(0)}. \quad (4.47)$$

So, the NI is a steady-state measure. The importance of the NI in CC-selection is explained as follows: if all the SISO controllers contain integral action and have positive loop gains, a negative value of the NI is a sufficient condition for instability of the closed-loop system with this particular variable pairing for *any* controller tuning. A positive value for NI is thus a *necessary* condition for stability, and candidate control configurations with $NI < 0$ are eliminated.

In [11, 18], the condition on the NI is generalized to block diagonal controllers and open-loop unstable systems. For block diagonal controllers, the Niederlinski index associated with the plant $P(s)$ is defined by:

$$NI = \det[P(0)\bar{P}(0)^{-1}]. \quad (4.48)$$

with $\bar{P} = \text{block diag}[P_{ii}]$. The NI is used as a necessary condition for the closed-loop stability of the decentralized control system:

Assume that:

- $P(s)$ and $\bar{P}(s)$ have the same RHP poles (4.49)

- $P(s)K(s)$ is strictly proper (4.50)

- $\bar{T}(s)$ is stable (4.51)

- $\bar{T}(0) = I$ (integral action) (4.52)

then the closed-loop system $T(s)$ will be stable only if:

$$NI > 0. \quad (4.53)$$

Here, T and \bar{T} are defined similarly as in Section 4.9. In [11], the NI is used in combination with the BRG (see Section 4.4) to formulate CC-selection rules for multiloop MIMO control systems, *i.e.*, for control systems with block diagonal controllers. The criterion for the CC-selection problem is based on *integrity* considerations: it is required that any decentralized control system is stabilized by a stable controller having integral action and maintains its nominal stability in the face of *any combination of loop failures*. A control system with these properties is termed to possess Decentralized Closed-Loop Integrity (DCLI). Note that a *single loop* may be a MIMO system itself, in the case of *block* diagonal controllers. Necessary conditions for DCLI are developed and CC-selection criteria based on DCLI are discussed. The results pertain only to stability, and dynamic performance is not addressed.

Before the rules are introduced, a few notational aspects have to be explained. Let J_k be the set of integers $J_k = \{1, \dots, k\}$ and J_l be a subset of J_k containing l ($1 \leq l \leq k$) elements of J_k . Let A_{J_l} denote the corresponding principal submatrices of A consisting of blocks A_{ij} with indices belonging to J_l . In the following discussion, " k -channel" plants are considered, *i.e.*, controllers $K(s)$ with k diagonal blocks are considered. The diagonal blocks P_{ii} ($i = 1, \dots, k$) constitute the subsystems under decentralized control. In the development of the selection rules, assumptions (4.49)-(4.52) are made and additionally $\det[P_{J_l}(0)] \neq 0$ for $l = 1, \dots, k$.

The following CC-selection rules based on the notion of DCLI are proposed and proven in [11]:

1. Select control configurations with $\det[\text{BRG}_{ii}[P(0)]] > 0$ for $i = 1, \dots, k$. In this case, one of the following properties is obtained:
 - The closed-loop system can be stabilized and remains stable after the failure of any i -th loop only if $\text{NI}[P(0)] > 0$.
 - The closed-loop system is unstable and so is the reduced system after the failure of any i -th loop if $\text{NI}[P(0)] < 0$.
2. Select control configurations with $\text{NI}[P(0)] > 0$, which is necessary for stabilizability.

Control configurations passing rules 1 and 2 may achieve DCLI against failures of any i -th loop, *i.e.*, may achieve DCLI if one controller block is out of service. The final CC-selection rule is now proposed as:

3. (a) Select configurations with $\det[\text{BRG}_{ii}[P_{J_l}(0)]] > 0$ for $l = 2, \dots, k-2$ and $i = 1, \dots, l$, or, equivalently:
 - (b) Select configurations with $\text{NI}[P_{J_l}(0)] > 0$ for $l = 2, \dots, k-2$.

Configurations passing rule 3 may possess DCLI against *any* combination of $(k-l)$ loop failures. Since the selection rules 1-3 are based on necessary conditions, a CC-selection procedure based on these rules need not be effective. By inspecting rules 1, 2 and 3 successively, an increasing amount of alternative configurations is eliminated; the remaining alternatives may be further screened using other methods.

The concept of DCLI is closely related with that of Decentralized Integral Controllability (DIC), which is introduced in [47, Chapter 14] in the context of multiloop SISO controllers. A plant P is DIC if there exists a diagonal controller with integral action such that:

- the closed-loop system is stable,
- the gains of any combination of loops can be reduced independently with a factor ϵ_i ($0 \leq \epsilon_i \leq 1$) without affecting the closed-loop stability.

This implies that the loops can be detuned or taken out of service while maintaining stability. In [47], necessary conditions for DIC are developed, which are commonly used for screening candidate control configurations. In [11], it is stated that DCLI is a necessary condition for DIC.

4.11 Singular value decomposition

In [34], an approach based on the SVD technique provides a quantitative measure of interactions in multiloop SISO control systems. In order to encompass both static and dynamic features, the analysis should be carried out over a range of frequencies of practical importance *e.g.*, over the frequency band of characteristic disturbances.

The systems considered are assumed to be described by the input-output relation $y(s) = P(s)u(s)$ where $P(s)$ is an $l \times m$ transfer function matrix; the number of measured variables l may thus differ from the number of manipulated variables m . Furthermore, the controller $K(s)$ is assumed to be diagonal with dimension $\min(l, m)$; which measurements or manipulated variables will be deleted depends on the control configuration to be proposed. The SVD of the transfer function matrix $P(s)$ can be written as [34]:

$$P(s) = W(s)\Sigma(s)V^H(s) \quad (4.54)$$

with:

$$\Sigma(s) = \begin{bmatrix} \Omega_{q \times q}(s) & 0_{q \times (m-q)} \\ 0_{(l-q) \times q} & 0_{(l-q) \times (m-q)} \end{bmatrix} \quad (4.55)$$

with $q = \text{rank}(P(s)) \leq \min(l, m)$. The entries of the diagonal matrix $\Omega(s)$ define the singular values of $P(s)$. With (4.54), the transfer function matrix $P(s)$ can be interpreted geometrically: an input vector in the direction of $v_i(s)$ propagates through the input space, is scaled by the gain $\sigma_i(s)$, and reappears in the output direction $w_i(s)$.

A measure is developed in terms of singular values and left and right singular vectors, which expresses the extent of interaction of the ij -th loop (" $y_i = p_{ij}u_j$ ") with other loops. Those loops that interact minimally with other loops, are called *natural* loops and are preferable to control the system.

The transfer function matrix $P(s)$ can be written in the following form:

$$P(s) = \sum_{i=1}^q \sigma_i Z_i(s) ; \quad Z_i(s) = w_i(s)v_i^H(s) \quad (4.56)$$

In this way, $P(s)$ is expressed as a linear combination of the q nodal contributions. Each of the nodal terms consists of a scaling factor σ_i and a rotation matrix Z_i . The entries ij with the largest absolute values in each Z_i define the natural loops y_i/u_j and thereby indicate the preferable IO-pairings [34, Appendix 1]. A frequency dependent interaction measure is developed, that quantifies the "difference" between each rotation matrix and a matrix of the same dimensions associated with the ideal natural loop, *i.e.*, a matrix with zero elements except one located at the entry ij defined by the preferable pairing y_i/u_j .

4.12 Structured singular value

In [18], the notion of Structured Singular Value (SSV) is used to define a new dynamic (*i.e.*, frequency dependent) interaction measure for square multivariable systems under feedback with diagonal or block diagonal controllers. It is stated that this measure can be used not only to predict the stability of decentralized control systems but also to measure the performance loss caused by these control structures.

Consider the undisturbed system in Fig. 4.1. The system $P(s)$ is assumed to be square ($m \times m$) and the controller is (block) diagonal. The method discussed in [18], is based on the interpretation of interactions in decentralized control systems as additive uncertainties, see Fig. 4.4. A block diagonal controller $K(s)$ is now to be designed for the system $\bar{P}(s) = \text{block diag}[P_{ii}(s)]$, such that the "nominal" closed-loop system with the transfer function matrix

$$\bar{T}(s) = \bar{P}(s)K(s)(I + \bar{P}(s)K(s))^{-1} \quad (4.57)$$

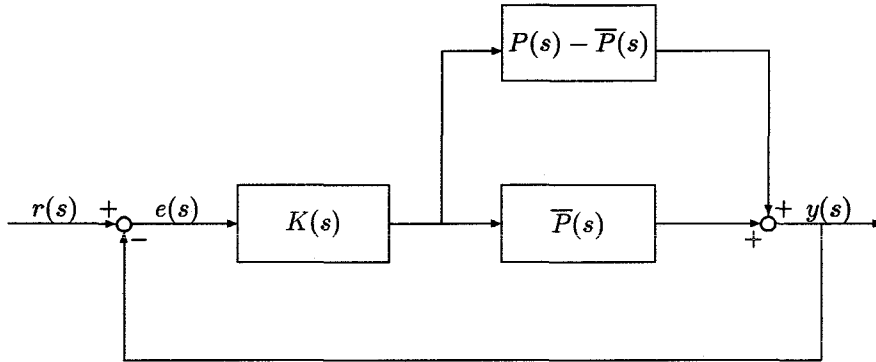


Figure 4.4: Framework of a "perturbed" decentralized control system

is stable. An *interaction measure* expresses the constraints imposed on the choice of the closed-loop transfer function matrix $\bar{T}(s)$ for the block diagonal system, which guarantee that the full closed-loop system $T(s)$ is stable. It is shown that the bound on the magnitude of $\bar{T}(s)$ is imposed by the "relative error matrix"

$$E(s) = (P(s) - \bar{P}(s))\bar{P}^{-1}(s) \quad (4.58)$$

arising from the approximation of the full system $P(s)$ by the block diagonal system $\bar{P}(s)$. More specifically: assume that $P(s)$ and $\bar{P}(s)$ have the same RHP-poles and that $\bar{T}(s)$ is stable. Then the closed-loop system $T(s)$ is stable if and only if [18]:

$$\bar{\sigma}[\bar{T}(j\omega)] < \mu^{-1}[E(j\omega)] \quad \forall \omega. \quad (4.59)$$

Equation (4.59) is in fact a robustness condition for $\bar{T}(s)$ to remain stable under the perturbation $E(s)$. It turns out that μ is the structured singular value for the analysis of feedback systems with structured uncertainties, see, e.g., [7, 38]. The value of μ depends on the structure assumed for $\bar{T}(s)$. Thus, for stability the magnitudes of the diagonal blocks \bar{T}_{ii} have to be constrained by the reciprocal of the SSV of the relative error matrix, which can be displayed on a magnitude-frequency diagram. Such plots can be used to predict stability and achievable performance of the decentralized control system.

It is not explicitly stated in [18], but it seems possible to use (4.59) in CC-selection problems, by imposing desirable properties on $\bar{T}(s)$ and then eliminating candidate configurations which do not satisfy this equation. For example, if offset-free performance is desirable, i.e., $\bar{T}(0) = I$ is imposed, candidate configurations for which $\mu[E(0)] \geq 1$ must be eliminated. Another possibility is to screen out configurations which cannot achieve the desired closed-loop bandwidth specified in \bar{T} .

In [59, 60], a CC-selection procedure is proposed consisting of two steps. The first step is to "rank" candidate configurations on a best-to-worst scale by applying the steady-state RGA or BRG. The second step is to apply SSV-theory for testing nominal Decentralized Integral Controllability (DIC, see Section 4.10 for the definition) and *combined* constraint satisfaction, robust stability and DIC (see also Section 3.11). In [59], it is claimed (not proven) that, under the assumptions of stable P , \bar{P} and \bar{T} , the unperturbed plant P is DIC with respect to \bar{T} if:

$$\mu(E(0)) < 1 \quad (4.60)$$

where μ is computed based on the structure denoted by \bar{T} . For the development of the combined test, which is also applied at steady state, the reader is referred to [59]. Both the μ -test for DIC and the combined μ -test can be used to eliminate undesirable candidate control configurations.

Chapter 5

Applications of control structure design

Applications of the various IO- and CC-selection methods discussed in Chapter 3 and 4 are reviewed in Section 5.1 of this chapter. All examples are based on numerical simulations; experimental results have not been encountered in literature. Section 5.2 proposes a different physical example, which incorporates a number of relevant aspects to be accounted for in CSD. This example is introduced to evaluate both existing IO- and CC-selection methods, and methods to be developed.

5.1 Applications from literature

In [58], the IO-selection method based on control power and speed (Section 3.1) is used to decide on regulatory control structures for a distillation column. Unfortunately, the approach is only qualitative, *i.e.*, control power and speed are designated "small", "moderate", respectively "strong", and "slow" or "fast".

In [46], the IO-selection method based on controllability and observability (Section 3.3) is illustrated by the following examples from process control: mixer-blender, double effect evaporator, 8-tray distillation column, and Williams-Otto plant. Detailed descriptions of these systems are not provided in [46]. The examples are used to illustrate the IO-selection procedure proposed; closed-loop evaluations of the selected IO-sets are *not* performed. In [16, 30, 61], examples from process control are used to illustrate some features of the structural concepts proposed in these papers. Again, only *structural* models are provided.

In [17], another process control example is treated, which is a combination of three units, *i.e.*, a mixer, a blender, and a heat exchanger. Linear models of the individual units are given. The IO-selection procedure based on cause-and-effect graphs (Section 3.4) is illustrated, but subsequent steps in control system design are *not* performed.

The IO-selection procedure based on accuracy of state estimates (Section 3.6) is performed for placement of temperature sensors (from seven collocation points) in a tubular reactor [32], a model of which is provided. The effect of varying measurement noise and system noise intensity on the optimal sensor locations is investigated. Again, an evaluation of alternative output sets by closed-loop simulations is *not* performed.

In [48], economically optimal IO-sets (Section 3.7) are selected for a double-effect evaporator and a flotation circuit. In both cases, some promising IO-sets are generated. The controllability of the alternative sets is evaluated by condition number plots. Literature to obtain plant models is referred to.

The use of the MRI (Section 3.8) for input selection is illustrated in [73] by means of a distillation column example. From plots of $\underline{\sigma}(P)$ associated with the candidate IO-sets, the optimal set is selected. Concepts based on RGA, NI, and Integral Controllability are used to decide on preferable

IO-pairings. The following analyses are performed to compare the candidate control structures. To investigate closed-loop stability (using SISO PI controllers), characteristic loci are plotted for the linearized system, and step responses are simulated for the nonlinear system. Furthermore, the alternative control structures are assessed for their ability to reject load disturbances, and their closed-loop robustness properties, using two simple measures [73].

In [54, 56], the IO-selection approach based on the condition number of the plant model (Section 3.9), is treated for a heat exchanger network with 8 candidate inputs and outputs. The system model can be obtained from [62].

In [21, 54], CSD is illustrated by an aircraft example, for which a reduced state space model (7 states) is given. Candidate square subsystems (IO-sets) are selected from 9 candidate outputs and 4 candidate inputs. For some subsystems selected, the CC-selection method based on performance degradation (Section 4.9) is illustrated. Closed-loop simulations to compare the "most promising" control structures are not performed. In [15], a different aircraft problem (obtained from [38, Appendix A]) is treated using the same criteria for IO-selection and CC-selection. Three candidate inputs and outputs are considered. Closed-loop simulations are performed, applying multivariable PI control [15, Appendix 3].

In [54, Chapter 7], a large-scale problem (50 states, 13 candidate outputs, 12 candidate inputs) from chemical process industry is studied: the Tennessee Eastman plant-wide control problem. CSD is performed for a linearized model, for which numerical values are also provided. The IO-selection method based on the condition number of the plant and the CC-selection method based on performance degradation are illustrated by this example. Some special interest is paid to the implications of integrators for IO-selection.

In [54, 56], CC-selection based on performance degradation is applied to an industrial variable cycle engine ($m = l = 3$). Unfortunately, details of this system are proprietary. A boiler furnace ($m = l = 4$) is proposed as a second example to illustrate the CC-selection approach. The control configurations selected are compared with those in [42], where the same system is considered and the BRG (Section 4.4) is used to decide on preferable control configurations.

In [31], an input selection procedure based on SVD (Section 3.10) is illustrated for a 20th order boiler model with 9 candidate inputs. To decide on the *number* of dominant inputs to be retained, LQ controllers with different numbers of inputs are designed, and the performance indices are studied. The SVD is often used to indicate proper locations for temperature measurements: in [6], locations for 2 temperature sensors in a 17-tray benzene-toluene distillation column are selected; in [44], placement of 2 temperature sensors in a 50-tray ethanol-water distillation column is considered. For both examples, the steady-state SVD is the only system data provided. Unfortunately, closed-loop simulations supporting the choice of the sensor locations are *not* performed.

The IO- and CC-selection methods based on the structured singular value (Sections 3.11 and 4.12) are clarified by some examples as well. In [37] and [35, Section 3.6], locations for 2 temperature measurements in a high-purity distillation column are selected, using design-specific screening tools. A detailed system description can be found in [47, Appendix A]. Only locations symmetric to the feed-tray are considered candidates. Disturbances, input uncertainty, and measurement noise are present. The results for the selection criterion applying the SSV, *i.e.*, based on robust performance, are compared with those resulting from a particular condition number criterion [35, Section 3.4.2]. Closed-loop simulations (response to step disturbances) with some alternative temperature sensor locations are performed under IMC control. Almost the same is done in [35, Appendix B]. In [35, Section 3.5], another distillation column example is studied. The transfer functions between the inputs and the candidate measured variables (temperatures) are available. General screening tools are applied to select the measurements. In [35, Section 4.7], a very brief discussion on measurement selection for a heavy oil fractionator can be found.

For the CSD method discussed in [59, 60], only input selection is illustrated with an example: an input set for a heavy oil fractionator is searched for, using μ -bounds. Unfortunately, the elaboration is very brief.

In [18], CC-selection based on SSV (Section 4.12) is applied to a distillation column, for which a 3×3 transfer function matrix $P(s)$ is given. The ability of a candidate control configuration to guarantee offset-free performance, and the limitations on the achievable bandwidth introduced by

the configuration, are addressed.

Of all tools for CC-selection, the use of RGA and closely related concepts (Section 4.4) is certainly most widespread. In [14], the steady-state RGA is used to select IO-pairings ($l = m = 5$) for a coke oven battery. A linearized state space model of this system (26 states) is provided. In [25], the concepts of frequency dependent PRGA and CLDG are illustrated for the fluid catalytic cracking process. Both a 5-state nonlinear system model and a linearized and reduced 2-state model for this process are given. Also the existence of RHP-zeros for candidate control configurations is addressed. In [23], the same system is considered. Existence of RHP-zeros, frequency dependent RGA and CLDG are used as criteria to assess the two control structures ($l = m = 2$) proposed. Moreover, the sensitivity to changes in operating point and parameter values, and the sensitivity to input uncertainty are studied for these control structures.

The use of frequency dependent PRGA and CLDG is also illustrated in [24, 63], where a 40-tray distillation column example with 2 inputs and 2 outputs is discussed. A reduced order (5 states) linear model is given. In [64], 5 alternative control structures (with different IO-sets) for a distillation column example are studied. The quality of the different control structures is assessed by frequency dependent RGA analysis and performance with "optimal" SISO PI controllers. The PRGA and RGA are also used in [26] and [71] to indicate interactions of the individual loops in distillation column control. In [66], the effect of large RGA elements is illustrated for a simplified distillation column model ($l = m = 2$). Closed-loop responses for 2 different control structures (with different input sets) are performed for diagonal and inverse-based controllers.

In [42], the use of frequency dependent BRG is illustrated for a boiler furnace (the same system is discussed in [54, 56]) and a system of heat-integrated reactors. Both systems are square with 4 inputs and outputs; references to obtain the models are provided. Closed-loop responses to setpoint changes are performed for the most promising control configurations. Unfortunately, the concept of DBRG, as proposed in [2, 3], is not illustrated with an example.

In [10], 2 distillation examples are presented to illustrate the use of the nonsquare steady-state RGA. In both examples ($l = 4, m = 3$ and $l = 9, m = 2$ respectively), this tool is used to square down the system, *i.e.*, to choose a square subsystem from the original nonsquare system in order to obtain a square control system. The results are compared with those from a SVD-based output reduction method. In [55], the nonsquare BRG is applied to a mixing tank with 2 outputs and 3 inputs; the transfer function matrix consists of first order transfer functions. The BRG's associated with the alternative control configurations are calculated, and the most preferable configurations are indicated. Applying IMC, the setpoint responses with the different configurations are compared.

The nonlinear BRG is applied to a chemical process (a model of which is provided) with 2 inputs and outputs [39, 41]. The IO-pairings indicated by the BRG's of the linearized model are compared with those resulting from the nonlinear BRG's. For the alternative configurations, closed-loop simulations to setpoint changes are discussed as well. Unfortunately, the example is not very detailed.

In [4], the concept of relative sensitivity (Section 4.5) is illustrated for a 3×3 system. The high-frequency asymptotes $\bar{\sigma}(\bar{Y}_i)$ associated with two alternative IO-pairings are considered. In [55], the "nonsquare relative sensitivity" is clarified for two alternative control configurations of a system with 2 outputs and 3 inputs.

The CC-selection approach based on the concept of interaction potential (Section 4.7) is illustrated for 3 systems with $l = m = 2$, and 2 systems with $l = m = 3$. The associated transfer function matrices are given; the individual elements represent first order processes with a dead time. For the 2×2 systems the tracking errors with different IO-pairings are compared. One of the examples compares the preferable IO-pairings indicated by the steady-state RGA with those obtained by the interaction potential.

In [43], the CC-selection criterion based on numerical invertibility (Section 4.8) is illustrated by 4 examples (one 2×2 system and three 3×3 systems; the physical interpretation of the examples is not always explained), and the recommended pairings are compared with those recommended by steady-state RGA and SVD. By these examples it is shown, that these pairing criteria not

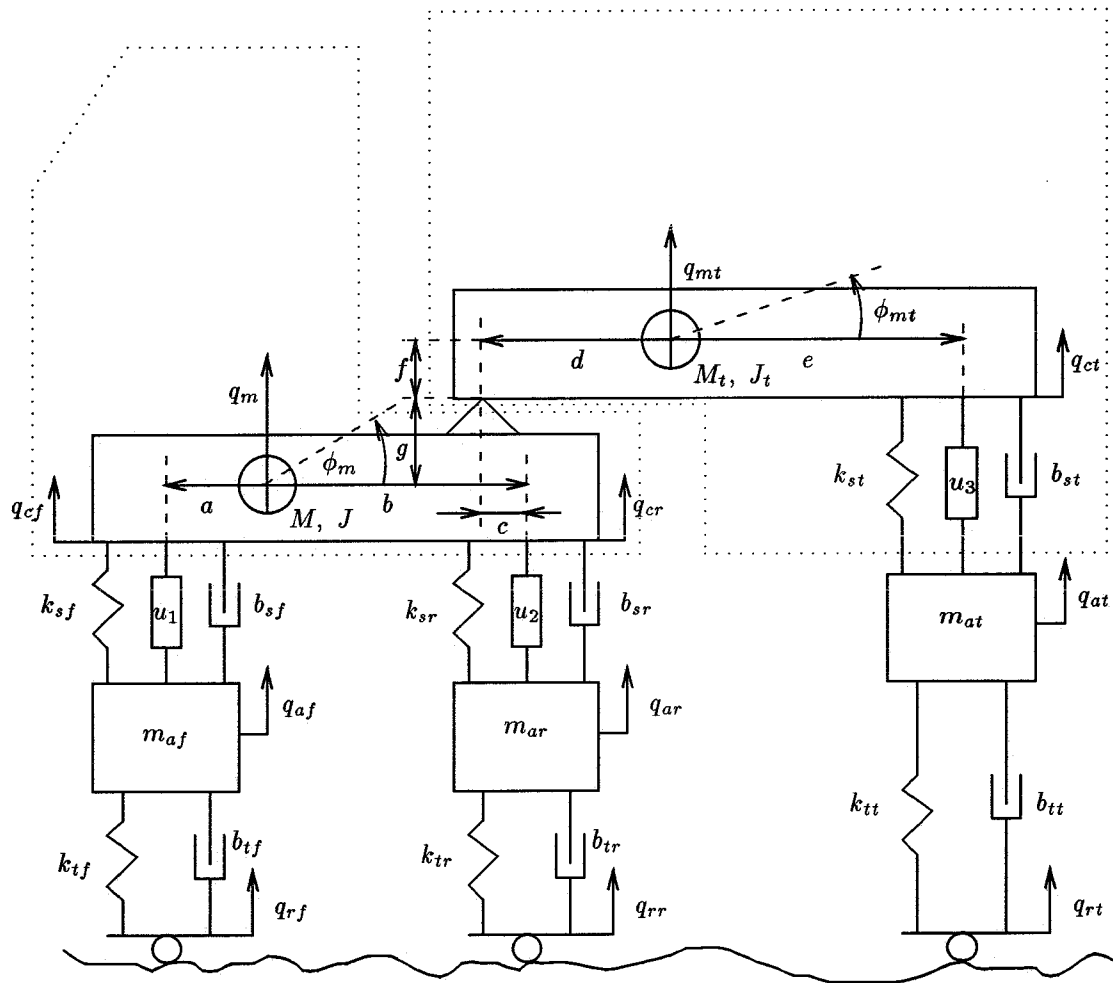


Figure 5.1: Six DOF model of a tractor-semitrailer combination

always yield the same preferable control configuration.

The CC-selection rules based on DCLI discussed in Section 4.10 are illustrated in [11] by a 4×4 transfer function matrix $P(0)$, representing a distillation column at steady state.

5.2 Proposal for a vehicle control example

At the faculty of Mechanical Engineering of the Eindhoven University of Technology, the development of (semi-)active suspension systems for tractor-semitrailer combinations is a major research topic. In Fig. 5.1, a 6 Degrees Of Freedom (DOF) model of the vehicle is depicted.

Three main design goals are distinguished. First, low values of the vertical and pitch accelerations of both the tractor and the semitrailer are important to guarantee good driver comfort, respectively cargo protection. In the second place, due to limited suspension working space, the suspension travel must be limited. Finally, the dynamic tire forces have to be kept low, in order to guarantee good handling and to minimize damaging the road surface.

An active suspension system is expected to perform better than a passive one for given driving

conditions (*e.g.*, speed and road surface). Moreover, contrary to passive suspension systems, the characteristics of (semi-) active suspensions can be adjusted to particular driving conditions, trying to improve the dynamic behaviour of the vehicle. The excitation of the rear wheels of the tractor is believed to be the main cause of the pitch motions of the chassis and the cabin. This is the reason that generally the passive suspension at the rear of the tractor is replaced by an active one.

Eliminating the effects of gravity and assuming the angles ϕ_m and ϕ_{mt} to be small, the 6 DOF model of the tractor-semitrailer combination can be represented by the following linear state space description:

$$\begin{aligned}\dot{x}(t) &= Ax(t) + B_u u(t) + B_w w(t) \\ y(t) &= Cx(t) + D_u u(t) + D_w w(t) \\ z(t) &= Ex(t) + F_u u(t) + F_w w(t)\end{aligned}\quad (5.1)$$

with:

$$\begin{aligned}\text{state variables:} & \quad x^T = [q_{af}, q_{ar}, q_{at}, q_m, \phi_m, \phi_{mt} \\ & \quad \dot{q}_{af}, \dot{q}_{ar}, \dot{q}_{at}, \dot{q}_m, \dot{\phi}_m, \dot{\phi}_{mt}] \\ \text{candidate manipulated variables:} & \quad u^T = [u_f, u_r, u_t] \\ \text{excitation by road surface:} & \quad w^T = [q_{rf}, q_{rr}, q_{rt} \\ & \quad \dot{q}_{rf}, \dot{q}_{rr}, \dot{q}_{rt}] \\ \text{variables to be controlled:} & \quad z^T = [q_{cf} - q_{af}, q_{cr} - q_{ar}, q_{ct} - q_{at} \\ & \quad q_{af} - q_{rf}, q_{ar} - q_{rr}, q_{at} - q_{rt} \\ & \quad \ddot{q}_m, \ddot{\phi}_m, \ddot{q}_{mt}, \ddot{\phi}_{mt}] \\ \text{candidate measured variables:} & \quad y^T = [q_{af} - q_{rf}, q_{ar} - q_{rr}, q_{at} - q_{rt} \\ & \quad q_{cf} - q_{rf}, q_{cr} - q_{rr}, q_{ct} - q_{rt} \\ & \quad q_{cf} - q_{af}, q_{cr} - q_{ar}, q_{ct} - q_{at} \\ & \quad \phi_m, \phi_{mt} \\ & \quad \dot{q}_{cf} - \dot{q}_{af}, \dot{q}_{cr} - \dot{q}_{ar}, \dot{q}_{ct} - \dot{q}_{at} \\ & \quad \ddot{q}_{af}, \ddot{q}_{ar}, \ddot{q}_{at} \\ & \quad \ddot{q}_{cf}, \ddot{q}_{cr}, \ddot{q}_{ct} \\ & \quad \ddot{q}_m, \ddot{\phi}_m, \ddot{q}_{mt}, \ddot{\phi}_{mt}].\end{aligned}$$

All displacements in x are relative to the equilibrium position. The vertical displacement at the front and the rear of the tractor (q_{cf} and q_{cr} respectively) and the vertical displacement at the rear of the semitrailer q_{ct} depend on the state variables in the following way:

$$\begin{aligned}q_{cf} &= q_m - a\phi_m \\ q_{cr} &= q_m + b\phi_m \\ q_{mt} &= q_m + (b - c)\phi_m + d\phi_{mt} \\ q_{ct} &= q_m + (b - c)\phi_m + (d + e)\phi_{mt}\end{aligned}\quad (5.2)$$

The various matrices playing a role in (5.1) can be found in Appendix A.

The first, second and third element in z (variables to be controlled) represent the suspension travels to be limited, while z_4 , z_5 and z_6 account for the dynamic tire forces. The relevant accelerations of both tractor and semitrailer are represented by the elements z_7 - z_{10} .

The column of candidate measurements consists of successively the displacement between the axles and the road surface, the displacement between the chassis/semitrailer and the road surface, the displacement between the chassis/semitrailer and the axles, the angular displacement of the tractor and the semitrailer, the velocities between chassis/semitrailer and the axles, the accelerations of the axles, the accelerations at the front and rear of the tractor and at the rear of the semitrailer, and finally the vertical and angular accelerations of the center of gravity of the tractor and the semitrailer.

As already noted, currently active suspension is applied to the rear wheels of the tractor only (u_r). It seems interesting to check if the correctness of this approach can be confirmed by the IO-selection methods discussed in this report. Moreover, it could be investigated how many and which type of measurements are suggested by the IO-selection methods, and if these measurements are consistent with what is done in practice ($q_{cf} - q_{af}$, $q_{cr} - q_{ar}$, \ddot{q}_{cf} and \ddot{q}_{cr}). CC-selection methods

may indicate if a centralized controller is preferable to a decentralized one. Furthermore, it seems interesting to investigate if it is possible to decouple the design of active suspensions for the tractor and the semitrailer by introducing decentralized control.

As it has been noted in the Introduction, during the stage of CSD it is important to account for, among others, modeling errors and nonlinearities. In the system model discussed here, these effects can easily be introduced. Therefore, this example seems suitable to evaluate CSD methods based on robust performance/stability, or to develop CSD methods for nonlinear systems.

In practice, modeling errors are always present. In the first place, unmodeled dynamics may occur due to neglected sensor and actuator dynamics, or unmodeled resonant modes, *e.g.*, due to flexibility of the chassis. Parameter errors, *e.g.* due to a varying weight of the cargo, or wrongly estimated mass, spring or damper parameters, are also very important. Nonlinearities are introduced, *e.g.*, if the angles ϕ_m and ϕ_{mt} are large, if a tire lifts off, or in case nonlinear damper or spring characteristics are considered. Moreover, in practice Coulomb friction is expected to play a role, which is also a nonlinearity not accounted for in the current model.

Chapter 6

Comparison

In this chapter, the different procedures for IO-selection and CC-selection are compared, based on the thirteen criteria listed in Chapter 1. The most promising methods are suggested and their merits and limitations are further discussed.

In Tables 6.1 and 6.2, respectively the IO-selection methods and CC-selection methods are assessed for their ability to address the desirable aspects discussed in the Introduction. For this purpose, four symbols are used, with the following meaning:

- + : the aspect of interest is *positively* addressed for the particular method
- 0 : the aspect of interest is *not satisfactorily* addressed for the particular method
- : the aspect of interest is *not* or *negatively* addressed for the particular method
- ? : the possibilities for the aspect of interest are *not exactly known* for this method.

In Table 6.1, the IO-selection procedures explained in Chapter 3 are qualified. It is concluded that only two methods explicitly account for robust stability: the method based on the condition number of the plant (Section 3.9) and the one based on the SSV (Section 3.11); the latter approach also accounts for robust performance. This small number is somewhat disappointing, since the issue of robust stability seems to be very important in modern control systems. According to Reeves [54, 56], the systems requiring control are becoming increasingly complex, leading to higher levels of system and environmental uncertainty. Therefore, it seems obvious that robust stability is not only accounted for during controller design, but also in the preceding stage of control system design, *i.e.*, in the CSD phase. Unfortunately, both methods are indirect and restricted to linear system descriptions; nonlinear equivalents have not been encountered in literature.

An important advantage of both methods is their controller independence. Both procedures provide efficient testing conditions for screening a large number of candidates. Furthermore, the conditions for viability of an IO-set can be weakened or strengthened, by which more or fewer candidate IO-sets can be eliminated. This can be done, *e.g.*, by imposing varying quantitative requirements on the desired closed-loop bandwidth and the system's uncertainty to be handled, or by modifying certain weighting functions. Moreover, the theory of both methods is rather complete and well developed, and although the development of the selection criteria requires thorough knowledge of control theory, the application of the tools seems straightforward.

The method based on the SVD is a direct one, *i.e.*, it is possible to directly select desirable measurements and manipulations from the SVD of the plant description that incorporates *all* candidate measurements and manipulations. Most of the other methods are based on the choice of a particular IO-set followed by a feasibility test. In the methods based on economics and accuracy of the state estimates, an optimization algorithm is used, which yields one "optimal" set of measured and manipulated variables; therefore, these methods are also considered direct ones. In the method based on accuracy of state estimates, only direct selection of the measurement *locations* is performed, after specification of the *number* of measurements. Advantages of the SVD method are that it is also applicable to nonsquare plants, and that it is independent of the

section	aspect												
	1:rob.stab.	2:nom.perf.	3:rob.perf.	4:complex.	5:general	6:nonlin.	7:K-indep.	8:direct	9:quant.	10:effic.	11:effect.	12:simple	13:theory
3.1:pow./speed	-	-	-	-	0	0	-	-	-	-	+	+	-
3.2:pole/zero	-	0	-	-	+	?	+	-	-	0	+	+	-
3.3:ctrb/obsvb	-	-	-	+	-	?	-	-	-	+	-	+	+
3.4:digraph	-	-	-	-	+	+	+	-	-	0	-	+	+
3.5:ach.perf.	-	+	-	+	0	?	-	-	+	-	+	-	-
3.6:accur.est.	-	-	-	+	-	-	+	+	+	0	+	-	0
3.7:econom.	-	-	-	-	-	0	-	+	+	-	+	-	+
3.8:MRI	-	-	-	+	-	-	+	-	+	+	-	+	-
3.9:cond.nr.	+	+	-	+	-	-	+	-	+	+	-	+	+
3.10:SVD	-	-	-	+	0	-	+	+	+	+	+	+	+
3.11:SSV	+	+	+	+	-	-	+	-	+	+	+	0	0

Table 6.1: The IO-selection methods compared

controller. Moreover, the theory is rather well developed and many examples in literature have proven its value. The scaling dependence of the SVD method is often mentioned as an important disadvantage.

The method based on the cause-and-effect graphs seems to be not only applicable to linear systems, but to nonlinear systems as well. However, it is expected that the procedure is not powerful enough to be applied on its own, in the sense that in many cases the selection criterion will pass a very large number of candidate IO-sets. In order to reduce the candidate IO-sets to a sufficiently small number, the IO-selection procedure has to be followed by another one, possibly making the *entire* IO-selection procedure inefficient.

The methods based on controllability criteria, locations of poles/zeros and achievable performance under optimal control are known to have nonlinear equivalents. However, if it is possible to translate these into useful tools for IO-selection is unknown, and has to be investigated in future research. Although in [16, 46] the concept of structural controllability is used to propose feasible IO-sets for nonlinear systems, the ability to address the original nonlinear system's controllability by requirements on its structural linearized representation is questionable. So, the practical usefulness of this concept must be investigated during future research. Unfortunately, theory and practical applications of the methods discussed in Sections 3.2 and 3.5 for the purpose of IO-selection, have not been found in literature, even not for linear systems. Moreover, for these methods the same remark applies as for the cause-and-effect graphs, *i.e.*, it is expected that the criteria are not very powerful, in the sense that a large number of candidate IO-sets may remain after screening. For further screening, additional aspects of feasibility must be addressed.

Selection of inputs and outputs based on control power and speed is qualitative and not very systematic. It seems not very useful, particularly not for large-scale systems. The method based on the MRI is theoretically not very well-founded and is only used for selecting manipulated variables.

In Table 6.2, the results for the CC-selection methods discussed in Chapter 4 are displayed. From this table, it is concluded that the method based on the SSV is the only one that explicitly accounts for robust stability. In the methods based on the RGA/BRG, robust stability is less rigorously addressed, since these methods are only provided with "indications" for circumstances under which stability for a particular control configuration may be endangered. For the method based on "performance degradation" (Section 4.9), it is doubtful if robust stability is really achieved for candidate configurations passing the criteria. Although it is stated in [54, 56], that robust stability can be accounted for by specifying the admissible difference between the ideal closed-loop transfer

section	aspect												
	1:rob.stab.	2:nom.perf.	3:rob.perf.	4:complex.	5:general	6:nonlin.	7:K-indep.	8:direct	9:quant.	10:effic.	11:effect.	12:simple	13:theory
4.1:fix.eig.	-	-	-	+	-	?	+	-	-	+	+	+	-
4.2:rel.degr.	-	-	-	+	+	+	+	+	0	+	+	+	-
4.3:ach.perf.	-	+	-	+	-	-	-	-	+	-	+	-	-
4.4:RGA/BRG	0	0	0	+	0	+	+	0	+	+	0	+	+
4.5:rel.sens.	-	0	-	+	0	-	-	-	+	+	+	0	0
4.6:CLDG	-	0	-	+	-	-	+	-	+	+	+	0	-
4.7:interact.pot.	-	+	-	+	0	-	-	0	+	-	+	-	+
4.8:num.invert.	-	-	-	+	-	-	+	-	+	+	?	+	+
4.9:perf.degr.	?	+	-	+	-	-	+	-	+	+	-	+	+
4.10:nom.stab./int.	-	-	-	+	0	-	0	-	+	0	-	+	+
4.11:SVD	-	-	-	+	0	-	+	+	+	+	+	+	-
4.12:SSV	+	+	?	+	0	-	+	-	+	+	+	0	0

Table 6.2: The CC-selection methods compared

function matrix \bar{T} and the real one T by means of a frequency dependent real-valued function d_T , this is not further explained or proven.

Unfortunately, no CC-selection procedure has been encountered which explicitly accounts for robust performance. However, maybe the method based on the SSV offers some possibilities, since in [37] it is stated (not elaborated!), that design-dependent screening tools for IO-selection based on robust performance can also be developed for (block) diagonal control configurations. If this is true, it would be possible to perform the IO-selection phase and CC-selection phase *simultaneously!*

The majority of the CC-selection methods considered are not "generally applicable", *e.g.*, the methods based on performance degradation are restricted to square (sub)systems; the methods based on nominal stability/closed-loop integrity and numerical invertibility are based on steady-state considerations, while the method based on relative sensitivity is developed for high frequencies; the procedures based on closed-loop disturbance gain, interaction potential and numerical invertibility seem to be useful for *diagonal* controllers only.

The CC-selection method applying the concept of relative degree seems to be the most general, although the systems are assumed to be square. Moreover, the method is applicable to *nonlinear* systems. Unfortunately, the method is believed to be useful only for a preliminary quantitative assessment of the alternative control configurations; the relative degree should be combined with other analytical tools, towards a more quantitative CC-selection procedure. In [41, 39], the BRG has been generalized for use in *nonlinear* systems. Unfortunately, the proposed method seems to require high computational effort and is not simple in use.

From Table 6.2, it is concluded that the methods based on relative degree and SVD directly yield one, or some, favorable control configurations: the entries in particular matrices indicate the favorable IO-pairings (this also applies for the "preliminary" IO-pairings in the method based on interaction potential). The method based on the relative gain is only direct if multiloop SISO controllers are considered and the RGA can be used; in case of block diagonal controllers, the individual BRG's have to be recomputed for each candidate control configuration and the method becomes indirect.

Selection of the control configuration based on a measure of performance degradation seems to be the most powerful, in the sense that the number of candidates that is termed viable can easily be affected by putting more or less severe requirements on the allowable performance degradation margin. The same may be achieved with the methods based on the SSV and interaction potential,

by varying specifications on the desired closed-loop system's "performance" \bar{T} . Since the method based on performance degradation uses necessary conditions for feasibility, it may be ineffective. The RGA/BRG method is provided with less efficient and less quantitative CC-selection rules. Therefore, the method is expected to be effective only if a *set* of guidelines for CC-selection (see also Section 4.4) is used to form a combined test for viability of a control configuration. Unfortunately, for some of the pairing rules based on the RGA, the prove is missing [49].

As a final remark, it is emphasized that the comparison of the IO/CC-selection procedures in this chapter is based on what has been found in literature. For an honest comparison this is not sufficient. Instead, a more detailed and more critical analysis of the various methods, in combination with a preliminary implementation using a representative example, must be performed. The tractor-semitrailer problem discussed in Section 5.2 could be used for this purpose.

Chapter 7

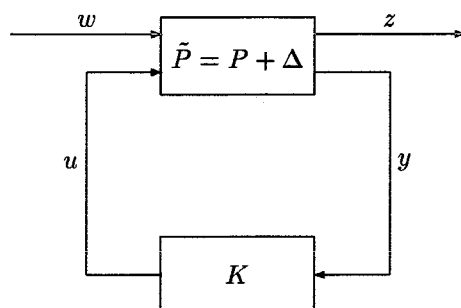
Recommendations for future research

In this chapter, some important issues for further research will be suggested. Firstly, the *goal* of CSD has to be clearly defined, since this lays the foundation for the development of any CSD method. Secondly, some recommendations for improvement of the most promising methods for linear control systems are proposed. In the last part of this section, some important issues a CSD method to be developed must account for, will be discussed.

In [21, 54, 56], the goal of control system design is formulated as follows: *Minimize control system complexity subject to the achievement of accuracy specifications in the face of uncertainty.* That is, robust performance must be achieved, but unnecessary control system complexity cannot be tolerated. One might wonder if this paradigm is correct and complete. Firstly, in most control systems, achieving the accuracy specifications is not the only control objective, *i.e.*, product yield, safety aspects and marginal costs are often important objectives as well, while *internal stability* of the control system is often a prerequisite. Furthermore, uncertainty may not only include modeling errors, measurement noises, and external disturbances, but also *failures* of (a combination of) sensors and actuators. Handling of *constraints* is also very important in practical control systems, *i.e.*, the input/output domain is limited, contrary to what is often assumed in controller design. Constraints may, *e.g.*, be imposed by actuator/sensor saturation, or safety aspects (*e.g.*, maximum pressures in drums, limited working space for mechanical manipulators). Consequently, the goal of control system design and therefore the goal of CSD, could be modified: *Minimize control system complexity subject to the achievement of the control objectives in the face of uncertainties, failures and constraints.*

Consider the representation of the uncertain feedback control system in Fig. 7.1. The *goal* of CSD is thus, to select an appropriate set of measured (y) and manipulated variables (u), that not only minimizes controller complexity, but also guarantees satisfactory control of the performance objectives (z) and internal stability, even in the presence of system (Δ) and environmental uncertainties (in w), actuator/sensor failures, and system constraints.

In this report, different approaches for CSD, particularly for linear control systems, have been discussed and compared. From Table 6.1, the preliminary conclusion is drawn, that with respect to the IO-selection phase the methods employing the condition number of the plant and the one based on the SSV seem the most promising for linear systems. The methods based on the performance degradation criterion under decentralized control, the RGA/BRG methods, and the method based on the SSV seem the most promising for CC-selection in linear systems. Successfully applying the condition number and cross-feed degradation criteria is thus one possible way to deal with CSD for linear control systems; this method is implemented in the MATLAB Control Configuration Design Toolbox [57]. However, this approach has some shortcomings to be solved in future research. Two important disadvantages will be explained:



u : manipulated variables
 w : exogeneous inputs, *e.g.*, measurement noise, reference trajectory
 z : outputs to be controlled
 y : measurements

Figure 7.1: General framework of a perturbed feedback control system

1. In the IO-selection phase, the measured and manipulated variables are selected to maintain stability of $P_y u$ (see Fig. 1.1) in the presence of modeling errors. In [54, 56], it is stated that one attempts to control z by controlling the measured variables y . Since this aspect is not explicitly accounted for during IO-selection (only robust stability is accounted for), it must have been assumed that the variables to be controlled z are always completely represented by the selected measurements y , *whatever* subset of the candidate measurements is selected. It is expected that this is not always possible in practice, since to properly control z , a particular number and type of measurements may always have to be performed. In the next phase, *i.e.*, the CC-selection phase, cross-feed performance degradation with respect to the *measured* variables is investigated, instead of addressing performance degradation in the real control objectives z . This seems contradictory and therefore the issue of *independent* treatment of measurements and variables to be controlled merits further investigation.
2. Although nonsquare systems P may be considered, the *subsystems* under decentralized control are always assumed to be square, *i.e.*, the number of inputs and outputs of one subsystem are the same. The same applies for most of the other approaches discussed. Since this assumption puts severe limitations on the control systems to be considered, extension of the theory to nonsquare subsystems seems to be worthwhile. So, controller matrices, *e.g.*, of the form:

$$K(s) = \begin{bmatrix} \times & 0 \\ \times & 0 \\ 0 & \times \end{bmatrix}, \text{ or } K(s) = \begin{bmatrix} \times & 0 & 0 & 0 & 0 \\ 0 & \times & \times & 0 & 0 \\ 0 & \times & \times & 0 & 0 \\ 0 & 0 & 0 & \times & \times \end{bmatrix}$$

need also be paid attention to. Maybe the nonsquare BRG [55] and the nonsquare RGA [10] are reasonable starting points. In [40], the mathematical development of interaction measures for systems whose inputs and/or outputs belong to more than one subsystem, *e.g.*,

$$K(s) = \begin{bmatrix} \times & 0 \\ \times & \times \\ 0 & \times \end{bmatrix},$$

is proposed as an important research topic.

Another approach for CSD for linear control systems, is one which employs the SSV theory. The method proposed in [35, 36, 37] is only explained for selecting secondary measurements in

inferential control systems. Therefore, a recommendation for future research is to extend the output selection approach to an input-output selection approach based on maintenance of *robust performance*. The SSV-based method discussed in [59] is provided with viability tests on constraint satisfaction, robust stability, DIC, and combinations of these aspects. Unfortunately, the tests are developed under the assumption of integral control action and are only applicable at steady-state. The development of more general, frequency-dependent, IO-selection and CC-selection criteria seems to be worthwhile.

The CC-selection method based on the notions of RGA and BRG may also be promising for use in CSD. However, to obtain a more effective and powerful CC-selection procedure, the selection criteria must be formulated more quantitatively, since unambiguous selection rules such as "avoid pairings with large RGA elements", or "prefer pairings with RGA values close to 1" have to be avoided. Furthermore, for some existing pairing rules, a theoretical justification has to be searched for.

In the last part of this section, a number of issues are proposed, which certainly must be accounted for in a CSD method to be developed. Although some of them have been mentioned briefly in the Introduction, their importance justifies some special attention.

In this report, CSD is split up in two stages, *i.e.*, the IO-selection and CC-selection phase. The same is usually done in literature: no method has been encountered which treats both stages simultaneously. It is expected that this alternative approach would yield a more direct and effective CSD procedure. An important recommendation is thus to examine if it is possible to develop a CSD method solving the IO-selection and CC-selection problems *simultaneously*. This implies that a selection criterion must be searched for which is (efficiently) applicable in both phases, *e.g.*, criteria using the SVD, SSV, or the achievable performance under optimal control, or one applying the notions of digraphs and relative degree. Particularly, it seems worthwhile to investigate the possibilities of a selection criterion based on robust performance, instead of successively performing IO-selection and CC-selection respectively based on robust stability and (nominal) performance requirements, see Sections 3.9 and 4.9.

Furthermore, a *direct* CSD method could be looked for, *i.e.*, a method which yields one, or maybe some, favorable control structures given the control objectives, system uncertainty (and eventually constraints and potential loop failures) and allowable complexity. In a direct approach, identifying all candidate control structures by testing for a specific criterion is not necessary. With the criteria discussed, it is expected that it is not possible to directly find the "optimal" control configuration, using only one criterion. Instead, different criteria must successively be applied, eliminating an increasing number of candidate control structures. Such an indirect method may be very time-consuming, especially for large-scale systems. For an indirect CSD method, development of *efficient* screening tools is therefore an important future research topic. A potential advantage of an indirect approach is, that it yields additional insight in the system to be controlled, which may be useful for the subsequent steps in control system design.

Ideally, a CSD method uses only *open-loop system data*, *i.e.*, information that is independent of the controller. The idea is to eliminate candidates for which a controller achieving the control objectives (in the presence of uncertainty etc.) does not exist, whatever design method is used. A controller-dependent selection criterion is inefficient if prior controller design is required, see, *e.g.*, Sections 3.5 and 4.3, and [3](DBRG). Such a criterion should be used only if the number of candidate control structures has been reduced to a sufficiently low level. On the other hand, a selection criterion which is developed under the assumption of a particular controller *type* may still be efficient, if it is not necessary to actually design the controller for each candidate control structure, see, *e.g.*, Section 4.5 and [37]. For an efficient approach to CSD, it is thus recommendable to develop selection criteria which do not require controller design, *i.e.*, are independent of controller tuning.

Complexity of the control system is mentioned as an essential issue in CSD. In future research, complexity must not only be addressed during CSD, but moreover an unambiguous *definition of complexity* has to be formulated. In [54, Section 5.4], it is stated that at least the following aspects must be addressed by the concept of "complexity":

- the number of selected measured and manipulated variables,
- the number of feedback interconnections between these variables,
- sensor and actuator costs,
- reliability and maintainability, and
- required effort for controller design and tuning.

Another important aspect of complexity is the *transparency* of the control system, and therefore of the control structure: for an operator it should be easily to understand how to influence the "behavior" of the system, by which he can intervene fast and correctly. In [54, Section 5.4], it is suggested that complexity must be addressed after a set of viable control structures has been generated. Minimum complexity might be achieved by constructing an objective function incorporating the aspects associated with complexity listed above, and minimizing this function over all viable control structures.

The main topic for future research is the development of a CSD method which is applicable to *nonlinear* control systems, *e.g.*, mechanical systems with multiplicative and goniometric terms for displacement and velocity, or processes with fractions or exponential terms for temperatures and product concentrations. At this moment, the IO-selection method based on cause-and-effect graphs and the CC-selection method based on relative degree, are the only ones directly applicable to nonlinear systems. However, these methods are expected to be not effective enough to perform CSD "on their own", but they could be used as initial screening tools.

The following procedures discussed in this report are known to have nonlinear equivalents: the IO-selection procedures based on controllability and observability, locations of poles and zeros, achievable performance under optimal control, and the IO-selection procedure based on economics; the CC-selection procedures based on fixed eigenvalues and BRG. Unfortunately, except for the nonlinear BRG, theory and applications for nonlinear CSD are lacking and merit further investigation. From Tables 6.1 and 6.2, it may be concluded that for linear systems, these methods (except for the BRG) suffer from some severe disadvantages, and therefore they are expected to be less promising for nonlinear control systems as well. Regarding the nonlinear BRG, theory is limited and some proves are missing. Moreover, this method requires solving nonlinear differential equations; therefore, the method is expected to be inefficient. Nevertheless, the nonlinear BRG is an interesting concept and certainly deserves thorough investigation.

Another important topic for future research is the development of a CSD method for nonlinear systems, which is provided with testing criteria for robust stability and/or robust performance (as for the methods discussed in Sections 3.9, 3.11 and 4.12). Preferably, for an efficient test on viability, the criterion is expressed in terms of norm bounds ("finite L_p -gains", see, *e.g.*, [52]) on particular desirable system properties (or nonlinear operators). It must be investigated if it is possible to formulate such requirements for nonlinear control systems. In [52], it is shown that for perturbed nonlinear systems in a specific form, a (dynamic) state feedback controller achieving robust stability exists, if the perturbation has finite L_2 -gain and a particular Hamilton-Jacobi equation (nonlinear analog of Riccati equation) has a solution.

Bibliography

- [1] M. Abdel-Mooty and J. Roorda. Optimal control configuration for active control of structural vibrations. In *Proc. of the ASCE Engineering Mechanics Specialty Conf. Mechanics Computing in 1990's and Beyond*, pages 761–765, USA, May 20-22 1991.
- [2] Y. Arkun. The dynamic block relative gain and its role in the design of robust decentralized control. In *Proc. of the 1986 American Control Conf.*, pages 11–16, 1986.
- [3] Y. Arkun. Dynamic block relative gain and its connection with the performance and stability of decentralized control structures. *Int. J. Control*, 46(4):1187–1193, 1987.
- [4] Y. Arkun. Relative sensitivity: a dynamic closed-loop interaction measure and design tool. *AIChE J.*, 34(4):672–675, April 1988.
- [5] L. Bakule and J. Lunze. *Decentralized design of feedback control for large-scale systems*. Institute of Information Theory and Automation of the Czechoslovak Academy of Science, 1988. supplement by *Kybernetika*, 24(3-6).
- [6] B.W. Bequette and T.F. Edgar. The equivalence of non-interacting control system design methods in distillation. In *Proc. of the 1986 American Control Conf.*, pages 31–37, 1986.
- [7] O.H. Bosgra and H. Kwakernaak. *Design methods for control systems*. spring term 1993-1994. lecture notes from the Dutch Graduate Network on Systems and Control.
- [8] S.P. Boyd and C.H. Barratt. *Linear controller design: limits of performance*. Prentice-Hall, Englewood Cliffs, New Jersey, 1991.
- [9] E.H. Bristol. On a new measure of interactions for multivariable process control. *IEEE Trans. Automat. Control*, AC-11(1):133–134, 1966.
- [10] J.-W. Chang and C.-C. Yu. The relative gain for non-square multivariable systems. *Chem. Eng. Sci.*, 45(5):1309–1323, 1990.
- [11] M.S. Chiu and Y. Arkun. Decentralized control structure selection based on integrity considerations. *Ind. Eng. Chem. Res.*, 29(3):369–373, 1990.
- [12] P. Daoutidis and C. Kravaris. Control structure selection: a geometric perspective. In *Proc. of the 1990 American Control Conf.*, pages 2126–2132, San Diego, May 23-25 1990.
- [13] P. Daoutidis, M. Soroush, and C. Kravaris. Feedforward/feedback control of multivariable nonlinear processes. *AIChE J.*, 36(10):1471–1484, October 1990.
- [14] I. Fletcher, I.G. French, and M. Hague. Modelling and control of a coke oven battery. In R. Whalley, editor, *Application of multivariable system techniques*, pages 225–233, University of Bradford, UK, March 1994.
- [15] N. Geerts. Control configuration design. trainee project Report WFW 94.005, Eindhoven University of Technology, Dept. of Mechanical Engineering, 1993. in Dutch.

- [16] A. Georgiou and C.A. Floudas. Structural analysis and synthesis of feasible control schemes - theory and applications. *Chem. Eng. Res. Des.*, 67:600-618, November 1989.
- [17] R. Govind and G.J. Powers. Control system synthesis strategies. *AIChE J.*, 28(1):60-73, January 1982.
- [18] P. Grosdidier and M. Morari. Interaction measures for systems under decentralized control. *Automatica*, 22(3):309-319, 1986.
- [19] G. Hoffmann. Aktive Dämpfung der Gier- und Nickschwingung an einem Roboterarm mit Dehnungs- und Beschleunigungssensoren. *Automatisierungstechnik*, 39:101-104 and 128-134, February and April 1991.
- [20] B.R. Holt and M. Morari. Design of resilient processing plants - vi. The effect of right-half-plane zeros on dynamic resilience. *Chem. Eng. Sci.*, 40:59-74, 1985.
- [21] R.F. Hoskin, C.N. Nett, and D.E. Reeves. Control configuration design for a mixed vector thrust ASTOVL aircraft in hover. 1991.
- [22] M. Hovd. *Studies on control structure selection and design of robust decentralized and SVD controllers*. PhD thesis, University of Trondheim, The Norwegian Institute of Technology, 1992.
- [23] M. Hovd and S. Skogestad. Impact of model uncertainty on control structure selection for the fluid catalytic cracking process. In *Proc. of the IFAC Conf. on Advanced Control of Chemical Processes*, pages 215-220, Toulouse, France, 1991.
- [24] M. Hovd and S. Skogestad. Simple frequency-dependent tools for control system analysis, structure selection and design. *Automatica - J. IFAC*, 28(5):989-996, July 1992.
- [25] M. Hovd and S. Skogestad. Procedure for regulatory control structure selection with application to the FCC process. *AIChE J.*, 39(12):1938-1953, December 1993.
- [26] M. Hovd and S. Skogestad. Robust sequential design of decentralized controllers. In *Proc. of the 2nd European Control Conference*, volume 2, pages 1050-1055, Groningen, The Netherlands, June 1993.
- [27] H.-P. Huang, M. Ohshima, and I. Hashimoto. Dynamic interaction and multiloop control system design. *J. Proc. Cont.*, 4(1):15-27, 1994.
- [28] A.G. de Jager. The design of decentralized control systems. Report B-286, Delft University of Technology, Laboratory for Measurement and Control, August 1983. in Dutch.
- [29] A.G. de Jager. Structural aspects of decentralized control. Report A-286, Delft University of Technology, Laboratory for Measurement and Control, March and April 1983. in Dutch.
- [30] R.D. Johnston and G.W. Barton. Structural equivalence and model reduction. *Int. J. Control*, 41(6):1477-1491, 1985.
- [31] J.P. Keller and D. Bonvin. Selection of inputs for the purpose of model reduction and controller design. In *Proc. of the IFAC 10th Triennial World Congress*, pages 209-214, Munich, 1987.
- [32] S. Kumar and J.H. Seinfeld. Optimal location of measurements in tubular reactors. *Chem. Eng. Sci.*, 33:1507-1516, 1978.
- [33] H. Kwakernaak and R. Sivan. *Linear optimal control systems*. Wiley-Interscience, New York, 1972.

- [34] H. Lau, J. Alvarez, and K.F. Jensen. Synthesis of control structures by singular value analysis: dynamic measures of sensitivity and interaction. *AIChE J.*, 31(3):427-439, March 1985.
- [35] J.H. Lee. *Robust inferential control: a methodology for control structure selection and inferential control system design in the presence of model/plant mismatch*. PhD thesis, California Institute of Technology, 1991.
- [36] J.H. Lee and M. Morari. Robust control structure selection and control system design methods applied to distillation column control. In *Proc. of the 29th Conf. on Decision and Control*, pages 2041-2046, Honolulu, Hawaii, December 1990.
- [37] J.H. Lee and M. Morari. Robust measurement selection. *Automatica*, 27(3):519-527, 1991.
- [38] J.M. Maciejowski. *Multivariable feedback design*. Addison-Wesley, Amsterdam, 1989.
- [39] M. Manousioutakis, V. and Nikolaou. Analysis of decentralized control structures for nonlinear systems. *AIChE J.*, 35(4):549-558, April 1989.
- [40] V. Manousiouthakis. Decentralized control systems for chemical plants. Issues and approaches. In *Proc. of the IFAC 10th Triennial World Congress*, pages 205-207, Munich, 1987.
- [41] V. Manousiouthakis and M. Nikolaou. On the nonlinear block relative gain. In *Proc. of the 1987 American Control Conference*, pages 2120-2125, 1987.
- [42] V. Manousiouthakis, R. Savage, and Y. Arkun. Synthesis of decentralized process control structures using the concept of block relative gain. *AIChE J.*, 32(6):991-1003, June 1986.
- [43] G. Mijares, J.D. Cole, N.W. Naugle, H.A. Preisig, and C.D. Holland. A new criterion for the pairing of control and manipulated variables. *AIChE J.*, 32(9):1439-1449, 1986.
- [44] C. Moore, J. Hackney, and Carter D. Selecting sensor location and type for multivariable processes. In *Proc. of the Shell Process Control Workshop*, pages 291-308, Houston, Texas, December 15-18 1987.
- [45] M. Morari. Design of resilient processing plants - iii. *Chem. Eng. Sci.*, 38(11):1881-1891, 1983.
- [46] M. Morari and G. Stephanopoulos. Studies in the synthesis of control structures for chemical processes: part ii. *AIChE J.*, 26(2):232-246, March 1980.
- [47] M. Morari and E. Zafiriou. *Robust process control*. Prentice-Hall, Englewood Cliffs, NJ, 1989.
- [48] L.T. Narraway and J.D. Perkins. Selection of process control structure based on linear dynamic economics. *Ind. Eng. Chem. Res.*, 32(11):2681-2692, 1993.
- [49] C.N. Nett and V. Manousiouthakis. Euclidean condition and block relative gain: connections, conjectures, and clarifications. *IEEE Trans. on Automat. Control*, AC-32(5):405-407, April 1987.
- [50] H. Nijmeijer. Principles of optimal control theory. memorandum 908, University of Twente, Faculty of Applied Mathematics, November 1990.
- [51] H. Nijmeijer and A.J. van der Schaft. *Nonlinear dynamical control systems*. Springer-Verlag, New York, 1990.
- [52] H. Nijmeijer and A.J. van der Schaft. *Stabilization and optimal control of nonlinear systems*. spring term 1994. lecture notes from the Dutch Graduate Network on Systems and Control.
- [53] N. Nishida, G. Stephanopoulos, and A.W. Westerberg. A review of process synthesis. *AIChE J.*, 27(3):321-351, May 1981.

- [54] D.E. Reeves. *A comprehensive approach to control configuration design for complex systems*. PhD thesis, Georgia Institute of Technology, 1991.
- [55] D.E. Reeves and Y. Arkun. Interaction measures for nonsquare decentralized control structures. *AIChE J.*, 35(4):603–613, 1989.
- [56] D.E. Reeves, C.N. Nett, and Y. Arkun. Control configuration design for complex systems: a practical theory. submitted to *IEEE Trans. on Automat. Control*, 1991.
- [57] D.E. Reeves, C.N. Nett, and Y. Arkun. *Control configuration design toolbox for use with MATLAB*, May 1991. user's guide, version 1.0.
- [58] J.E. Rijnsdorp. *Integrated process control and automation*. Elsevier, Amsterdam, 1991.
- [59] D.E. Rivera. A structured singular value test for variable selection and pairing. In *Proc. of the 1989 American Control Conf.*, pages 560–565, Pittsburgh, 1989.
- [60] D.E. Rivera, A. Desai, T. Beaumariage, and C. Roberts. A knowledge-based framework for controller structure selection in intelligent process control. In *Proc. of the 1993 American Control Conf.*, pages 1890–1894, San Francisco, June 1993.
- [61] L.W. Russell and J.D. Perkins. Towards a method for diagnosis of controllability and operability problems in chemical plants. *Chem. Eng. Res. Des.*, 65:453–461, November 1987.
- [62] R.A. Savage. Block relative gain and its application to the synthesis of control structures for chemical plants. Master's thesis, Rensselaer Polytechnic Institute, 1984.
- [63] S. Skogestad and M. Hovd. Use of frequency-dependent RGA for control structure selection. In *Proc. of the 1990 American Control Conf.*, pages 2133–2139, San Diego, CA, 1990.
- [64] S. Skogestad, P. Lundström, and E.W. Jacobsen. Selecting the best distillation control configuration. *AIChE J.*, 36(5):753–764, May 1990.
- [65] S. Skogestad and M. Morari. Effect of disturbance directions on closed-loop performance. *Ind. Eng. Chem. Res.*, 26(10):2029–2035, 1987.
- [66] S. Skogestad and M. Morari. Implications of large RGA elements on control performance. *Ind. Eng. Chem. Res.*, 26(11):2323–2330, 1987.
- [67] J.-J.E. Slotine and W. Li. *Applied nonlinear control*. Prentice-Hall, NJ, 1991.
- [68] G. Stephanopoulos. *Chemical process control. An introduction to theory and practice*. Prentice-Hall, Englewood Cliffs, NJ, 1984.
- [69] G. Strang. *Linear algebra and its applications*. Harcourt Brace Jovanovich, San Diego, third edition, 1988.
- [70] H. Tamura and T. Yoshikawa. *Large scale systems control and decision making*. Marcel Dekker Inc., New York, 1990.
- [71] E.A. Wolff and S. Skogestad. Control configuration selection for distillation columns under temperature control. In *Proc. of the 2nd European Control Conf.*, volume 2, pages 637–642, Groningen, The Netherlands, June 1993.
- [72] D.R. Yang, D.E. Seborg, and D.A. Mellichamp. Combined balance control structure selection for distillation columns. *Ind. Eng. Chem. Res.*, 30(9):2159–2168, 1991.
- [73] C.-C. Yu and W.L. Luyben. Design of multiloop SISO controllers in multivariable processes. *Ind. Eng. Chem. Process. Des. Dev.*, 25(2):498–503, 1986.

Appendix A

Tractor-semitrailer model

In this appendix, the system matrices in the state space description (5.1) of the tractor-semitrailer model are given in detail.

The equations of motion of the tractor-semitrailer can be written as:

$$M\ddot{q} + B\dot{q} + Kq = u^* + w^* \quad (\text{A.1})$$

with:

$$q = \begin{bmatrix} q_{af} \\ q_{ar} \\ q_{at} \\ q_m \\ \phi_m \\ \phi_{mt} \end{bmatrix}$$

$$u^* = \begin{bmatrix} u_f \\ u_r \\ u_t \\ -u_f - u_r - u_t \\ au_f - bu_r - (b-c)u_t \\ -(d+e)u_t \end{bmatrix} = B_u^* u; \quad B_u^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \\ a & -b & -(b-c) \\ 0 & 0 & -(d+e) \end{bmatrix}$$

$$w^* = \begin{bmatrix} k_{tf}q_{rf} + b_{tf}\dot{q}_{rf} \\ k_{tr}q_{rr} + b_{tr}\dot{q}_{rr} \\ k_{tt}q_{rt} + b_{tt}\dot{q}_{rt} \\ 0 \\ 0 \\ 0 \end{bmatrix} = B_w^* w; \quad B_w^* = \begin{bmatrix} k_{tf} & 0 & 0 & b_{tf} & 0 & 0 \\ 0 & k_{tr} & 0 & 0 & b_{tr} & 0 \\ 0 & 0 & k_{tt} & 0 & 0 & b_{tt} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and u and w as defined in Section 5.2. Furthermore:

$$M = \begin{bmatrix} m_{af} & 0 & 0 & 0 & 0 & 0 \\ 0 & m_{ar} & 0 & 0 & 0 & 0 \\ 0 & 0 & m_{at} & 0 & 0 & 0 \\ 0 & 0 & 0 & (M + M_t) & M_t(b-c) & M_t d \\ 0 & 0 & 0 & M_t(b-c) & J + M_t(b-c)^2 + M_t g^2 & M_t d(b-c) + M_t fg \\ 0 & 0 & 0 & M_t d & M_t d(b-c) + M_t fg & J_t + M_t d^2 + M_t f^2 \end{bmatrix}$$

$$B = \begin{bmatrix} b_{tf} + b_{sf} & 0 & 0 & -b_{sf} & ab_{sf} & 0 \\ 0 & b_{tr} + b_{sr} & 0 & -b_{sr} & -bb_{sr} & 0 \\ 0 & 0 & b_{tt} + b_{st} & -b_{st} & -(b-c)b_{st} & -(d+e)b_{st} \\ -b_{sf} & -b_{sr} & -b_{st} & B_{44} & B_{45} & B_{46} \\ ab_{sf} & -bb_{sr} & -(b-c)b_{st} & B_{54} & B_{55} & B_{56} \\ 0 & 0 & -(d+e)b_{st} & B_{64} & B_{65} & B_{66} \end{bmatrix}$$

in which:

$$\begin{aligned}
B_{44} &= b_{sf} + b_{sr} + b_{st} \\
B_{45} &= -ab_{sf} + bb_{sr} + (b-c)b_{st} \\
B_{46} &= (d+e)b_{st} \\
B_{54} &= B_{45} \\
B_{55} &= a^2b_{sf} + b^2b_{sr} + (b-c)^2b_{st} \\
B_{56} &= (b-c)(d+e)b_{st} \\
B_{64} &= B_{46} \\
B_{65} &= B_{56} \\
B_{66} &= (d+e)^2b_{st}
\end{aligned}$$

$$K = \begin{bmatrix} k_{tf} + k_{sf} & 0 & 0 & -k_{sf} & ak_{sf} & 0 \\ 0 & k_{tr} + k_{sr} & 0 & -k_{sr} & -bk_{sr} & 0 \\ 0 & 0 & k_{tt} + k_{st} & -k_{st} & -(b-c)k_{st} & -(d+e)k_{st} \\ -k_{sf} & -k_{sr} & -k_{st} & K_{44} & K_{45} & K_{46} \\ ak_{sft} & -bk_{sr} & -(b-c)k_{st} & K_{54} & K_{55} & K_{56} \\ 0 & 0 & -(d+e)k_{st} & K_{64} & K_{65} & K_{66} \end{bmatrix}$$

in which:

$$\begin{aligned}
K_{44} &= k_{sf} + k_{sr} + k_{st} \\
K_{45} &= -ak_{sf} + bk_{sr} + (b-c)k_{st} \\
K_{46} &= (d+e)k_{st} \\
K_{54} &= K_{45} \\
K_{55} &= a^2k_{sf} + b^2k_{sr} + (b-c)^2k_{st} \\
K_{56} &= (b-c)(d+e)k_{st} \\
K_{64} &= K_{46} \\
K_{65} &= K_{56} \\
K_{66} &= (d+e)^2k_{st}.
\end{aligned}$$

Choosing $x^T = [q^T \dot{q}^T]$, (A.1) can be written in state space form:

$$\dot{x} = Ax + B_u u + B_w w \quad (\text{A.2})$$

with:

$$A = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}B \end{bmatrix}; \quad B_u = \begin{bmatrix} 0 \\ M^{-1} \end{bmatrix} B_u^*; \quad B_w = \begin{bmatrix} 0 \\ M^{-1} \end{bmatrix} B_w^*$$

The variables to be controlled z are expressed in terms of x , u and w by the relation:

$$z = Ex + F_u u + F_w w \quad (\text{A.3})$$

with:

$$E = \begin{bmatrix} -1 & 0 & 0 & 1 & -a & 0 & 0 & \cdots & 0 \\ 0 & -1 & 0 & 1 & b & 0 & 0 & \cdots & 0 \\ 0 & 0 & -1 & 1 & b-c & d+e & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ A_{10,1} & A_{10,2} & A_{10,3} & A_{10,4} & A_{10,5} & A_{10,6} & A_{10,7} & \cdots & A_{10,12} \\ A_{11,1} & A_{11,2} & A_{11,3} & A_{11,4} & A_{11,5} & A_{11,6} & A_{11,7} & \cdots & A_{11,12} \\ E_{9,1} & E_{9,2} & E_{9,3} & E_{9,4} & E_{9,5} & E_{9,6} & E_{9,7} & \cdots & E_{11,12} \\ A_{12,1} & A_{12,2} & A_{12,3} & A_{12,4} & A_{12,5} & A_{12,6} & A_{12,7} & \cdots & A_{12,12} \end{bmatrix}$$

where:

$$E_{9,j} = A_{10,j} + (b-c)A_{11,j} + dA_{12,j}$$

$$F_u = \begin{bmatrix} 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \\ B_{u10,1} & B_{u10,2} & B_{u10,3} \\ B_{u11,1} & B_{u11,2} & B_{u11,3} \\ F_{u9,1} & F_{u9,2} & F_{u9,3} \\ B_{u12,1} & B_{u12,2} & B_{u12,3} \end{bmatrix}$$

where:

$$F_{u9,j} = B_{u10,j} + (b - c)B_{u11,j} + dB_{u12,j}$$

$$F_w = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ B_{w10,1} & B_{w10,2} & B_{w10,3} & B_{w10,4} & B_{w10,5} & B_{w10,6} \\ B_{w11,1} & B_{w11,2} & B_{w11,3} & B_{w11,4} & B_{w11,5} & B_{w11,6} \\ F_{w9,1} & F_{w9,2} & F_{w9,3} & F_{w9,4} & F_{w9,5} & F_{w9,6} \\ B_{w12,1} & B_{w12,2} & B_{w12,3} & B_{w12,4} & B_{w12,5} & B_{w12,6} \end{bmatrix}$$

where:

$$F_{w9,j} = B_{w10,j} + (b - c)B_{w11,j} + dB_{w12,j}$$

Here, $A_{i,j}$ denotes the element of A on the i -th row, in the j -th column.
For the candidate measured variables, a similar relation can be written:

$$y = Cx + D_u u + D_w w \quad (\text{A.4})$$

with:

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -a & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & b & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & b - c & d + e & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & -a & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & b & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & b - c & d + e & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & -a & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & b & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & b - c & d + e \\ A_{7,1} & A_{7,2} & A_{7,3} & A_{7,4} & A_{7,5} & A_{7,6} & A_{7,7} & A_{7,8} & A_{7,9} & A_{7,10} & A_{7,11} & A_{7,12} \\ A_{8,1} & A_{8,2} & A_{8,3} & A_{8,4} & A_{8,5} & A_{8,6} & A_{8,7} & A_{8,8} & A_{8,9} & A_{8,10} & A_{8,11} & A_{8,12} \\ A_{9,1} & A_{9,2} & A_{9,3} & A_{9,4} & A_{9,5} & A_{9,6} & A_{9,7} & A_{9,8} & A_{9,9} & A_{9,10} & A_{9,11} & A_{9,12} \\ C_{18,1} & C_{18,2} & C_{18,3} & C_{18,4} & C_{18,5} & C_{18,6} & C_{18,7} & C_{18,8} & C_{18,9} & C_{18,10} & C_{18,11} & C_{18,12} \\ C_{19,1} & C_{19,2} & C_{19,3} & C_{19,4} & C_{19,5} & C_{19,6} & C_{19,7} & C_{19,8} & C_{19,9} & C_{19,10} & C_{19,11} & C_{19,12} \\ C_{20,1} & C_{20,2} & C_{20,3} & C_{20,4} & C_{20,5} & C_{20,6} & C_{20,7} & C_{20,8} & C_{20,9} & C_{20,10} & C_{20,11} & C_{20,12} \\ A_{10,1} & A_{10,2} & A_{10,3} & A_{10,4} & A_{10,5} & A_{10,6} & A_{10,7} & A_{10,8} & A_{10,9} & A_{10,10} & A_{10,11} & A_{10,12} \\ A_{11,1} & A_{11,2} & A_{11,3} & A_{11,4} & A_{11,5} & A_{11,6} & A_{11,7} & A_{11,8} & A_{11,9} & A_{11,10} & A_{11,11} & A_{11,12} \\ C_{23,1} & C_{23,2} & C_{23,3} & C_{23,4} & C_{23,5} & C_{23,6} & C_{23,7} & C_{23,8} & C_{23,9} & C_{23,10} & C_{23,11} & C_{23,12} \\ A_{12,1} & A_{12,2} & A_{12,3} & A_{12,4} & A_{12,5} & A_{12,6} & A_{12,7} & A_{12,8} & A_{12,9} & A_{12,10} & A_{12,11} & A_{12,12} \end{bmatrix}$$

where:

$$\begin{aligned} C_{18,j} &= A_{10,j} - aA_{11,j} \\ C_{19,j} &= A_{10,j} + bA_{11,j} \\ C_{20,j} &= A_{10,j} + (b-c)A_{11,j} + (d+e)A_{12,j} \\ C_{23,j} &= A_{10,j} + (b-c)A_{11,j} + dA_{12,j} \end{aligned}$$

$$D_u = \begin{bmatrix} 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \\ B_{u7,1} & B_{u7,2} & B_{u7,3} \\ B_{u8,1} & B_{u8,2} & B_{u8,3} \\ B_{u9,1} & B_{u9,2} & B_{u9,3} \\ D_{u18,1} & D_{u18,2} & D_{u18,3} \\ D_{u19,1} & D_{u19,2} & D_{u19,3} \\ D_{u20,1} & D_{u20,2} & D_{u20,3} \\ B_{u10,1} & B_{u10,2} & B_{u10,3} \\ B_{u11,1} & B_{u11,2} & B_{u11,3} \\ D_{u23,1} & D_{u23,2} & D_{u23,3} \\ B_{u12,1} & B_{u12,2} & B_{u12,3} \end{bmatrix}$$

where:

$$\begin{aligned} D_{u18,j} &= B_{u10,j} - aB_{u11,j} \\ D_{u19,j} &= B_{u10,j} + bB_{u11,j} \\ D_{u20,j} &= B_{u10,j} + (b-c)B_{u11,j} + (d+e)B_{u12,j} \\ D_{u23,j} &= B_{u10,j} + (b-c)B_{u11,j} + dB_{u12,j} \end{aligned}$$

$$D_w = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 \\ B_{w7,1} & B_{w7,2} & B_{w7,3} & B_{w7,4} & B_{w7,5} & B_{w7,6} \\ B_{w8,1} & B_{w8,2} & B_{w8,3} & B_{w8,4} & B_{w8,5} & B_{w8,6} \\ B_{w9,1} & B_{w9,2} & B_{w9,3} & B_{w9,4} & B_{w9,5} & B_{w9,6} \\ D_{w18,1} & D_{w18,2} & D_{w18,3} & D_{w18,4} & D_{w18,5} & D_{w18,6} \\ D_{w19,1} & D_{w19,2} & D_{w19,3} & D_{w19,4} & D_{w19,5} & D_{w19,6} \\ D_{w20,1} & D_{w20,2} & D_{w20,3} & D_{w20,4} & D_{w20,5} & D_{w20,6} \\ B_{w10,1} & B_{w10,2} & B_{w10,3} & B_{w10,4} & B_{w10,5} & B_{w10,6} \\ B_{w11,1} & B_{w11,2} & B_{w11,3} & B_{w11,4} & B_{w11,5} & B_{w11,6} \\ D_{w23,1} & D_{w23,2} & D_{w23,3} & D_{w23,4} & D_{w23,5} & D_{w23,6} \\ B_{w12,1} & B_{w12,2} & B_{w12,3} & B_{w12,4} & B_{w12,5} & B_{w12,6} \end{bmatrix}$$

where:

$$\begin{aligned} D_{w18,j} &= B_{w10,j} - aB_{w11,j} \\ D_{w19,j} &= B_{w10,j} + bB_{w11,j} \\ D_{w20,j} &= B_{w10,j} + (b-c)B_{w11,j} + (d+e)B_{w12,j} \\ D_{w23,j} &= B_{w10,j} + (b-c)B_{w11,j} + dB_{w12,j} \end{aligned}$$