REACTION AND FLOW VARIANTS/INVARIANTS FOR THE ANALYSIS OF CHEMICAL REACTION DATA

THÈSE N^{O} 1861 (1998)

PRÉSENTÉE AU DÉPARTEMENT DE GÉNIE MÉCANIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES TECHNIQUES

PAR

Michael AMRHEIN

Dipl.-Ing. Verfahrenstechnik, Universität Stuttgart, Allemagne
de nationalité allemande

acceptée sur proposition du jury:

Prof. D. Bonvin, directeur de thèse
Prof. G. Bastin, rapporteur
Prof. S. Morgenthaler, rapporteur
Prof. A. Renken, rapporteur
Dr M. Schumacher, rapporteur
Dr B. Srinivasan, rapporteur

Lausanne, EPFL
1998

Contents

in Liebe an Gabriela und Thomas Yann

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Angestellter am Institut d'Automatique (IA) der École Polytechnique Fédérale in Lausanne in Zusammenarbeit mit der Lonza AG, Visp, Wallis. Sie wurde von der schweizerischen Kommission für Technologie und Innovation (Projekt 2892.1) gefördert, der ich in diesem Zusammenhang meinen Dank ausspreche.

Außerdem möchte ich allen meinen Freunden und Bekannten danken, die in irgendeiner Weise, sei es mit guten Worten oder mit wissenschaftlichen Ratschlägen, zum Entstehen dieser Arbeit beitragen haben.

Den Professoren D. Bonvin und R. Longchamp danke ich für die Ermöglichung dieser Arbeit an ihrem Institut. Besonders Herrn Prof. D. Bonvin, meinem Doktorvater, bin ich zutiefst zum Dank verpflichtet für sein ungebrochenes Interesse an dieser Arbeit, für die zahlreichen, langen, inspirierenden Diskussionen und für die Endkorrektur des vorliegenden Berichtes.

Den Professoren G. Bastin, S. Morgenthaler und A. Renken und den Doktoren M. M. Schumacher und B. Srinivasan gilt mein Dank für die freundliche Übernahme des Koreferates und Herrn Prof. R. Longchamp für den Prüfungsvorsitz.

Srinis Engagement, seine geniale Intuition und sein "sektionistisches" Denken waren essentiell in dieser Arbeit, wofür ich ihm meinen tiefsten Dank ausdrücken möchte. Mein Dank gilt ihm auch für die Freundschaft, die uns während all dieser Jahre verbunden hat.

Für die zahlreichen Ratschläge industrieller und experimenteller Natur möchte ich meinem Freund und Kollegen, Dr. Martin Schumacher von der Lonza AG, danken. Weiterhin gilt mein Dank Stefan Lauwiner von der Lonza AG, Prof. J. Besse und dem in den Ferien tödlich verunglückten Dominique Bamat von der Ingenieurschule Sitten, Wallis, für ihre Hilfe bei den experimentellen Arbeiten. Auch Prof. S. Morgenthaler vom Lehrstuhl für Angewandte Statistik der EPFL schulde ich Dank für die freundschaftliche Zusammenarbeit.

Allen meinen Kollegen möchte ich ein herzliches Dankeschön aussprechen für die hervorragende und kameradschaftliche Stimmung am Institut. Unvergeßlich bleiben die "Orgies" und Ausflüge IN und EX. Mein besonderer Dank gilt in diesem Zusammenhang Marie-Claire, unserer Sekretärin und "Mutter" aller, meinen Bürokollegen Alejandro und Sandra, sowie Denis, Yves, Christophe + Christophe und Sergio für ihre Hilfe in Soft- und Hardwareangelegenheiten, als auch Christina für die zahlreichen Diskussionen verfahrenstechnischer Art. Meinem Freund Thomas möchte ich an dieser Stelle für sein Wirken als Koch außerhalb der Kantinenzeiten danken.

Meinen Eltern und meiner Schwester zolle ich Dank für ihren liebevollen Beistand. Schließlich und endlich wollte ich meinen zwei Liebsten, meiner Frau Gabriela und meinem Sohn Thomas Yann, danken für ihr geduldiges Einverständnis für meine langen Nächte und Wochenenden am Institut. Gabrielas Liebe und moralische Unterstützung waren unerläßlich.

Abstract

This dissertation is concerned with the development of a methodology and appropriate tools for the investigation of chemical reaction systems using measured data. More specifically, the determination of reaction stoichiometry and kinetics from concentration or, preferably, spectral measurements is considered.

The main contribution of this work is the derivation of a nonlinear transformation of the dynamic model that enables the separation of the evolution of the states into three parts: (i) the reaction-variant part (related to the reactions), (ii) the reaction-invariant and flow-variant part (related to the inlet and outlet streams), and (iii) the reaction- and flow-invariant part (related to the initial conditions). This transformation is very helpful in the analysis of concentration and spectral data.

Dynamic model

First-principles models of reaction systems are gaining importance in chemical and biotechnological production. They can considerably reduce process development costs and be used for simulation, model-based monitoring, control, and optimization, thus leading to improved product quality, productivity, and process safety. These models include information regarding both the chemical reactions and the operational mode of the reaction system.

For the analysis of these models, it is important to distinguish between the states that depend on the reactions and those which do not. The concept of reaction invariants is extended to include the flow invariants of reaction systems with inlet and outlet streams. A nonlinear transformation of the first-principles dynamic model to normal form is proposed. Model reduction, state accessibility, and feedback linearizability are analyzed in the light of this transformation.

Concentration data

Concentration data collected from reaction systems are highly structured, a result of the underlying reactions and the presence of material exchange terms. It is shown that concentration data can be analyzed in the framework of the three-level decomposition provided by the transformation to normal form. The resulting factorization, termed *the factorization of concentration data*, enables (i) the separation of the reaction and flow variants/invariants and (ii) the segregation of the dynamics (extents of reaction, integral of flows) from the static information (stoichiometry, initial and inlet concentrations).

Using the factorization of concentration data, it is possible to isolate the reaction variant part by subtracting the reaction-invariant part from measured concentrations. The reaction-variant part is often unknown, since it depends on the kinetic description (typically the main difficulty in modeling chemical reaction systems). In contrast, the reaction-invariant part is usually known or measured. It is shown that, in cases where the reaction variants can be computed from the concentrations of a few measured species, the concentrations of the remaining species can be reconstructed using the

known reaction-invariant part.

Target factor analysis has been used successfully with concentration data to determine, *without knowledge of reaction kinetics*, the number of reactions and the corresponding stoichiometries. It is shown that, when only the reaction-variant part of the data is considered, existing target factor-analytical techniques can be readily applied. However, if target factor analysis needs to be applied directly to measured concentrations, knowledge of reaction-invariant relationships is required to specify necessary *and* sufficient conditions for the acceptance of stoichiometric targets.

Spectral data

In current practice, concentration measurements during the course of a reaction are generally not available, neither on-line nor off-line. Owing to new measurement technologies, spectral measurements are now available in both the laboratory and production. Various spectral instruments enable non-destructive indirect concentration measurement of most of the species *in-situ*/on-line during the course of a reaction. Measurements are available at high sampling rates and delay-free at low costs. Furthermore, in most cases, the spectral data are linear, i.e., the mixture spectrum is a linear combination of the pure-component spectra weighted by the concentrations. It is shown that the three-level interpretation provided by the transformation to normal form is applicable to spectral data from reacting mixtures.

Similarly to traditional wet-chemical analysis methods, a calibration model must also be estimated that provides concentration estimates from spectral measurements. All calibration methods require that a new spectrum lies in the space spanned by the calibration spectral data (*space-inclusion condition*). To verify this space-inclusion condition, it is proposed to build a calibration model for the reaction-variant part only. Once the reaction variants are predicted from a new spectrum, the (known) reaction invariants can be added to reconstruct the concentrations.

Concentration measurements for some species of interest are often not available due to difficulties/costs in sampling, sample preparation, and development of analytical techniques. Thus, traditional calibration of spectral measurements for the purpose of concentration estimation is not possible. Instead, explicit or implicit knowledge about the kinetic structure will be used (prior knowledge about the reaction-variant part), thus enabling the formulation of factor-analytical methods as a calibration problem.

For pedagogical reasons, the results are developed for isothermal, constant-density reaction systems with inlet and outlet streams. The results are then extended to various scenarios such as reaction systems with varying density and temperature. Furthermore, factorizations of concentration data are presented that include temperature or calorimetric measurements. Several special cases are considered, encompassing continuous stirred-tank reaction systems, semibatch and batch reaction systems, systems with reactions in quasi-equilibrium conditions, and non-reacting mixtures with closure.

Version abrégée

Cette dissertation traite du développement d'une méthodologie et d'outils appropriés pour l'étude de systèmes réactionnels chimiques sur la base de mesures experimentales. Plus particulièrement, on considère la détermination de la stœchiométrie et de la cinétique sur la base de mesures de concentration ou, de préférence, de mesures spectrales.

La contribution principale de ce travail réside dans la dérivation d'une transformation non linéaire du modèle dynamique qui permet de séparer l'évolution des états en trois parties: (i) la partie liée aux variants réactionnels (réactions), (ii) la partie liée aux invariants réactionnels *et* variants de flux (entrées et sorties), et (iii) la partie liée aux invariants réactionnels et de flux (conditions initiales). La transformation s'avère très utile par l'analyse de mesures de concentration et spectrales.

Modèle dynamique

Les modèles de connaissance de systèmes réactionnels gagnent de l'importance dans le développement et la production chimiques et biotechnologiques. Ils sont très utiles pour la simulation, la supervision, la commande, et l'optimisation, qui mènent tous à une amélioration de la qualité du produit, de la productivité et de la sécurité du processus. Ces modèles incluent l'information concernant les réactions chimiques et le mode opératoire du système réactionnel.

Pour l'analyse de ces modèles, il est important de distinguer les états qui dépendent des réactions de ceux qui n'en dépendent pas. Dans ce travail, le concept d'invariants réactionnels est étendu pour inclure les invariants de flux pour les systèmes réactionnels avec entrées et sorties. On propose une transformation non linéaire du modèle dynamique vers une forme normale. Les concepts de réduction du modèle, d'accessibilité d'état et de linéarisabilité par feedback sont ensuite analysés à partir de cette transformation.

Mesures de concentration

Les mesures de concentration obtenues à partir de systèmes réactionnels sont corrélées entre elles du fait de la présence de réactions et de termes d'alimentation du réacteur. On démontre que les mesures de concentration peuvent être analysées dans le cadre de la décomposition à trois niveaux basée sur la transformation sous forme normale. La factorisation résultant, appelé *factorisation des mesures de concentrations*, permet (i) la séparation de variants/invariants réactionnels et de flux, et (ii) la séparation des éléments dynamiques (avancements de réaction, effets des débits d'alimentation) de l'information purement statique (stœchiométrie, concentrations initiales et d'alimentation).

La factorisation des mesures de concentrations permet d'isoler la partie liée aux variants réactionnels en soustrayant des concentrations mesurées la partie connue liée aux invariants réactionnels. En effet, les variants réactionnels sont souvent inconnus, puisqu'ils dépendent de la description cinétique qui représente la difficulté principale de modélisation des systèmes réactionnels. En revanche, la partie liée aux invariants réac-

tionnels est habituellement connue ou mesurée. On démontre que, dans les cas où les variants réactionnels peuvent être calculés à partir des concentrations de quelques substances mesurées, les concentrations des autres substances peuvent être reconstruites en utilisant la partie connue liée aux invariants réactionnels.

L'analyse factorielle avec cibles a été employée avec succès sur des mesures de concentration pour déterminer, sans connaissance de la cinétique de réaction, le nombre de réactions et les stœchiométries correspondantes. On démontre que, en considérant uniquement la partie liée aux variants réactionnels des mesures, l'analyse factorielle est applicable. Cependant, si l'analyse factorielle doit être appliquée directement sur des mesures de concentration, la connaissance de certains invariants réactionnels est exigée afin de pouvoir spécifier des conditions nécessaires *et* suffisantes pour l'acceptation des cibles stœchiométriques.

Mesures spectrales

Dans la pratique, les mesures de concentration ne sont généralement pas disponibles pendant la réaction. Grâce à de nouvelles technologies de mesure, les mesures spectrales sont maintenant disponibles en laboratoire et en production. Divers instruments permettent une mesure indirecte et non destructive de concentration pour la plupart des substances, in-situ et en-ligne pendant une réaction. Les mesures sont disponibles à des fréquences d'échantillonnage rapides, sans retard et à de faibles coûts. En outre, dans la plupart des cas, les mesures spectrales sont linéaires, c'est-à-dire que le spectre d'un mélange est une combinaison linéaire des spectres des composants purs pondérés par les concentrations. On démontre que l'interprétation à trois niveaux basée sur la transformation sous forme normale est également applicable aux mesures spectrales de mélanges réactionnels.

Comme avec les méthodes traditionnelles d'analyse, il convient également d'estimer un modèle de calibration afin d'estimer des concentrations à partir de mesures spectrales. Toutes les méthodes de calibration exigent qu'un nouveau spectre se situe dans l'espace déterminé par les mesures spectrales de calibration (condition d'inclusion). Pour vérifier cette condition, on propose de développer un modèle de calibration uniquement pour la partie liée aux variants réactionnels. Une fois ceux-ci estimés à partir d'un nouveau spectre, les concentrations sont calculées en y ajoutant la contribution des invariants réactionnels (connue à partir des conditions expérimentales).

Les mesures de concentration pour certaines substances d'intérêt ne sont souvent pas disponibles dû aux difficultés/coûts dans le prélèvement, la préparation d'échantillon et le développement des techniques analytiques. Ainsi, la calibration traditionnelle de mesures spectrales pour l'estimation de concentrations n'est pas possible. Pour pallier au manque de mesures de concentration, la connaissance explicite ou implicite de la structure cinétique est utilisée, permettant ainsi de formuler l'approche d'analyse factorielle comme un problème de calibration.

Pour des raisons pédagogiques, les résultats sont développés pour les systèmes réactionnels isothermes, de densité constante et avec entrées/sorties. Les résultats sont ensuite étendus à d'autres scénarios tels que les systèmes réactionnels à densité et température variables. On présente aussi des factorisations de mesures de concentrations qui permettent d'inclure des mesures de température. Plusieurs cas spéciaux sont considérés:

les systèmes réactionnels continus de type CSTR, les systèmes réactionnels en semi-batch et batch, les systèmes avec des réactions en pseudo-équilibre, et les mélanges non réactionnels avec contraintes sur les concentrations.

Nomenclature

Matrices

Matrices are denoted by capital boldface latin or greek letters. If not otherwise stated in the main text, the matrix dimension is the one given in parentheses.

Latin letters

$\mathbf{A}$	spectral matrix ($K \times L$)	$\boldsymbol{\mathcal{A}}$	dynamic matrix ($n \times n$)
$\mathbf{B}$	regressor matrix	$\boldsymbol{\mathcal{B}}$	input matrix ($n \times m$)
$\mathbf{C}$	molar concentration matrix ($K \times S$)	$\boldsymbol{\mathcal{C}}_{p,in}$	diagonal matrix of the specific heat capacities of the inlets ($p \times p$)
$\mathbf{D}$	molar concentration matrix in reaction-variant form (molar RV-concentration matrix; $K \times S$)	$\mathbf{E}$	matrix of extinction coefficients or pure-component spectra ($S \times L$)
$\mathbf{F}$, $\mathbf{G}$	noise matrices	$\boldsymbol{\mathcal{F}}$	general matrix function
$\mathbf{H}$	spectral matrix in reaction-variant form (RV-spectral matrix; $K \times L$)	$\mathbf{I}$	identity matrix
$\mathbf{J}$	projection matrix	$\mathbf{J}_c$	centering matrix (special projection matrix)
$\mathbf{L}$	matrix used in Section 3.1 ($S \times \varsigma - R$)	$\mathbf{M}$	$\operatorname{diag}(\mathfrak{m}(1), \mathfrak{m}(2), \ldots, \mathfrak{m}(K))$ ($K \times K$)
$\mathbf{M}_w$	diagonal matrix of molecular weights ($S \times S$)	$\boldsymbol{\mathcal{M}}$	matrix used in Section 3.1
$\mathbf{N}$	stoichiometric matrix ($R \times S$)	$\mathbf{P}/\boldsymbol{\mathfrak{P}}$	abstract/physical loading matrix
$\boldsymbol{\mathcal{P}}$	selection matrix; matrix used in Section 3.1 ($S \times S - R$)	$\mathbf{Q}$	matrix used in Section 3.2.3 ($S \times S - \varsigma$)
$\boldsymbol{\mathcal{Q}}$	matrix used in Section 3.1($S \times S - \varsigma$)	$\boldsymbol{\mathcal{Q}}_{in}$	matrix of transformed inlet volumetric flowrates for K observations ($p \times S$)
$\mathbf{R}$, $\mathbf{R}^{\dagger}$, $\mathbf{R}^{\ddagger}$	rotation matrices	$\boldsymbol{\mathcal{R}}_r$	matrix of transformed kinetic rates for K observations ($R \times S$)
$\mathbf{T}/\boldsymbol{\mathfrak{T}}$	abstract/physical scores matrix	$\mathbf{U}$	input or design matrix

Latin letters (continued)

$\mathbf{V}$	diagonal matrix of V for K measurements ($K \times K$)	$\mathbf{W}$	weight fraction matrix ($K \times S$)
$\mathbf{X}$	matrix of reaction variants ($K \times R$)	$\mathbf{Y}$	output matrix
$\mathbf{Z}$	matrix of reaction invariants *and* flow variants ($K \times p$)		

Greek letters

$\mathbf{\Gamma}$	transpose of the reaction-invariance matrix ($S \times N$)	$\mathbf{\Theta}$	reaction spectral matrix $\mathbf{NE}$ ($R \times L$)
$\mathbf{\Lambda}$	$\mathrm{diag}\,(\lambda(1), \ldots, \lambda(K))$ ($K \times K$)	$\mathbf{\Sigma}$	matrix of singular values or eigenvalues in decreasing order
$\mathbf{\Upsilon}$	matrix of left singular vectors	$\mathbf{\Phi}$	diagonal matrix related to density
$\mathbf{\Omega}$	matrix of right singular vectors		

Vectors

Vectors are denoted by boldface small latin or greek letters or boldface symbols.

Latin letters

$\mathbf{a}$	spectrum (L-dim.)	$\mathbf{b}$	column of $\mathbf{B}$ or $\mathcal{B}$
$\mathbf{c}$	molar concentrations (S-dim.)	$\mathbf{d}$	molar concentration vector in reaction-variant form (molar RV-concentrations; S-dim.)
$\mathbf{e}$	pure-component spectrum (L-dim.)	$\mathbf{f}$	noise vector
$\mathbf{h}$	spectral vector in reaction-variant form (RV-spectral vector; L-dim.)	$-\Delta\mathbf{h}_R$	vector of reaction enthalpies [J/mol] (R-dim.)
$\boldsymbol{k}$	vector of equilibrium constants	$\mathbf{m}$	baseline slope (K-dim.)
$\boldsymbol{m}$	masses of the various species (S-dim.)	$\mathbf{n}$	reaction stoichiometry (S-dim.)
$\boldsymbol{n}$	number of moles (S-dim.)	$\mathbf{p}$	column of $\mathbf{P}$ or $\mathcal{P}$

Latin letters (continued)

$\mathbf{q}$	volume flowrates (p-dim.)	$\boldsymbol{q}_T$	extended flow vector
$\mathbf{r}$	reaction rates (R-dim.)	$\mathbf{t}$	column of $\mathbf{T}$
$\mathcal{T}$	temperature vector (K-dim.)	$\mathbf{u}$	column of $\mathbf{U}$
$\mathbf{w}$	weight fractions (S-dim.)	$\mathbf{x}$	vector of reaction variants (R-dim.)
$\boldsymbol{x}$	state vector	$\mathbf{y}$	output vector
$\mathbf{z}$	reaction-invariant and flow-variant vector (p-dim.)	$\boldsymbol{z}$	transformed state vector

Greek letters

$\boldsymbol{\alpha}$, $\boldsymbol{\beta}$	coefficient/constant vectors	$\boldsymbol{\gamma}$	column of $\boldsymbol{\Gamma}$; constant vector
$\boldsymbol{\varepsilon}$	error vector	$\boldsymbol{\theta}$	parameter vector
$\boldsymbol{\lambda}$	channel number vector (L-dim.)	$\boldsymbol{\nu}$	inlet mass flow-rates (p-dim.)
$\bar{\boldsymbol{\rho}}$	pure-component density vector (S-dim.)	$\boldsymbol{\upsilon}$	column of $\boldsymbol{\Upsilon}$
$\boldsymbol{\omega}$	column of $\boldsymbol{\Omega}$		

Symbols

$\mathbf{0}$	vector or matrix of appropriate dimension with all elements being 0	$\mathbf{1}$	vector or matrix of appropriate dimension with all elements being 1

Scalars

Scalars are denoted by plain small or capital latin or greek letters. Chemical species are abbreviated by $\mathcal{X}_i$, $i = 1, \ldots, S$.

S_b has two meanings depending on the context: (i) "S_b species" means the set of species tagged by S_b, and (ii) "$a + S_b$" means the sum of a and the cardinality S_b of the particular species subset S_b. Similarly, R_b has two meanings depending on the context: (i) "R_b reactions" means the particular reaction subset, and (ii) "$a + R_b$" means the sum of a and the cardinality R_b of the particular reaction subset R_b.

Small latin letters

c	molar concentration [M]	c_p	specific heat capacity [J/°K]
f, g, h	scalar functions	g	entry in $\boldsymbol{\Gamma}$

Small latin letters (continued)

k	time instant	k	mass action function
m	number of inputs	$\mathfrak{m}$	total mass [g]
$\mathfrak{m}_j$	mass of the jth species in a mixture [g]	n	number of states
n_{ij}	stoichiometric coefficient of the jth species in the ith reaction	$\mathfrak{n}$	element of $\boldsymbol{n}$
p	number of inlet streams or inputs	q	number of outputs
p	pressure; peak number	q_{in}	inlet volumetric flowrate [l h^{-1}]
q_{out}	outlet volumetric flowrate [l h^{-1}]	r	reaction rate [M h^{-1}]
$\mathfrak{r}_{ij}$	elementary function of c_j for the ith reaction	t	time [h]
u	element of $\mathbf{U}$; input	w	weight fraction [wt-%]
w_n	mole fraction [n-%]	$\mathcal{w}$	window size
x	extent of reaction or reaction variant [mol]	z	reaction invariant *and* flow variant [l]

Capital latin letters

A	number of factors (principal components, etc.)	A	heat-transfer area [m^2]
B	number of number of batch runs, additions, or experiments	E	normalized activation energy [K]
I	index set	K	number of observations (measurements, objects, etc.)
K	equilibrium constant	L	number of channels (wavelengths, wavenumbers, peaks, etc.)
M_w	molecular weight [g mol^{-1}]	M	stirrer torque [Nm]
N	number of types of atoms, electrical charges, etc.	O	instrumental output
$\dot{Q}_{co}$	heating/cooling power and dissipated heat power [J/h]	R	number of reactions
R	gas constant	S	number of reacting species (see Chapters 2–4); number of absorbing species (see Chapter 5)

Capital latin letters (continued)

S_k	number of known species for which the concentrations are measured in the calibration step	S_m	number of species for which the concentrations are measured (see Chapter 4); number of absorbing species observed in the selected spectral data (see Chapter 5)
S_{rn}	number of non-absorbing reaction species	S_{na}	number of non-reacting absorbing species
S_p	number of species known *a priori* to be present in the mixture for which the pure-component spectra are available	S_r	the total number of reacting species
S_{ra}	number of non-absorbing reacting species	S_t	total number of species
S_{ta}	the total number of absorbing species	S_x	number of species that have nonzero elements in $\mathbf{C}_x$; the total number of species that have nonzero elements in $\mathbf{C}_x$ with $\mathbf{C}_x$ fulfilling $\mathbf{C}_{x,n} = \mathbf{R}^{\ddagger}\,\mathbf{C}_x$ (see Section 5.3.1)
T	temperature in [°C] or [K]	U	global heat transfer coefficient through the wall in [W m^{-2} K^{-1}]
V	total volume [l]	$\dot{Q}_{sti}$	dissipated stirring power [J/h]

Small greek letters

α, β, γ, ϑ, η, μ	coefficients/constants	γ_{sn}	number of atoms, electrical charges, etc. of the nth type in the sth species
δ	dilution; Dirac impulse	ε	error
θ	parameter	κ	nominal (or specific) rate constant or pre-exponential factor
λ	state related to the outflow; channel number, wavelength, wavenumber, etc.	ν	inlet mass flowrate [g h^{-1}]; wavenumber
ρ	density [g l^{-1}]	$\bar{\rho}$	pure component density [g l^{-1}]

Small greek letters (continued)

σ	singular value; standard deviation	ς	rank of $[\mathbf{N}^{\mathrm{T}}, \mathbf{C}_{in}]$
ω	angular speed [(rad) s^{-1}]		

Operators, spaces, and special functions

AUTO$(\cdot)$	autocorrelation of a vector	$\mathcal{C}(\cdot)$	covariance matrix
diag $(\mathbf{y})$; diag $(\mathbf{Y})$	diagonal matrix containing the elements of the vector $\mathbf{y}$; vector containing the diagonal elements of the matrix $\mathbf{Y}$		
D	differential or difference operator	$\mathcal{D}(\cdot)$	dispersion matrix
Δ	difference	EE	the standard deviation of the embedded error or the average additive prediction error or the bias
IE	the standard deviation of the extracted or estimation error	ME	the standard deviation of the measurement error
$\mathcal{N}(\cdot)$	null space of a matrix	NEE	normalized bias
NIE	normalized estimation error	NRMS	normalized root mean squared error
prank $(\cdot)$	pseudo-rank or approximate-rank of a matrix	rank $(\cdot)$	mathematical rank of a matrix
$\mathcal{R}(\cdot)$	correlation matrix or reachable set	$\mathbb{R}^{p\times q}$	real space of dimension p by q
RMS	root mean squared error	$\mathcal{S}_c(\cdot)$	subspace spanned by the columns of a matrix (column space)
$\mathcal{S}_r(\cdot)$	subspace spanned by the rows of a matrix (row space)	$\mathcal{V}(\cdot)$	variance
τ	transformation (diffeomorphism)	$\mathcal{V}$	neighborhood
$\|\cdot\|_F$	Frobenius or Euclidian (matrix) norm	$\|\cdot\|_p$	p-norm of vectors and matrices
$\equiv$	definition	$:=$	redefinition

Subscripts

Small letters

a	abstract	c	related to $\mathbf{c}$, or partial derivative with respect to $\mathbf{c}$; related to calibration set
cv	related to cross validation	d	dependent
e	extended	ext	external
f	final	g	augmented (appended)
i, j	running indices	id	independent
in	inlet stream	k	related to the S_k species
l	locally; dependent	m	related to the S_m species
max	maximal	min	minimal
n	new; related to mole fraction	$\mathfrak{n}$	related to $\boldsymbol{n}$
rn	related to the S_{rn} species	na	related to the S_{na} species
o	observed	out	related to outlet streams
p	dimension of $\mathbf{1}$ or $\mathbf{0}$-vectors; related to projection error; related to the S_p species	q	related to inlet streams
r	related to the S_r species; related to reaction variants	ra	related to the S_{ra} species
s	setpoint; partition	t	related to the S_t species; partition
ta	related to the S_{ta} species	tar	target
u	unmeasured; unknown, interferent	w	related to weight fractions
x	experimental	$\mathfrak{x}$	related to $\boldsymbol{x}$
y, z	partitions	$\mathfrak{z}$	related to $\boldsymbol{z}$ or $\mathfrak{Z}$

Capital letters

A	related to the number of factors	L	related to number of channels
R	related to the reactions	T	partial derivative with respect to temperature T; related to temperature

Symbols

0	initial	(I)	without a quantity I

Symbols (continued)

$-$	before	$+$	after

Superscripts

Superscripts can have different meanings when applied to scalars, vectors or matrices.

Unique meaning

b	identifier of a subset of species	p	product
r	reactant	T	transpose of a vector or matrix
$-$	consumed; before limit	$*$	optimal
$\perp$	orthogonal complement	$\star$	related

Applied to matrices

$+$	unique Moore–Penrose pseudo-inverse	$+T$	pseudo-inverse transpose
-1	inverse	$-T$	inverse transpose

Applied to scalars

$+$	produced	$-$	consumed

Accents

Accents can have different meanings when applied to scalars, vectors or matrices.

Unique meaning

$\dot{a}$	derivative with respect to time; finite difference with respect to time	$\hat{a}$	estimate
$\tilde{a}$	noisy or random quantity	$\acute{a}$	at a final state or at equilibrium

Applied to matrices

$\bar{\mathbf{A}}_c$	column mean-centered	$\bar{\mathbf{A}}_r$	row mean-centered

Applied to vectors

$\bar{\mathbf{a}}$	mean	$\bar{\mathbf{a}}_c$	column mean
$\bar{\mathbf{a}}_r$	row mean		

Applied to scalars

$\bar{a}$	maximal	$\underline{a}$	minimal

1

Introduction

1.1 Motivation

In recent years, constantly changing market conditions have forced many chemical producers to move to the turbulent world of multiproduct plants in order to provide rapid product changeovers. The economic environment calls for reduced development times with the goal of shorter times to market. Also, while guaranteeing safety and environmental aspects, product quality and productivity improvements of existing chemical processes become increasingly important to compete against other providers, especially when product patents expire. Consequently, the important goals of the chemical, pharmaceutical and biotechnological industries can be summarized as follows: (i) short development time (e.g., fast scale-up) and (ii) safe manufacturing of products of reproducible quality in short production times.

The chemical reactor is the heart of a chemical plant. It involves *chemical reactions* that convert specific feed material (or reactants) into marketable products. For reasons such as safety, product quality, and productivity, inlet and outlet streams are often used in industry. Furthermore, since isothermal operation is sometimes difficult to achieve or simply undesirable (nonoptimal), the temperature is varied through external heating/cooling. The resulting *reaction system* encompasses the reactions *and* the physical processes (inlets/outlets, heating/cooling, etc.).

Operating strategies (e.g., initial conditions, feeding profiles), monitoring methods (e.g., for product quality), and control designs (e.g., for reactor temperature, pH) of reaction systems are developed rather conservatively from laboratory tests (in the process development phase) and then, in some cases, adjusted and improved slowly over time on the production units. The reason for this commonly-encountered procedure is often the lack of a reasonable *first-principles model* and appropriate *chemical measurement methods* that can provide frequent concentration estimations at low cost.

First-principles models describe the state evolution (the concentrations, the temperature, and the volume) by means of conservation equations of differential nature (molar balances, heat balance, continuity equation) and constitutive equations of algebraic nature (e.g., equilibrium or rate expressions). They usually include information regarding both the underlying reactions (e.g., *stoichiometry*[1], heats of reaction, reaction

[1] In biotechnological reaction systems, yield coefficients are often used instead of stoichiometries.

rates) and the operational mode of the reaction system (e.g., initial conditions, material exchange terms, operational constraints).

Two types of concentration measurements are commonly found in industry: weight fraction and molar concentration. However, concentration measurements during the course of a reaction are not generally available, neither on-line nor off-line. If concentrations are measured, they are mostly used to provide post-production quality control assessment and demonstrate waste discharge compliance (American Chemical Society *et al.*, 1996).

Concentrations are rarely analyzed for samples taken *during* the course of a reaction. There are several reasons for this, including difficulties/costs in sampling, sample preparation, and development of the wet-chemical analysis methods. Owing to new measurement technologies, spectral measurements (such as vibrational or emission spectroscopy) are now available in both the laboratory and production. Various spectral instruments enable non-destructive indirect concentration measurement of most species *in-situ*/on-line during a reaction. Measurements are available at high sampling rates, low costs and delay-free.

Based on these measurements and prior knowledge regarding the reaction system, it would be challenging to meet the industrial goals mentioned above by solving the following problems:

(P1) Quantitative estimation/prediction of concentrations.
(P2) Building first-principles models.
(P3) Monitoring, control, and optimization of reaction systems.

The solution to Problem P2 usually reduces time and costs in the laboratory as it enables simulation of various scenarios, and the solution to Problem P3 typically makes production processes safer and more efficient. First-principles models also play a key role in solving Problem P3. The availability of concentration estimates (solution to Problem P1) forms the basis for solving Problems P2 and P3: For building first-principles models, concentration estimates must be available *off-line*, while for monitoring, control, and optimization of reaction systems, it is necessary to have them *on-line*.

This dissertation develops a methodology and appropriate tools that help solve the three problems for homogeneous reaction systems with inlet and outlet streams.

Quantitative estimation/prediction of concentrations

For traditional wet-chemical analyses, a calibration model for the estimation of concentrations is typically required. This holds true also for the estimation of concentrations from spectral data. For building a calibration model, all the species expected to be present in a mixture of unknown concentrations must be available and varied independently. For *non-reacting (static) calibration data*, numerous mixtures with concentrations selected according to an experimental plan are usually prepared (Box *et al.*, 1978). The difficulty with this approach is that possibly highly-reactive intermediates must be available — a requirement that is often difficult and sometimes impossible to meet. This difficulty is circumvented by taking samples during the course of the

reaction. This way, the intermediates are produced by the reactions, and the need to vary them externally disappears.

However, the concentration measurements of such samples are often difficult to obtain for the reasons mentioned above. Thus, traditional calibration for the purpose of concentration estimation is often not feasible. For such situations, factor-analytical techniques are available to estimate concentrations from spectral measurements. They require some *prior knowledge* of, for example, the kinetic *structure*.

Concentration data contain *superposed* information regarding both the underlying reactions (*reaction-variant part*) and the operational mode of the reactor (including the rates of mixing and transport processes; *reaction-invariant part*). It would be very helpful to be able to separate these two parts and to identify the reaction stoichiometries (and the reaction rates) directly from the reaction-variant part.

Building first-principles models

The rates of transport processes can usually be estimated adequately from the properties of e.g. the individual species, the flow characteristics (Perry and Green, 1984). However, the kinetic description represents the main difficulty in modeling chemical reaction systems. For most industrially-relevant reactions, the kinetic parameters cannot be estimated reliably from theory and, thus, must be determined experimentally from concentration data. This estimation requires the system stoichiometry *and* candidate kinetic structures to be available. In practice, however, it would be preferable to identify the reaction stoichiometries from the available data independently of reaction kinetics. In other words, instead of fitting a global model (stoichiometric and kinetic) to measured data, it would be better to proceed in two steps: (i) determine the reaction stoichiometries and extents of reaction from measured data without knowledge of kinetics, and (ii) determine the kinetic structure for each reaction individually from the corresponding extent of reaction computed in the first step. Only that step is considered in this dissertation.

Postulate

For batch reaction systems, the reaction invariants are constants. However, for reaction systems with inlet and/or outlet streams, some of the reaction invariants vary with the inlet and/or outlet streams. For such systems, the concept of flow variants/invariants can help in the analysis of reaction systems.

The following postulate can be formulated:

> The concept of *reaction and flow variants/invariants* helps in the analysis of methods for solving Problems P1–P3. Furthermore, this concept is crucial for fine-tuning methods such as multivariate calibration and curve-resolution techniques.

This postulate, which forms the basis of this dissertation, needs to be investigated and proven. For this purpose, three representations will be derived for reaction systems and reaction data: a first-principles dynamic model, a factorization of concentration data, and a factorization of spectral data. In the three representations, the concept of reaction and flow variants/invariants will enable to separate the contribution of the

reactions (*reaction variants*) from the direct contribution of the inlet streams (*reaction invariants and flow variants*) and the terms related to the initial conditions (*reaction and flow invariants*).

1.2 State of the art

Reaction variants/invariants

For the analysis of first-principles models, it is important to distinguish between the states that depend on the reactions and those which do not. A *linear* transformation was proposed by Waller and Mäkilä (1981) to separate the reaction-variant states (same dimension as the number of independent reactions) from the reaction-invariant states. This transformation requires the knowledge of stoichiometry but not of kinetics. Even when inlet and outlet streams are present, the reaction invariants can be separated from the reaction variants by the transformation proposed by Waller and Mäkilä (1981). However, the reaction invariants are influenced by the flows and, thus, evolve with time (Fjeld *et al.*, 1974). The transformation of Waller and Mäkilä (1981) does not separate out the reaction *and* flow invariants.

Concentration measurements

Concentration measurements are often used for discriminating among reaction mechanisms (stoichiometry and kinetics) proposed by the chemist, and for estimating parameters in kinetic models (Himmelblau *et al.*, 1967; Hill, 1977; Perry and Green, 1984). These activities usually require the knowledge of candidate kinetic structures.

When the kinetic structure is not available, concentration data can still be used for two important activities: identification of stoichiometric models (using target factor analysis, TFA) and on-line state reconstruction.

TFA is a useful multivariate analysis tool to determine the number of independent reactions and the corresponding stoichiometries (Hamer, 1989; Bonvin and Rippin, 1990). However, the approach is restricted to cases where the reaction-variant part in the measured data can be isolated from the reaction-invariant part.

On-line state reconstruction enables to reconstruct the concentrations of the species of interest from those species which are routinely measured on-line. Under certain conditions, on-line state reconstruction can also be implemented without knowledge of kinetics. An advantage of this method is that it only requires an estimate of the initial concentrations of the species to be reconstructed. Bastin and Dochain (1990) showed that asymptotic convergence to the true concentrations can be guaranteed (*asymptotic observer*).

Spectral measurements

Spectral data are typically analyzed using multivariate calibration or curve resolution methods (e.g., factor analysis) depending on whether or not measured concentrations are available.

Multivariate calibration methods can predict concentrations from measured spectral

data on the basis of a calibration model (Martens and Naes, 1989; Brown, 1993). For the *species of interest*, the spectra and the corresponding concentrations are usually available for some observations (e.g., from wet-chemical analyses). The calibration model is also valid when the concentrations of some interferents (species of no interest) are unknown in the calibration step (Krutschkoff, 1967; Martens and Naes, 1989; Crocombe *et al.*, 1984; Haaland and Thomas, 1988). Also, it is a well-known fact that the region of applicability is wide enough to include a certain class of nonlinear models (Gemperline *et al.*, 1991; Bak and Larsen, 1995).

Most of the calibration methods have been developed for non-reacting spectral data. In contrast to non-reacting mixtures, however, data from reacting mixtures are structured due to the underlying chemical reactions, and (black-box) calibration methods often fail due to collinearity of concentrations.

Factor-analytical techniques involve two steps: (i) principal component analysis (Jackson, 1991), (ii) rotation of scores and loading matrices into physically-meaningful quantities using prior knowledge about the underlying system and the type of instrumental response. They have been used successfully for *non-reacting mixtures* to resolve the pure-component spectra and the corresponding concentrations from spectral data. For a detailed review, the reader is referred to Malinowski (1991); Tauler *et al.* (1995); Gemperline (1989); Hamilton and Gemperline (1990); some important FA techniques include TFA (Malinowski, 1991), evolving factor analysis (EFA; Maeder, 1987; Keller *et al.*, 1992; Brereton and Elbergali, 1994; Toft, 1995), iterative target factor analysis (IT-TFA; Vandeginste *et al.*, 1987), alternating regression (AR; Karjalainen, 1989; Tauler *et al.*, 1995), interactive self-modeling multivariate analysis (Windig, 1990); window evolving factor analysis (WEFA; Keller and Massart, 1991; Malinowski, 1992), incremental target factor analysis (IncTFA; Prinz, 1992), heuristic evolving latent projection (HELP; Kvalheim and Liang, 1992).

TFA can be used to individually test the existence of candidate species (described by pure-component target spectra) in spectral data. Once pure-component target spectra for all species absorbing in the unknown mixtures have been accepted, the corresponding concentration matrix can be reconstructed. However, the problem with such an approach is that pure-component target spectra must be available from a database.

For some types of spectral instruments (e.g., mid-infrared, nuclear magnetic resonance, mass spectroscopy), normalized pure-component target spectra exist, which can be used for spectral identification or finger printing using TFA (de Jong, 1991; Vogels *et al.*, 1993). However, they cannot be used for *quantitative* concentration estimation. Also, for numerous types of spectral instruments, pure-component target spectra are not available from a database, since they often depend on the specific instrument, the experimental conditions, and physico-chemical interactions between the species (Burns and Ciurzak, 1992).

Most of the FA techniques rely on the assumption that the rank of the spectral data is determined by the number of absorbing species. For reacting mixtures, however, this assumption is usually not satisfied.

1.3 Objectives of the dissertation

This work contributes to the system-theoretical analysis of dynamic reaction models. It introduces the concept of reaction and flow variants/invariants. In the light of this concept, model reduction, state accessibility and the construction of a linearizing feedback are investigated. Furthermore, special emphasis is drawn on the implications with regard to this concept on the utilization of concentration and spectral data. More specifically, the following activities are studied: (i) the reconstruction of concentrations for all species from the concentrations of a few measured species (on-line state reconstruction), (ii) the identification of stoichiometric models based on concentration data using TFA, and (iii) the estimation of concentrations from spectral measurements using multivariate calibration and factor analysis. For the three representations proposed in this dissertation, Table 1.1 summarizes the corresponding objectives and tools used. Note that most of the tasks do *not* require knowledge of kinetics.

Reaction and flow variants/invariants

The concept of reaction invariants of Waller and Mäkilä (1981) will be extended to include flow invariants of reaction systems with inlet and outlet streams, thus leading to a decomposition of the state evolution into reaction and flow variants/invariants. In contrast to Waller and Mäkilä (1981) who used a *linear* transformation, a *nonlinear* transformation will be proposed here. Furthermore, the concept will be extended to include the heat balance equation. Model reduction, state accessibility, and feedback linearizability will also be studied.

Concentration measurements

Concentration data will be analyzed using the knowledge gained from the system-theoretical analysis of the dynamic model. A data pre-treatment will be proposed that separates the reaction-variant and reaction-invariant parts.

The convergence of the asymptotic observer proposed by Bastin and Dochain (1990) will be studied in the light of the reaction and flow variants/invariants.

The TFA results of Hamer (1989) and Bonvin and Rippin (1990) will be extended to handle measured concentration data directly. Furthermore, the case of unmeasured species will be considered. Special emphasis will be given to the specification of necessary and sufficient conditions for the acceptance of stoichiometric targets.

Spectral measurements

Spectral data from reacting mixtures are factorized using the knowledge gained from the two previous analyses. Also, a data pre-treatment will be proposed that separates the reaction-variant and reaction-invariant parts. Certain useful results regarding the determination of the number of reactions and absorbing species from spectral data will be derived.

For the calibration of reacting spectral data, methods will be proposed that specify the validity range of a calibration model in terms of the experimental conditions of a new process run. Furthermore, methods will be proposed that enable choosing the reaction

Table 1.1. The representations proposed in this dissertation for investigating reaction systems with the corresponding objectives and tools. There is also an indication of whether or not kinetic knowledge is required.

	Dynamic model	Factorization of concentration data	Factorization of spectral data
Objectives in this study	Study of system properties	On-line state reconstruction; Stoichiometric modeling	Concentration estimation
Tools used	Model reduction[a] State accessibility[a] Feedback linearizability[b]	Asymptotic observer[a] Target factor analysis[a]	Calibration[a] Factor analysis[c]

[a] No knowledge of kinetics.

[b] Knowledge of kinetics required only for implementing of the linearizing feedback.

[c] Some knowledge of kinetic *structure* required.

invariants for the new set independently of those of the calibration set.

For situations where concentration measurements are not available for calibration, FA will be used. Methods will be proposed that (i) render most existing FA methods applicable to reacting spectral data, and (ii) construct a calibration set from implicit or explicit knowledge of kinetic structure with the goal of substituting for the missing calibration concentration measurements. The latter FA method will be formulated in the framework of calibration of reacting spectral data.

1.4 Organization of the dissertation

For pedagogical reasons, the results of this dissertation are developed for isothermal, constant-density reaction systems with inlet and outlet streams (*basic dynamic model*). The results are then extended to various scenarios such as reaction systems with varying density and temperature. Also, factorizations of *extended concentration data* that include temperature or calorimetric measurements will be presented. Furthermore, for spectral data, results are first developed under the assumption that all species both react and absorb. The results are then extended to take into account the presence of non-absorbing reacting species and/or of non-reacting absorbing species.

Several special cases are considered. These include continuous stirred-tank reaction systems, semibatch and batch reaction systems, systems with reactions in quasi-equilibrium conditions, and non-reacting mixtures with closure.

The main concepts and results are illustrated by simulated examples. For the sake of readability, the numerical values of the parameters for the simulated examples and the proofs of the main results are pushed to Appendices A and H, respectively. Furthermore, the main assumptions for the various representations (dynamic model and

factorizations of concentration and spectral data) are given in Appendix D.

An introduction to the main tools used in this dissertation such as on-line state reconstruction, TFA, multivariate calibration, and FA are given in Appendices B (linear algebra) and C (statistics). The applications of these tools to reaction systems are presented in the particular sections of the main text. The factorization of general (linear) spectral data is derived in Appendix F, and a non-exhaustive list of instruments, for which the results developed in this dissertation are applicable, is provided. Local rank useful for factor analysis of spectral data is studied in Appendix G.

Chapter 2 derives the basic dynamic model, its extensions, and special cases of reaction systems. Based on this dynamic model, the concept and some implications of reaction and flow variants/invariants are introduced in Chapter 3. Chapter 4 analyzes concentration data by introducing a factorization of concentration data. Furthermore, the importance of this factorization on the following two activities is illustrated: on-line state reconstruction and stoichiometric modeling. Based on the knowledge gained in the previous chapters, Chapter 5 derives a factorization of spectral data and illustrates its importance on the following two activities: multivariate calibration and curve-resolution (factor analysis) for the estimation of concentrations from spectral data. Chapter 6 presents concluding remarks and an outlook for future work.

2

Modeling chemical reaction systems

In current practice of the chemical/biotechnological industry, after screening experiments have been performed, response surface models (RSM; Box *et al.*, 1978) are often used in process development to describe the relationships between some important factors (e.g., initial conditions, catalyst type) and responses (e.g., quality and productivity measures). Since the structure is usually highly empirical (e.g., a polynomial model including linear, quadratic and interaction terms), RSMs are only valid within the experimental range. Thus, if the goal of an RSM is to determine the optimal values of dominant factors, numerous experiments using widely varying conditions are required for building the RSM.

Alternatively, first-principles models (kinetic and thermodynamic) can be used to simulate a wide range of conditions outside the experimental range (extrapolation). This way, process development and process design using flowsheet programs along with tools for capacity planning and scheduling for complete production lines can be greatly simplified, and, for example, the large costs of pilot experiments can often be avoided. Furthermore, model-based monitoring, control, and optimization in both process development and production can be used, thereby leading to improved product quality, productivity, and process safety.

Herein, a first-principles (dynamic) model for nonisothermal homogeneous reaction systems with inlet and outlet streams and varying density will be derived.

Without loss of generality, it will be assumed throughout the dissertation that (i) the reactions are independent, i.e., both the stoichiometries (describing the reaction network) and the reaction kinetics are linearly independent and the stoichiometries are constant, and (ii) the inlets are independent, i.e., both the inlet concentrations and flowrates are linearly independent and the inlet concentrations are constant.

In many practical applications, however, the stoichiometries and/or kinetics are linearly dependent. In some cases, the stoichiometries can even be time-varying. Furthermore, owing to variability in the raw materials or in the preprocessing steps, the inlet concentrations might vary and/or be linearly dependent. Also, the inlet flowrates might be linearly dependent due to feeding at proportional rates. For meeting the assumptions just mentioned, two transformations will be proposed: the *transformation to independent reactions* and the *transformation to independent inlets*.

For modeling, control and optimization of reaction systems, it is important to be able to influence the reactions in a desired manner by manipulating the available experimental

conditions (initial conditions, input variables). In the context of experimental planning for modeling, it will be shown that impulse additions and multiple process runs can be cast in the framework of *generalized inlet streams* that can be handled in a way similar to standard inlet streams.

In the next section, the basics of stoichiometry and kinetics are briefly revisited. Section 2.2 introduces the transformation to independent reactions and provides useful results regarding the theoretical stoichiometric space of a reaction system. Section 2.3 derives the mole and heat balance equations, and the continuity equation for reaction systems with (i) constant density and temperature (basic model), (ii) varying density and constant temperature, and (iii) constant density and varying temperature. Extensions to the basic model are provided in Section 2.4 such as (i) expression of the mole balance equations in weight fractions, (ii) varying inlet concentrations (transformation to independent inlets), and (iii) impulse additions and multiple process runs (generalized inlets). Special cases are studied in Section 2.5 such as continuous stirred-tank reaction systems (CSTRs), batch and semibatch reaction systems, and systems with reactions in quasi-equilibrium conditions.

2.1 Reaction network and kinetics

2.1.1 Reaction network

A chemical reaction converts specific feed material (or reactants) into certain products. It follows a quantitative relationship that is usually described by the *stoichiometry*.

Let $\mathcal{X}_j$ denote the jth reacting species ($j = 1, \ldots, S$), where S denotes the total number of reacting species. The reaction network involves a set of R coupled irreversible reactions that are described as:

$$\sum_{j=1}^{S} n_{ij}^{-} \mathcal{X}_j \rightarrow \sum_{j=1}^{S} n_{ij}^{+} \mathcal{X}_j, \quad n_{ij}^{-},\, n_{ij}^{+} \geq 0, \quad i = 1, \ldots, R \tag{2.1}$$

with n_{ij}^{-} and n_{ij}^{+} being nonnegative numbers expressing the relative number of moles of the jth component that is consumed or produced by the ith reaction (Perry and Green, 1984).

The species subsets $\mathcal{X}^r$ and $\mathcal{X}^p$ are particularly useful throughout this work:

- *A reactant* is a species that appears on the left-hand side of at least one reaction:

$$\mathcal{X}^r = \{\mathcal{X}_j |\ n_{ij}^{-} > 0 \text{ for some } i\}, \quad I^r = \{j |\ n_{ij}^{-} > 0 \text{ for some } i\}.$$

  Let also I^{ri} denote the set of indices for the reactants involved in the ith reaction.

- *A product* is a species that appears on the right-hand side of at least one reaction:

$$\mathcal{X}^p = \{\mathcal{X}_j |\ n_{ij}^{+} > 0 \text{ for some } i\}, \quad I^p = \{j |\ n_{ij}^{+} > 0 \text{ for some } i\}.$$

If a species $\mathcal{X}_j$ is both reactant and product in the Reaction i and if $n_{ij}^+ > n_{ij}^- > 0$, then it is called an *autocatalyst*; if $n_{ij}^+ = n_{ij}^- > 0$, then $\mathcal{X}_j$ is a *catalyst*. In contrast to catalysts, *solvents* are not considered as one of the S reacting species and, thus, does not appear in the reaction scheme.

By defining the *stoichiometric coefficient* $n_{ij} \equiv n_{ij}^+ - n_{ij}^-$, the $R \times S$ *stoichiometric matrix* $\mathbf{N}$ becomes:

$$\mathbf{N} = \begin{bmatrix} n_{11} & n_{12} & \cdots & n_{1S} \\ \cdots & \cdots & \cdots & \cdots \\ n_{R1} & n_{R2} & \cdots & n_{RS} \end{bmatrix}. \tag{2.2}$$

Thus, reactants and products have negative and positive entries in $\mathbf{N}$, respectively.

Reversible reactions refer to the case in which conversion of reactants to products is far from complete, and the reverse reaction becomes important. A reversible reaction is usually denoted as:

$$\sum_{j=1}^{S} n_{ij}^- \mathcal{X}_j \rightleftharpoons \sum_{j=1}^{S} n_{ij}^+ \mathcal{X}_j, \quad n_{ij}^-, n_{ij}^+ \geq 0. \tag{2.3}$$

However, similarly to Bastin *et al.* (1992), the consistency with the description of a system of irreversible reactions (2.1) requires to encode the reversible reaction (2.3) as two separate irreversible reactions:

$$\begin{aligned} &\sum_{j=1}^{S} n_{ij}^- \mathcal{X}_j \rightarrow \sum_{j=1}^{S} n_{ij}^+ \mathcal{X}_j, \\ &\sum_{j=1}^{S} n_{ij}^+ \mathcal{X}_j \rightarrow \sum_{j=1}^{S} n_{ij}^- \mathcal{X}_j, \end{aligned} \qquad n_{ij}^-, n_{ij}^+ \geq 0. \tag{2.4}$$

Thus, each reversible reaction is accounted for twice in the reaction network given by (2.2):

$$\mathbf{N} = \begin{bmatrix} \cdots & \cdots & \cdots & \cdots \\ n_{i1} & n_{i2} & \cdots & n_{iS} \\ -n_{i1} & -n_{i2} & \cdots & -n_{iS} \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}. \tag{2.5}$$

Remark 2.1 (Yield coefficients in biotechnological reaction systems)
In biotechnological reaction systems, n_{ij} denotes a *yield coefficient*, i.e., the consumption or production rate of the jth species in the ith reaction per consumption or production rate of a *normalizing species*. The coefficient of the normalizing species is defined as 1. The major difference between stoichiometric coefficients in chemical and biotechnological reaction systems is that, in chemical reaction systems, n_{ij} usually is an integer, while in biotechnological reaction systems, it is typically a real number.

Example 2.2 (Reversible reaction — esterification)
Catalyzed by sulfuric acid ($\mathcal{X}_5$), ethanol ($\mathcal{X}_1$) reacts with acetic acid ($\mathcal{X}_2$) in a reversible reaction to produce acetic acid ethyl ester ($\mathcal{X}_3$) and water ($\mathcal{X}_4$). The reaction network and the stoichiometric matrix are according to (2.5):

$$\mathcal{X}_1 + \mathcal{X}_2 + \mathcal{X}_5 \rightleftharpoons \mathcal{X}_3 + \mathcal{X}_4 + \mathcal{X}_5 \quad (2.6) \qquad \text{with} \qquad \mathbf{N} = \begin{bmatrix} -1 & -1 & 1 & 1 & 0 \\ 1 & 1 & -1 & -1 & 0 \end{bmatrix}.$$

Note that since $n_{i5}^- = n_{i5}^+ = 1$, $n_{i5} = 0$ $(i = 1, 2)$. Thus, $R = 2$ and $S = 5$. All species are reactants and products, i.e., $I_r = I_p = \{1, 2, \ldots, 5\}$.

Example 2.3 (Cyclic reaction)
Consider three consecutive reactions where $\mathcal{X}_1$ reacts to $\mathcal{X}_2$, $\mathcal{X}_2$ to $\mathcal{X}_3$, and $\mathcal{X}_3$ to $\mathcal{X}_1$. Since the initial reactant $\mathcal{X}_1$ is also the final product of the consecutive reactions, this reaction system is called *cyclic*.

$$\mathcal{X}_1 \rightarrow \mathcal{X}_2 \rightarrow \mathcal{X}_3 \rightarrow \mathcal{X}_1 \qquad (2.7) \qquad \text{with} \qquad \mathbf{N} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}.$$

All the species are reactants and products and, thus, $I_r = I_p = \{1, 2, 3\}$.

Example 2.4 (Parallel reactions — dimerisation and hydration)
1-phenyl-1-butoxyethan ($\mathcal{X}_3$) is produced in a hydration reaction from styrene ($\mathcal{X}_1$) and n-butanol ($\mathcal{X}_2$). In an irreversible side reaction, $\mathcal{X}_1$ dimerizes to (trans)-1,3-diphenylbut-1-ene ($\mathcal{X}_4$). There also exists a diastereoisomere of $\mathcal{X}_4$, (cis)-1,3-diphenylbut-1-ene. However, the two diastereoisomeres are considered here as one species (see the chemical structures in Appendix A.4). The main reaction can be modeled as:

(a) An irreversible reaction. Thus:

$$\begin{aligned} \mathcal{X}_1 + \mathcal{X}_2 &\rightarrow \mathcal{X}_3 \\ 2\mathcal{X}_1 &\rightarrow \mathcal{X}_4 \end{aligned} \quad (2.8) \qquad \text{with} \qquad \mathbf{N} = \begin{bmatrix} -1 & -1 & 1 & 0 \\ -2 & 0 & 0 & 1 \end{bmatrix}.$$

The reactants and products can be identified as $I_r = \{1, 2\}$ and $I_p = \{3, 4\}$, respectively.

(b) A reversible reaction. Thus:

$$\begin{aligned} \mathcal{X}_1 + \mathcal{X}_2 &\rightleftharpoons \mathcal{X}_3 \\ 2\mathcal{X}_1 &\rightarrow \mathcal{X}_4 \end{aligned} \quad (2.9) \qquad \text{with} \qquad \mathbf{N} = \begin{bmatrix} -1 & -1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ -2 & 0 & 0 & 1 \end{bmatrix}.$$

The reactants and products can be identified as $I_r = \{1, 2, 3\}$ and $I_p = \{1, 2, 3, 4\}$, respectively.

Example 2.5 (Ethanolysis reaction system)
In two successive irreversible ethanolysis reactions (de Vallière, 1989), phthalyl chloride monoethyl ester ($\mathcal{X}_3$) and phthalic diethylester ($\mathcal{X}_5$) are produced from phthalyl chloride ($\mathcal{X}_1$) and ethanol ($\mathcal{X}_2$). Both reactions also produce hydrochloric acid ($\mathcal{X}_4$). It is assumed that $\mathcal{X}_2$ also reacts with $\mathcal{X}_4$ in a reversible side reaction to ethyl chloride ($\mathcal{X}_6$) and water ($\mathcal{X}_7$). The reaction system can be described by the following reaction scheme:

$$\begin{aligned} &\mathcal{X}_1 + \mathcal{X}_2 \xrightarrow{r_1} \mathcal{X}_3 + \mathcal{X}_4 \\ &\mathcal{X}_3 + \mathcal{X}_2 \xrightarrow{r_2} \mathcal{X}_5 + \mathcal{X}_4 \\ &\mathcal{X}_4 + \mathcal{X}_2 \overset{r_3}{\rightleftharpoons} \mathcal{X}_6 + \mathcal{X}_7 \end{aligned} \quad (2.10) \qquad \text{with} \qquad \mathbf{N} = \begin{bmatrix} -1 & -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & -1 & -1 \end{bmatrix}.$$

Thus, $R = 4$ and $S = 7$. The reactants and products can be identified as $I_r = \{1, 2, 3, 4, 6, 7\}$ and $I_p = \{2, 3, 4, 5, 6, 7\}$, respectively.

2.1.2 Reaction kinetics

Let the molar concentration $c_j(k)$ be defined as:

$$c_j(k) \equiv \frac{\mathfrak{n}_j(k)}{V(k)}, \quad \forall j = 1, \dots, S, \tag{2.11}$$

where $\mathfrak{n}_j$ is the number of moles of the jth species, and V the reactor volume. The unit is usually given in [mol l^{-1}] or, abbreviated, in [M].

The *reaction rate*, r_i, expresses the rate of reactants consumption or products formation in the ith reaction. It often is a nonlinear function of the S-dimensional molar concentration vector $\mathbf{c}$, satisfying Assumptions A1–A2 (see Appendix D). The R reaction rates are represented in the R-dimensional vector of reaction rates $\mathbf{r}$, also called the *reaction kinetics* or *conversion rates*.

In many practical applications, a reaction rate can be modeled as (Bastin *et al.*, 1992):

$$r_i(\mathbf{c}, \boldsymbol{\theta}^i) = \kappa_i \prod_{j=1}^{S} \mathfrak{r}_{ij}(c_j, \boldsymbol{\theta}^{ij}), \quad \boldsymbol{\theta}^i \equiv \begin{bmatrix} \kappa_i \\ \boldsymbol{\theta}^{i1} \\ \dots \\ \boldsymbol{\theta}^{iS} \end{bmatrix}, \tag{2.12}$$

where κ_i is a nominal (or specific) *rate constant*. The *elementary kinetic function* $\mathfrak{r}_{ij}$ with the corresponding kinetic parameter vector $\boldsymbol{\theta}^{ij}$ ($j \in I^{ri}$) represents the influence of c_j on the ith reaction rate. In vector notation,

$$\mathbf{r} = \mathbf{r}(\mathbf{c}, \boldsymbol{\theta}), \tag{2.13}$$

where $\boldsymbol{\theta}$ is the kinetic parameter vector containing $\boldsymbol{\theta}^i$ for $i = 1, \dots, R$.

Basically, three extreme situations can be distinguished:

(S1) $\mathcal{X}_j$ *does not affect* the reaction rate r_i:

$$\mathfrak{r}_{ij}(c_j) = 1.$$

(S2) $\mathcal{X}_j$ is an *activator* for the ith reaction, i.e., it *positively affects* the reaction rate r_i. This can be modeled by a positive, monotonic increasing and concave function of c_j:

$$\mathfrak{r}_{ij}(c_j) \geq 0, \ \frac{\partial \mathfrak{r}_{ij}(c_j)}{\partial c_j} > 0, \ \frac{\partial^2 \mathfrak{r}_{ij}(c_j)}{\partial c_j^2} < 0.$$

(S3) $\mathcal{X}_j$ is an *inhibitor* for the ith reaction, i.e., it *negatively affects* the reaction rate r_i. This can be modeled by a positive, monotonic decreasing and convex function of c_j:

$$\mathfrak{r}_{ij}(c_j) \geq 0, \ \frac{\partial \mathfrak{r}_{ij}(c_j)}{\partial c_j} < 0, \ \frac{\partial^2 \mathfrak{r}_{ij}(c_j)}{\partial c_j^2} > 0.$$

From knowledge of the reaction mechanism, the elementary kinetic functions can be derived (see Ch. 4 of Perry and Green, 1984)). In many practical situations, however, the exact reaction mechanism is unknown, and a model for the reaction rates cannot be derived from the stoichiometry. Thus, the elementary kinetic functions must be modeled empirically. For example, the three extreme situations S1–S3 can be modeled by a single rational fraction:

$$\mathfrak{r}_{ij}(c_j, \boldsymbol{\theta}^{ij}) = \left(\frac{\alpha_{ij} + \beta_{ij}\, c_j}{1 + \gamma_{ij}\, c_j} \right)^{\vartheta_{ij}}, \quad \boldsymbol{\theta}^{ij} = \begin{bmatrix} \alpha_{ij} & \beta_{ij} & \gamma_{ij} & \vartheta_{ij} \end{bmatrix}^{\mathrm{T}}, \tag{2.14}$$

where α_{ij}, β_{ij}, and γ_{ij} are constant nonnegative coefficients, and ϑ_{ij} a constant coefficient. For $\vartheta_{ij} = 1$, the asymptotes are given by $\lim_{c_j \to 0} \mathfrak{r}_{ij} = \alpha_{ij}$ and $\lim_{c_j \to \infty} \mathfrak{r}_{ij} = \beta_{ij}/\gamma_{ij}$ for $\gamma_{ij} \neq 0$. Thus, equation (2.14) represents the kinetics of an activator when $\beta_{ij}/\gamma_{ij} > \alpha_{ij}$, and of an inhibitor when $\beta_{ij}/\gamma_{ij} < \alpha_{ij}$ (see Figure 2.1a and b). By combining two elementary functions (2.14), activation at low concentrations and inhibition at high concentrations can be modeled (see Figure 2.1c):

$$\begin{aligned} \mathfrak{r}_{ij}(c_j, \boldsymbol{\theta}^{ij}) &= \left(\frac{\alpha_{ij,1} + \beta_{ij,1}\, c_j}{1 + \gamma_{ij,1}\, c_j} \right) \left(\frac{1 + \beta_{ij,2}\, c_j}{1 + \gamma_{ij,2}\, c_j} \right) \\ \text{with } & \frac{\beta_{ij,1}\beta_{ij,2}}{\gamma_{ij,1}\gamma_{ij,2}} < \alpha_{ij,1}, \qquad \alpha_{ij,1}\,(\beta_{ij,2} - \gamma_{ij,1} - \gamma_{ij,2}) + \beta_{ij,1} > 0, \end{aligned} \tag{2.15}$$

where $\boldsymbol{\theta}^{ij} = [\alpha_{ij,1},\, \beta_{ij,1}, \beta_{ij,2},\, \gamma_{ij,1},\, \gamma_{ij,2}]^{\mathrm{T}}$. Table 2.1 lists some of the commonly found reaction types and corresponding coefficients.

Remark 2.6 (Catalyzed reaction systems)
In contrast to an autocatalyst, a catalyst $\mathcal{X}_j$ does not get consumed by the reaction and, thus, $n_{ij} = 0$. However, the catalyst may influence the elementary kinetic function

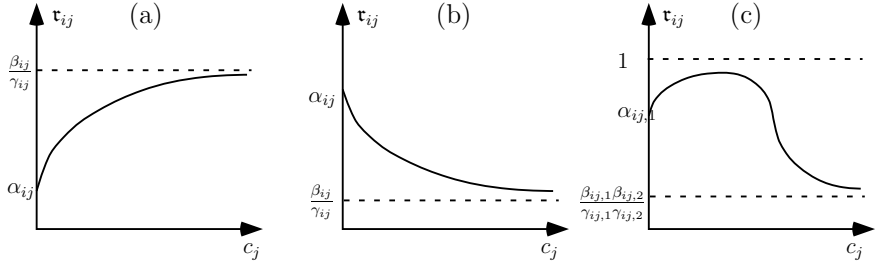

Figure 2.1. Elementary kinetic function $\mathfrak{r}_{ij}(c_j)$ modeling: (a) an activator, (b) an inhibitor, and (c) an activator at low concentrations and inhibitor at high concentrations. It is assumed that γ_{ij}, $\gamma_{ij,1}$, $\gamma_{ij,2} \neq 0$.

Table 2.1. Reaction types and corresponding coefficients of the elementary kinetic functions $\mathfrak{r}_{ij}$.

Model	α_{ij}	β_{ij}	γ_{ij}	ϑ_{ij}
Power law (linear) $c_j^{\vartheta_{ij}}$	0	1	0	$\neq 0$ (= 1 if linear)
Michaëlis–Menten $\frac{c_j}{\mu_m+c_j}$	0	μ_m^{-1}	μ_m^{-1}	1
Hyperbolic inhibition $\frac{\mu_h}{\mu_h+c_j}$	1	0	μ_h^{-1}	1

$\mathfrak{r}_{ij}$ in a nonlinear manner depending, e.g., on transport phenomena (see Ch. 4 of Perry and Green, 1984)).

2.2 Independent reactions

2.2.1 Practical situations of dependent stoichiometries and kinetics

In practice, reactions can exhibit (S1) linearly-dependent stoichiometries, (S2) linearly-dependent reaction rates on the time interval $[t_0, t_K]$, and (S3) time-varying stoichiometries. Note that the reaction rates $\mathbf{r}(t)$ are said to be linearly dependent on the time interval $[t_0, t_K]$ if there exist real numbers $\alpha_1, \alpha_2, \ldots, \alpha_R$, not all zero, such that $\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{r}(t) = 0$ for *all* $t \in [t_0, t_K]$, where $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_R]^{\mathrm{T}}$ (Chen, 1984). Otherwise, $\mathbf{r}(t)$ is said to be linearly independent on $[t_0, t_K]$.

Some practical Situations S1–S3 are listed next.

(S1) Linearly-dependent stoichiometries can occur, for e.g.:

- *Reversible reactions.* In (2.5), the two rows in $\mathbf{N}$ are multiples of each other with the factor -1.
- Reaction systems with $R > S$. Since $\operatorname{rank}(\mathbf{N}) \leq \min(R, S) = S$ (see Appendix B.2 for the definition of rank), the stoichiometries of $(S - R)$ reactions are linearly dependent.
- Cyclic reactions.

(S2) Linearly-dependent kinetics on the interval $[t_0, t_K]$ can occur, for e.g.:

- Reaction systems in *dynamic equilibrium.* After a transition time, the kinetic rates become linearly dependent (see Example 2.4a below)
- Biotechnological reactions due to the coupling of energy production and consumption (Duboc, 1997).

(S3) Some biotechnological reaction systems exhibit time-varying yield coefficients (stoichiometries) due to:

- Changing energy coupling (Duboc, 1997).
- Changing reaction conditions such as temperature or pH during a batch or from batch to batch (Chen *et al.*, 1995).

2.2.2 Theoretical stoichiometric space and transformation to independent stoichiometries

Consider Situation S1 (with constant stoichiometries and linearly-independent kinetics) where R_N rows of $\mathbf{N}$ are linearly independent ($\text{rank}\,(\mathbf{N}) = R_N$). Thus, $\mathbf{N}$ can be decomposed as follows:

$$\mathbf{N} = \boldsymbol{\mathcal{P}}_{r1}\check{\mathbf{N}}, \tag{2.16}$$

where $\check{\mathbf{N}}$ is an $R_N \times S$ matrix containing the R_N linearly-independent stoichiometries, and $\boldsymbol{\mathcal{P}}_{r1}$ an $R \times R_N$ matrix of full rank.

Let $\boldsymbol{\Gamma}$ be an $S \times N$ matrix, where γ_{sn} indicates the number of atoms, electrical charges, etc. of the nth type in the sth chemical species, and N the number of quantities conserved. For example, a column of $\boldsymbol{\Gamma}$ can contain the molecular weights of the S chemical species. Note that $\boldsymbol{\Gamma}$ is a matrix with only nonnegative elements.

Since any possible reaction (characterized by $\mathbf{n}^{\mathrm{T}}$, a row of $\mathbf{N}$) must necessarily satisfy the conservation equation for the N conserved quantities, it follows:

$$\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{n} = \mathbf{0}_N \quad \text{or} \quad \mathbf{n}^{\mathrm{T}}\boldsymbol{\Gamma} = \mathbf{0}_N^{\mathrm{T}} \tag{2.17}$$

and, for the R reactions:

$$\mathbf{N}\,\boldsymbol{\Gamma} = \mathbf{0}_{R\times N} \quad \text{or} \quad \boldsymbol{\Gamma}^{\mathrm{T}}\,\mathbf{N}^{\mathrm{T}} = \mathbf{0}_{N\times R}. \tag{2.18}$$

In mathematical terms, the reaction stoichiometries must lie in the null space of the *reaction-invariance matrix* $\boldsymbol{\Gamma}^{\mathrm{T}}$ (Waller and Mäkilä, 1981; Bonvin and Rippin, 1990). However, the reaction-invariance matrix only determines the *stoichiometric space* spanned by rows of $\mathbf{N}$ or $\mathbf{R}\,\mathbf{N}$, where $\mathbf{R}$ is an $R \times R$ matrix of full rank. In other words, the stoichiometric matrix can be determined from $\boldsymbol{\Gamma}^{\mathrm{T}}$ except for a rotation matrix. For this, it is sometimes said that $\boldsymbol{\Gamma}^{\mathrm{T}}$ determines the *theoretical stoichiometric space*. From the structure of $\boldsymbol{\Gamma}$, certain useful results will be derived.

Proposition 2.7 *The following properties result from* (2.18)*:*

(a) *The number of independent reactions obeys* $R_N < S$.

(b) $(S - R_N)$ *linearly-independent columns of* $\boldsymbol{\Gamma}$ *are required to span the complete null space of* $\mathbf{N}$*, and thus,* $N \geq S - R_N$.

(c) *For any* $\gamma \times S$ *concentration matrix* $\check{\mathbf{C}}$*, if* $\text{rank}\left(\check{\mathbf{C}}\,\boldsymbol{\Gamma}\right) = \gamma$*, then* $\text{rank}\left(\left[\begin{smallmatrix}\mathbf{N}\\ \check{\mathbf{C}}\end{smallmatrix}\right]\right) = R_N + \gamma$.

(d) *For any nonzero concentration vector* $\mathbf{c}$*,* $\mathbf{c} \notin \mathcal{S}_r(\mathbf{N})$*, where* $\mathcal{S}_r(\cdot)$ *denotes the subspace spanned by the rows of a matrix (row space).*

(See Appendix H.1 for proof)

2.2.3 Transformation to independent kinetics

Consider Situation S2 (with constant linearly-independent stoichiometries) where R_r elements of $\mathbf{r}$ are linearly independent on $[t_0, t_K]$. Thus,

$$\mathbf{r}(t) = \boldsymbol{\mathcal{P}}_{r2}\check{\mathbf{r}}(t), \tag{2.19}$$

where $\check{\mathbf{r}}(t)$ is a vector containing the R_r linearly-independent reaction rates on $[t_0, t_K]$, and $\boldsymbol{\mathcal{P}}_{r2}$ an $R \times R_r$ matrix of full rank.

2.2.4 Transformations to independent reactions

Consider the (general) case that describes Situations S1–S3. A transformation will be proposed that takes a system with *linearly-dependent* time-varying stoichiometries and/or reaction rates to a system with *linearly-independent* stoichiometries and reaction rates. This transformation uses the concept of *independent reactions* for the *reaction term* that expresses the reactant consumption and product formation. The reaction term is most easily understood for an S-component mixture exhibiting R reactions with constant temperature, density, and pressure and no inlet and outlet stream. It is generally assumed that the reaction system is homogeneous, i.e., mixing is ideal and the concentrations in the outlet(s) are equal to that inside the reactor. For such reaction systems, the dynamic behavior is completely described by the following molar balance equations (Perry and Green, 1984):

$$\dot{\mathbf{c}} = \mathbf{N}^{\mathrm{T}}\,\mathbf{r}(\mathbf{c}), \quad \mathbf{c}(0) = \mathbf{c}_0, \tag{2.20}$$

where $\mathbf{c}_0$ is the initial molar concentration vector and . $\mathbf{N}^{\mathrm{T}}\mathbf{r}(\mathbf{c})$ models the reaction term.

Definition 2.8 (Independent reactions) *A set of reactions is said to be independent if both the stoichiometries and the kinetics are linearly independent (on the time interval $[t_0, t_K]$), and the stoichiometries are constant.*

Two transformations to independent reactions are proposed: (T1) for time-varying $\mathbf{N}$ and (T2) for constant $\mathbf{N}$.

(T1) Let K be the number of time instants considered, and t_0 and t_K the initial and final time instants, respectively. Consider the reaction term $\mathbf{N}^{\mathrm{T}}(k)\,\mathbf{r}(k)$ for the kth time instant in the time interval $[t_0, t_K]$, and let the $S \times K$ matrix $\boldsymbol{\mathcal{F}}_R$ be defined as:

$$\boldsymbol{\mathcal{F}}_R \equiv [\mathbf{N}^{\mathrm{T}}(1)\,\mathbf{r}(1), \ldots, \mathbf{N}^{\mathrm{T}}(K)\,\mathbf{r}(K)]. \tag{2.21}$$

Let $R^\star \equiv \mathrm{rank}\,(\boldsymbol{\mathcal{F}}_R)$. Using for example singular value decomposition (SVD; see Appendix B.3), $\boldsymbol{\mathcal{F}}_R$ can be expressed as

$$\boldsymbol{\mathcal{F}}_R = \mathbf{N}^{\star\mathrm{T}}\,\boldsymbol{\mathcal{R}}_r, \tag{2.22}$$

where $\mathbf{N}^\star$ is the $R^\star \times S$ constant linearly-independent (abstract) stoichiometric matrix and $\boldsymbol{\mathcal{R}}_r$ the $R^\star \times K$ matrix of the corresponding linearly-independent

(abstract) reaction rates for the K time instants considered. The $R^\star$-dimensional reaction rate vector $\mathbf{r}^\star(k)$ at the kth time instant is the kth column of $\mathcal{R}_r$. Thus,

$$\mathbf{N}^{\mathrm{T}}(k)\,\mathbf{r}(k) = \mathbf{N}^{\star\mathrm{T}}\,\mathbf{r}^\star(k). \tag{2.23}$$

(T2) For constant $\mathbf{N}$, it is proposed to proceed as follows:

(a) Select the R_N rows of $\mathbf{N}$ that are linearly independent according to (2.16).

(b) Select the R_r elements of $\mathbf{r}$ that are linearly independent on the time interval $[t_0, t_K]$ according to (2.19).

(c) From (2.16) and (2.19), compute an $R_N \times R_r$ matrix defined as

$$\boldsymbol{\mathcal{P}}_r \equiv \boldsymbol{\mathcal{P}}_{r1}^{\mathrm{T}}\,\boldsymbol{\mathcal{P}}_{r2}. \tag{2.24}$$

(d) For the case $R_N \geq R_r$, construct the constant linearly-independent stoichiometric matrix $\mathbf{N}^\star = \boldsymbol{\mathcal{P}}_r^{\mathrm{T}}\,\check{\mathbf{N}}$, and the linearly-independent reaction rate vector $\mathbf{r}^\star(k) = \check{\mathbf{r}}(k)$, with $R^\star = R_r$. Otherwise, $\mathbf{N}^\star = \check{\mathbf{N}}$ and $\mathbf{r}^\star(k) = \boldsymbol{\mathcal{P}}_r\,\check{\mathbf{r}}(k)$ with $R^\star = R_N$. Thus, the number of independent reactions $R^\star$ is given by $\min(R_N, R_r)$. In summary, the following transformation is obtained:

$$\mathbf{N}^{\mathrm{T}}\,\mathbf{r}(k) = \check{\mathbf{N}}^{\mathrm{T}}\,\boldsymbol{\mathcal{P}}_r\,\check{\mathbf{r}} = \mathbf{N}^{\star\mathrm{T}}\,\mathbf{r}^\star(k). \tag{2.25}$$

The structure of (2.20) is conserved for time-varying linearly-dependent stoichiometries and reaction rates except for the following redefinitions: $\mathbf{N} := \mathbf{N}^\star$, $\mathbf{r} := \mathbf{r}^\star$, and $R := R^\star$. Thus, the assumption that the stoichiometries and the reaction rates are linearly independent and the stoichiometries are constant, is without loss of generality.

Examples will illustrate Transformation T2 for Situations S2 and S3.

Example 2.2a (Dependent stoichiometries I; cont'd from page 12**)**
Assume power-law kinetics, $\mathbf{r} = [\kappa_1 c_1 c_2 c_5,\ \kappa_2 c_3 c_4 c_5]^{\mathrm{T}}$. Since only one row of $\mathbf{N}$ is linearly independent (say, the first), $R_N = 1$ and

$$\mathbf{N} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} -1 & -1 & 1 & 1 & 0 \end{bmatrix} = \boldsymbol{\mathcal{P}}_{r1}\,\check{\mathbf{N}}.$$

Since the reaction rates are linearly independent, it follows that $R_r = R = 2$, $\boldsymbol{\mathcal{P}}_{r2} = \mathbf{I}_R$, and $\boldsymbol{\mathcal{P}}_r = \boldsymbol{\mathcal{P}}_{r1}^{\mathrm{T}}$, where $\mathbf{I}_R$ is the R-dimensional identity matrix. Furthermore, since $R_N < R_r$, it follows that $R^\star = R_N = 1$ and

$$\mathbf{N}^\star = \check{\mathbf{N}} = \begin{bmatrix} -1 & -1 & 1 & 1 & 0 \end{bmatrix}, \qquad \mathbf{r}^\star = \boldsymbol{\mathcal{P}}_r\,\mathbf{r} = r_1 - r_2.$$

Example 2.3a (Dependent stoichiometries II; cont'd from page 12**)**
Assume power-law kinetics, $\mathbf{r} = [\kappa_1 c_1,\ \kappa_2 c_2,\ \kappa_3 c_3]^{\mathrm{T}}$. Since $\mathrm{rank}\,(\mathbf{N}) = 2$, it is sufficient to select two linearly-independent stoichiometries, e.g., the stoichiometries of the first two reactions. Thus, $R_N = 2$ and

$$\mathbf{N} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \boldsymbol{\mathcal{P}}_{r1}\,\check{\mathbf{N}}.$$

Since the reaction rates are linearly independent, it follows that $R_r = R = 3$, $\mathcal{P}_{r2} = \mathbf{I}_R$, and $\mathcal{P}_r = \mathcal{P}_{r1}^{\mathrm{T}}$. Furthermore, since $R_N < R_r$, it follows that $R^\star = R_N = 2$ and

$$\mathbf{N}^\star = \check{\mathbf{N}} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}, \qquad \mathbf{r}^\star = \mathcal{P}_r\,\mathbf{r} = \begin{bmatrix} r_1 - r_3 \\ r_2 - r_3 \end{bmatrix}.$$

The physical reaction pathway $\mathcal{X}_1 \to \mathcal{X}_2 \to \mathcal{X}_3 \to \mathcal{X}_1$ is reduced to $\mathcal{X}_1 \to \mathcal{X}_2 \to \mathcal{X}_3$, where the third (physical) reaction is accounted for in the reaction rates of the two independent reactions.

Example 2.4a (Dependent kinetics; cont'd from page 12**)**
Consider the irreversible parallel reaction system (2.8). Assume power-law kinetics for the second reaction, $r_2 = \kappa_2 c_1^2$. Furthermore, let us assume that and the first kinetic rate being linearly dependent on the second one for $c_2 > 0$, or zero otherwise:

$$r_1 = \begin{cases} \alpha\, r_2, & \forall\, c_2 > 0 \\ 0, & \text{otherwise.} \end{cases}$$

This type of kinetics is called *discontinuous and linearly-dependent kinetics*. Since only one element of $\mathbf{r}$ is linearly independent (say, the second), $R_r = 1$ and according to (2.19):

$$\mathbf{r} = \begin{bmatrix} \alpha \\ 1 \end{bmatrix} r_2 = \mathcal{P}_{r2}\,\check{\mathbf{r}}.$$

Two cases can be distinguished:

(a) $c_2 = 0$: Since only the second reaction is active, $\mathbf{N}^\star = [-2,\ 0,\ 0,\ 1]$ and $\mathbf{r}^\star = r_2$.

(b) $c_2 > 0$: Since the stoichiometries are linearly independent, it follows that $R_N = R = 2$, $\check{\mathbf{N}} = \mathbf{N}$, $\mathcal{P}_{r1} = \mathbf{I}_R$, and $\mathcal{P}_r = \mathcal{P}_{r2}$. Furthermore, since $R_N > R_r$, $R^\star = R_r = 1$ and

$$\mathbf{N}^\star = \mathcal{P}_r^{\mathrm{T}}\check{\mathbf{N}} = \mathcal{P}_r^{\mathrm{T}}\mathbf{N} = \begin{bmatrix} -(2+\alpha) & -\alpha & \alpha & 1 \end{bmatrix}, \qquad \mathbf{r}^\star = \check{\mathbf{r}} = r_2.$$

From $\mathbf{N}^\star$, the following non-physical pathway is obtained: $(2+\alpha)\,\mathcal{X}_1 + \alpha\,\mathcal{X}_2 \to \alpha\,\mathcal{X}_3 + \mathcal{X}_4$. This is equivalent to (2.8).

Example 2.9 (Dependent kinetics and stoichiometries)
Consider the following reaction network:

$$\mathcal{X}_1 \rightleftharpoons \mathcal{X}_2 \to \mathcal{X}_3 \to \mathcal{X}_4 \qquad (2.26) \qquad \text{with} \qquad \mathbf{N} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

Assume the kinetic rates of the last two reactions to be linearly dependent on the first two: $\mathbf{r} = [r_1,\ r_2,\ \alpha\, r_1,\ \beta\, r_2]^{\mathrm{T}}$ with α and β being positive coefficients. Since only the first, third, and fourth stoichiometry are linearly independent, $R_N = 3$ and according to (2.16):

$$\mathbf{N} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} = \mathcal{P}_{r1}\,\check{\mathbf{N}}.$$

Also, since only two elements of $\mathbf{r}$ are linearly independent (say, the first two), $R_r = 2$ and according to (2.19):

$$\mathbf{r} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \alpha & 0 \\ 0 & \beta \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \mathcal{P}_{r2}\,\check{\mathbf{r}}.$$

Thus,

$$\mathcal{P}_r = \mathcal{P}_{r1}^{\mathrm{T}} \mathcal{P}_{r2} = \begin{bmatrix} 1 & -1 \\ \alpha & 0 \\ 0 & \beta \end{bmatrix}.$$

Since $R_r < R_N$, $R^\star = R_r = 2$ and

$$\mathbf{N}^\star = \mathcal{P}_r^{\mathrm{T}} \check{\mathbf{N}} = \begin{bmatrix} -1 & (1-\alpha) & \alpha & 0 \\ 1 & -1 & -\beta & \beta \end{bmatrix}, \qquad \mathbf{r}^\star = \check{\mathbf{r}} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}.$$

From $\mathbf{N}^\star$, the following non-physical reaction pathway is obtained:

$$\begin{aligned} \mathcal{X}_1 + \alpha\,\mathcal{X}_2 &\rightarrow \mathcal{X}_2 + \alpha\,\mathcal{X}_3 \\ \mathcal{X}_2 + \beta\,\mathcal{X}_3 &\rightarrow \mathcal{X}_1 + \beta\,\mathcal{X}_4. \end{aligned}$$

This is equivalent to (2.26).

2.3 Dynamic models for reaction systems with inlet and outlet streams

In Section 2.2, a dynamic model for isothermal, isobaric, constant-density homogeneous reaction systems with no inlet and outlet stream was derived. Here, the assumptions of constant temperature and density and of no inlet and outlet stream are removed. It is solely assumed that the pressure-dependency of the physical properties (e.g., density, heat capacity) is negligible. The dynamic behavior of an S-component mixture will be described by: (i) S mole balance equations for $\mathbf{c}$, (ii) one continuity equation for V, (iii) one heat balance equation for T, and (iv) the corresponding initial conditions. Figure 2.2 illustrates a reactor with the different terms needed to build the balance equations.

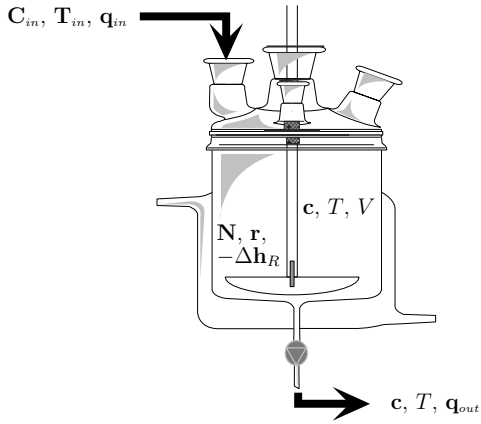

Figure 2.2. Reaction system with p inlet and q outlet streams.

Without loss of generality, it is assumed that (1) the reaction term is modeled in terms of R *independent* reactions, (2) only one outlet stream is present, and (3) both the inlet concentrations and the inlet volumetric flowrates are linearly independent, and the inlet concentrations are constant. The first assumption was studied in Subsection 2.2, and the second and third assumptions will be discussed in Subsections 2.3.1 and Subsection 2.4.2, respectively.

2.3.1 Mole balance equations

The mole balance equations for a homogeneous reaction system with p inlet and q outlet streams, S species, and R independent reactions are given by:

$$\frac{d(V\,\mathbf{c})}{dt} = \mathbf{N}^{\mathrm{T}}\,\mathbf{r}(\mathbf{c},T)\,V + \mathbf{C}_{in}\,\mathbf{q}_{in} - q_{out}\,\mathbf{c}, \quad (V\,\mathbf{c})(0) = V_0\mathbf{c}_0, \qquad \mathbf{C}_{in} \equiv [\mathbf{c}_{in}^1, \mathbf{c}_{in}^2, \cdots, \mathbf{c}_{in}^p], \tag{2.27}$$

where $\mathbf{C}_{in}$ is the $S \times p$ inlet molar concentration matrix, $\mathbf{c}_{in}^i$ the molar concentrations of the ith inlet stream, $\mathbf{q}_{in}$ the p-dimensional vector of volumetric flowrates, and V_0 the initial volume.

Note that $\mathbf{r}$ is now also a function of T. It can be modeled by the Arrhenius equation:

$$\kappa_i = \kappa_{i0}\,\mathrm{e}^{-E/T}, \quad E = \frac{E^\star}{\mathsf{R}},$$

where κ_{i0} is the frequency factor, $E^\star$ the activation energy, and E the activation energy normalized by the gas constant, R.

Owing to the assumption of homogeneity,

$$q_{out} = \mathbf{1}_q^{\mathrm{T}}\mathbf{q}_{out}, \tag{2.28}$$

where $\mathbf{1}_q$ is an q-dimensional vector with all elements 1.

The last two terms in (2.27) indicate the reactants being supplied to and the products/reactants removed from the system (the so-called *material exchange terms*).

2.3.2 Continuity equation

The continuity equation is given by:

$$\dot{\mathfrak{m}} = \mathbf{1}_p^{\mathrm{T}}\,\boldsymbol{\nu}_{in} - \nu_{out}, \quad \mathfrak{m}(0) = \mathfrak{m}_0, \tag{2.29}$$

where $\mathfrak{m}$ is the total mass, $\boldsymbol{\nu}_{in}$ and ν_{out} the p-dimensional inlet mass flowrate vector and outlet mass flowrate, respectively, and $\mathfrak{m}_0$ the initial total mass.

Using the density, ρ, equation (2.29) can be rewritten as:

$$\frac{d}{dt}(\rho\,V) = \rho\,\dot{V} + V\,\dot{\rho} = \mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Phi}_{in}\,\mathbf{q}_{in} - \rho\,q_{out}, \tag{2.30}$$

which gives:

$$\dot{V} = \frac{1}{\rho}\,\mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Phi}_{in}\,\mathbf{q}_{in} - q_{out} - \frac{V}{\rho}\,\dot{\rho}, \qquad V(0) = V_0 \tag{2.31}$$

with $\boldsymbol{\Phi}_{in} \equiv \mathrm{diag}\,(\rho_{in}^1, \ldots, \rho_{in}^p)$, ρ_{in}^i the density of the ith inlet stream, and V_0 the initial volume.

Let ρ depend on $\mathbf{c}$ and T, $\boldsymbol{\rho}_c \equiv \partial\rho/\partial\mathbf{c}$ be an S-dimensional vector, and $\rho_T \equiv \partial\rho/\partial T$. With $\dot{\rho} = (\boldsymbol{\rho}_c^{\mathrm{T}}\,\dot{\mathbf{c}} + \rho_T\,\dot{T})$ and $\dot{\mathbf{c}}$ calculated from (2.27) as:

$$\dot{\mathbf{c}} = \mathbf{N}^{\mathrm{T}}\,\mathbf{r}(\mathbf{c}, T) + \frac{1}{V}\,\mathbf{C}_{in}\,\mathbf{q}_{in} - \frac{q_{out}}{V}\,\mathbf{c} - \frac{\dot{V}}{V}\,\mathbf{c} \tag{2.32}$$

equation (2.31) becomes:

$$\dot{V} = \frac{1}{\rho}\,\mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Phi}_{in}\,\mathbf{q}_{in} - q_{out} - \frac{1}{\rho}\,\boldsymbol{\rho}_c^{\mathrm{T}}\,\left(V\,\mathbf{N}^{\mathrm{T}}\,\mathbf{r} + \mathbf{C}_{in}\,\mathbf{q}_{in} - q_{out}\,\mathbf{c} - \dot{V}\,\mathbf{c}\right) - \frac{V}{\rho}\,\rho_T\,\dot{T}. \tag{2.33}$$

Solving (2.33) for $\dot{V}$ gives with $V(0) = V_0$:

$$\dot{V} = \frac{1}{\rho - \boldsymbol{\rho}_c^{\mathrm{T}}\,\mathbf{c}}\left[\left(\mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Phi}_{in} - \boldsymbol{\rho}_c^{\mathrm{T}}\,\mathbf{C}_{in}\right)\,\mathbf{q}_{in} + (\boldsymbol{\rho}_c^{\mathrm{T}}\,\mathbf{c} - \rho)\,q_{out} - V\,\boldsymbol{\rho}_c^{\mathrm{T}}\,\mathbf{N}^{\mathrm{T}}\,\mathbf{r} - V\,\rho_T\,\dot{T}\right]. \tag{2.34}$$

2.3.3 Heat balance equation

Under the assumptions that the pressure-dependency of the physical properties and the mixing enthalpies are negligible, the heat balance equation reads:

$$\frac{d\,(\mathfrak{m}\,\mathfrak{c}_p\,T)}{dt} = V\,(-\Delta\mathbf{h}_R^{\mathrm{T}})\,\mathbf{r}(\mathbf{c}, T) + \mathbf{T}_{in}^{\mathrm{T}}\,\boldsymbol{\mathfrak{C}}_{p,in}\,\boldsymbol{\nu}_{in} - T\,\mathfrak{c}_p\,\nu_{out} + \dot{Q}_{ext}, \quad T(0) = T_0, \tag{2.35}$$

where $\mathfrak{c}_p$ is the specific heat capacity, $\Delta\mathbf{h}_R$ the R-dimensional vector of reaction enthalpies, $\mathbf{T}_{in}$ the p-dimensional inlet temperature vector, $\boldsymbol{\mathfrak{C}}_{p,in}$ the p-dimensional diagonal matrix with elements being specific heat capacities of the inlet streams, $\dot{Q}_{ext}$ the external heat power (see Remark 2.10 below), and T_0 the initial temperature.

The structures of (2.27) and (2.35) are similar except for the additional term $\dot{Q}_{ext}$ in (2.35). Also, the accumulation terms of (2.27) and (2.35) are somewhat different, since the total mass $\mathfrak{m}$ appears in (2.35), while the volume appears in (2.27). However, by introducing the auxiliary variables $T^{\dagger} \equiv \mathfrak{c}_p\,\rho\,T$ and $\mathbf{T}_{in}^{\dagger} \equiv \boldsymbol{\mathfrak{C}}_{p,in}\,\boldsymbol{\Phi}_{in}\,\mathbf{T}_{in}$, (2.35) can be brought into a form where the volume appears in the accumulation term:

$$\frac{d\,(V\,T^{\dagger})}{dt} = V\,(-\Delta\mathbf{h}_R)^{\mathrm{T}}\,\mathbf{r}(\mathbf{c}, T) + \mathbf{T}_{in}^{\dagger\mathrm{T}}\,\mathbf{q}_{in} - T^{\dagger}\,q_{out} + \dot{Q}_{ext}, \quad T^{\dagger}(0) = T_0^{\dagger}. \tag{2.36}$$

Remark 2.10 (Modeling of $\dot{Q}_{ext}$)
The main sources of external heat power $\dot{Q}_{ext}$ are external heating/cooling power, $\dot{Q}_{co}$, and dissipated stirrer power, $\dot{Q}_{sti}$, i.e., $\dot{Q}_{ext} = \dot{Q}_{co} + \dot{Q}_{sti}$. For both terms, models are proposed below.

In practical applications, heating or cooling power is transferred to the reaction system using, for example, a jacket. Thus, by assuming constant heat capacity for the jacket

flow stream, $\mathsf{c}_{p,co}$, the heat balance equation for the jacket can be derived as (Perry and Green, 1984):

$$\mathsf{m}_{co}\,\mathsf{c}_{p,co}\,\frac{dT_{co}}{dt} = \nu_{co}\,\mathsf{c}_{p,co}\,(T_{co,in} - T_{co,out}) + \dot{Q}_{co}, \quad \dot{Q}_{co} = \mathsf{U}\,\mathsf{A}\,(T - T_{co}),$$

where m_{co} is the total mass of the heating/cooling liquid in the jacket, T_{co}, $T_{co,in}$, and $T_{co,out}$ the average, inlet, and outlet temperature of the heating/cooling liquid, respectively, ν_{co} the mass flowrate of the heating/cooling liquid, U the overall heat transfer coefficient through the wall, and A the heat-transfer area. Note that for the control of the reactor temperature T, ν_{co} is commonly chosen as the manipulated variable. Also, an additional relationship between T_{co}, $T_{co,in}$, and $T_{co,out}$. In practice, it is often assumed that $T_{co,out} = T_{co}$.

If the residence time m_{co}/ν_{co} in the jacket is small with respect to the smallest time constant of the reaction system, the dynamic of the jacket can be neglected by using the concept of singular perturbation (Keener, 1988) and, thus:

$$\dot{Q}_{co} = \nu_{co}\,\mathsf{c}_{p,co}\,(T_{co,out} - T_{co,in})\,. \tag{2.37}$$

Since the quantities on the right side of (2.37) are often measured, $\dot{Q}_{co}$ can be estimated. This is contrast to $\dot{Q}_{co} = \mathsf{U}\,\mathsf{A}\,(T - T_{co})$ where U and A are usually difficult to estimate (and might vary during the course of the reaction).

The dissipated stirrer power $\dot{Q}_{sti}$ can be modeled as:

$$\dot{Q}_{sti} = \mathsf{M}\,\omega,$$

where M is the stirrer torque and ω the angular speed. Both M and ω can be measured and, thus, $\dot{Q}_{sti}$ estimated.

2.3.4 Overview of dynamic models

The (basic) model for reaction systems with constant density and temperature is given in Table 2.2. Its extensions to varying density and temperature are also provided. For the three models, Assumptions A1–3 are verified (See Appendix D).

Let $\boldsymbol{n}$ be the S-dimensional vector of the number of moles with $\boldsymbol{n} = V\,\mathbf{c}$. For the first two models in Table 2.2, $\boldsymbol{n}$ and V are considered the $(S+1)$ states and $\mathbf{q}_{in}$ and q_{out} the $(p+1)$ *linearly-independent* inputs.

For the model with varying temperature, the following quantities are defined: $\boldsymbol{n}_T \equiv \left[{}_{(V}{}^{\boldsymbol{n}}{}_{T^\dagger)}\right]$ $(S{+}1{\times}1)$, $\mathbf{c}_T \equiv \left[{}^{\mathbf{c}}_{T^\dagger}\right]$, $(S{+}1{\times}1)$, $\mathbf{N}_T \equiv [\mathbf{N},\, -\Delta\mathbf{h}_R]$ $(R{\times}S{+}1)$, $\mathbf{C}_{T,in} = \begin{bmatrix} \mathbf{C}_{in} & \mathbf{0}_S \\ \mathbf{T}_{in}^{\dagger_T} & 1 \end{bmatrix}$ $(S+1\times p+1)$, and $\mathbf{q}_{Tin} = \left[{}^{\mathbf{q}_{in}}_{\dot{Q}_{ext}}\right]$ $(p+1\times 1)$. (2.40) shows that the external heat flowrate $\dot{Q}_{ext}$ is considered as an additional inlet stream ($p_T \equiv p+1$) and the modified temperature ($VT^\dagger$) is modeled as an additional species ($S_T = S+1$). Thus, such a system includes (S_T+1) states and (p_T+1) inputs.

In the case of time-varying reaction enthalpies that are linearly dependent on the stoichiometries, the number of independent reactions $R^\star$ is determined by Transformation T1 in Subsection 2.2.4 with the redefinition $\mathbf{N}(k) := \mathbf{N}_T(k)$, where $\mathbf{N}_T(k) \equiv$

Table 2.2. The basic model for reaction systems with constant density and temperature and its extensions to varying density and temperature. Second column: assumptions in addition to Assumptions A1–3; third column: derived from equations; fourth column: model equations.

Model	Ass.	From	Equations
Basic	$\rho = \not{c}$, $T = \not{c}$	(2.27), (2.34)	$\dot{\boldsymbol{n}} = \frac{d(V\,\mathbf{c})}{dt} = V\,\mathbf{N}^{\mathrm{T}}\,\mathbf{r}(\mathbf{c}) + \mathbf{C}_{in}\,\mathbf{q}_{in} - q_{out}\,\mathbf{c}, \quad \boldsymbol{n}(0) = V_0\mathbf{c}_0 = \boldsymbol{n}_0$ $\dot{V} = (\mathbf{1}_p^{\mathrm{T}}\mathbf{q}_{in}) - q_{out}, \qquad V(0) = V_0$ (2.38)
Varying-density	$T = \not{c}$	(2.27), (2.34)	$\dot{\boldsymbol{n}} = \frac{d(V\,\mathbf{c})}{dt} = V\,\mathbf{N}^{\mathrm{T}}\,\mathbf{r}(\mathbf{c}) + \mathbf{C}_{in}\,\mathbf{q}_{in} - q_{out}\,\mathbf{c}, \qquad \boldsymbol{n}(0) = V_0\mathbf{c}_0 = \boldsymbol{n}_0$ $\dot{V} = \frac{1}{\rho - \boldsymbol{\rho}_c^{\mathrm{T}}\,\mathbf{c}}\left[\left(\mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Phi}_{in} - \boldsymbol{\rho}_c^{\mathrm{T}}\,\mathbf{C}_{in}\right)\,\mathbf{q}_{in} + (\boldsymbol{\rho}_c^{\mathrm{T}}\,\mathbf{c} - \rho)\,q_{out} - V\,\boldsymbol{\rho}_c^{\mathrm{T}}\,\mathbf{N}^{\mathrm{T}}\,\mathbf{r}\right], \quad V(0) = V_0$ (2.39)
Noniso-thermal	$\rho = \not{c}$	(2.27), (2.34), (2.35)	$\dot{\boldsymbol{n}}_T = \frac{d(V\,\mathbf{c}_T)}{dt} = V\,\mathbf{N}_T^{\mathrm{T}}\,\mathbf{r}(\mathbf{c}_T, T) + \mathbf{C}_{T,in}\,\mathbf{q}_{T,in} - q_{out}\,\mathbf{c}_T, \;\; \boldsymbol{n}_T(0) = V_0\mathbf{c}_{T,0} = \boldsymbol{n}_{T,0}$ $\dot{V} = (\mathbf{1}_p^{\mathrm{T}}\mathbf{q}_{in}) - q_{out}, \qquad V(0) = V_0$ (2.40)

$[\mathbf{N}(k),\ -\Delta\mathbf{h}_R(k)]$, and $\Delta\mathbf{h}_R(k)$ the R-dimensional linearly-dependent time-varying reaction enthalpy vector.

In the case of time-varying inlet concentrations, modified temperature, and a linearly-dependent external heat flowrate, the number of independent inlets p_T is determined by the transformation presented in Subsection 2.4.2 except for $\mathbf{C}_{in}(k) := \mathbf{C}_{T,in}(k)$, $\mathbf{q}_{in}(k) := \mathbf{q}_{T,in}(k)$, and $p := p_T$.

2.4 Extensions

In this section, the following extensions to the basic model (see 2.38) are provided: (i) expression of the mole balance equations in weight fractions, (ii) varying inlet concentrations (independent inlets), and (iii) impulse additions and multiple process runs (generalized inlets)

2.4.1 Reaction systems expressed in weight fractions

When the density varies, it is often useful to express the dynamic model in terms of weight fractions, w_j, $(i = 1, \dots, S)$ defined as:

$$w_j = \frac{\mathfrak{m}_j}{\mathfrak{m}}, \quad \forall j = 1, \dots, S, \qquad \mathbf{1}_S^{\mathrm{T}}\mathbf{w} = 1, \tag{2.41}$$

where $\mathfrak{m}_j$ is the mass of the jth species and $\mathbf{w}$ the S-dimensional weight fraction vector. Note that the weight fractions are closed, since $\mathbf{1}_S^{\mathrm{T}}\mathbf{w}(k) = 1$. The unit is usually given in [wt.-%].

Model (2.38) can be rewritten in terms of weight fractions. By substituting: (i) $V\,\mathbf{c} = \mathbf{M}_w^{-1}\,\boldsymbol{m}$, (ii) $\mathbf{C}_{in}\,\mathbf{q}_{in} = \mathbf{M}_w^{-1}\,\mathbf{W}_{in}\boldsymbol{\nu}_{in}$, and (iii) $q_{out}\,\mathbf{c} = \nu_{out}\,\mathbf{M}_w^{-1}\,\mathbf{w}$ into equation (2.38):

$$\begin{aligned}
\dot{\boldsymbol{m}} = \frac{d(\mathfrak{m}\,\mathbf{w})}{dt} &= \frac{\mathfrak{m}}{\rho(\mathbf{w})}\,\mathbf{N}_w^{\mathrm{T}}\,\mathbf{r}_{\mathfrak{m}}(\mathbf{w},\,\mathfrak{m}) + \mathbf{W}_{in}\,\boldsymbol{\nu}_{in} - \nu_{out}\,\mathbf{w},\ \boldsymbol{m}(0) = \mathfrak{m}_0\,\mathbf{w}_0 = \boldsymbol{m}_0 \\
\dot{\mathfrak{m}} &= \mathbf{1}_p^{\mathrm{T}}\boldsymbol{\nu}_{in} - \nu_{out}, \qquad\qquad \mathfrak{m}(0) = \mathfrak{m}_0, \\
\text{with}\quad \mathbf{N}_w &= \mathbf{N}\,\mathbf{M}_w,\ \mathbf{W}_{in} = \left[\mathbf{w}_{in}^1, \cdots, \mathbf{w}_{in}^p\right],
\end{aligned} \tag{2.42}$$

where $\boldsymbol{m}$ is an S-dimensional vector indicating the masses of the various species, $\mathbf{r}_{\mathfrak{m}}$ the R-dimensional reaction rate vector expressed as a function of $\mathbf{w}$ and $\mathfrak{m}$, $\mathbf{w}_{in}^i$ the weight fractions of the ith inlet stream (without loss of generality, assumed to be constant), $\boldsymbol{\nu}_{in}$ the p-dimensional inlet mass flowrate vector, ν_{out} the outlet mass flowrate, $\mathbf{M}_w = \mathrm{diag}\,(M_{w,1}, \dots, M_{w,S})$, $M_{w,i}$ the molecular weight of the ith species, and $\rho(\mathbf{w})$ the density expressed as a function of $\mathbf{w}$.

2.4.2 Varying inlet concentrations

Practical situations with time-varying inlet concentrations are listed below.

(1) Variability in the raw materials or in the preprocessing steps,

(2) Gaseous outlet from liquid reactions. In this case, the corresponding outlet flowrate $q_{in}(t)$ is negative.

Similarly to the transformation in Section 2.2, a transformation will be proposed that takes the system with *linearly-dependent* inlet streams of *varying* inlet concentrations to a system with *linearly-independent* inlet streams of *constant* inlet concentrations. This transformation uses the concept of *independent inlet streams*.

Definition 2.11 (Independent inlet streams) *A set of inlet streams is said to be independent if (i) both the inlet concentrations and the corresponding flowrates are linearly independent (on the time interval $[t_0, t_K]$), and (ii) the inlet concentrations are constant.*

Consider the inlet term $\mathbf{C}_{in}(k)\,\mathbf{q}_{in}(k)$ and $\mathbf{1}_p^{\mathrm{T}}\mathbf{q}_{in}(k)$ at the kth time instant in the time interval $[t_0, t_K]$, and let the $S \times K$ matrix $\boldsymbol{\mathcal{F}}_{in}$, and the $p \times K$ matrix $\boldsymbol{\mathcal{Q}}_{in}$ be defined as:

$$\boldsymbol{\mathcal{F}}_{in} \equiv [\mathbf{C}_{in}(1)\,\mathbf{q}_{in}(1), \ldots, \mathbf{C}_{in}(K)\,\mathbf{q}_{in}(K)], \qquad \boldsymbol{\mathcal{Q}}_{in} \equiv [\mathbf{q}_{in}(1), \ldots, \mathbf{q}_{in}(K)]. \tag{2.43}$$

Let $\alpha \equiv \mathrm{rank}\,(\boldsymbol{\mathcal{F}}_{in})$. $\boldsymbol{\mathcal{F}}_{in}$ can be decomposed using, for example, SVD such that

$$\boldsymbol{\mathcal{F}}_{in} = \bar{\mathbf{C}}_{in}\,\bar{\boldsymbol{\mathcal{Q}}}_{in}, \tag{2.44}$$

where $\bar{\mathbf{C}}_{in}$ and $\bar{\boldsymbol{\mathcal{Q}}}_{in}$ are the $S \times \alpha$ matrix of constant (abstract) inlet concentrations and the $\alpha \times K$ matrix of corresponding (abstract) inlet volumetric flowrates for the K time instants considered, respectively. α represents the number of species that can be independently manipulated by the inlet streams. The α-dimensional inlet volumetric flowrate vector $\bar{\mathbf{q}}_{in}(k)$ at the kth time instant is the kth column of $\bar{\boldsymbol{\mathcal{Q}}}_{in}$ $(k = 1, \ldots, K)$.

The effect of time-varying inlet concentrations on the volume is eliminated by an additional inlet stream $q_{in,aug}$:

$$\begin{aligned} \dot{V} &= \mathbf{1}_p^{\mathrm{T}}\mathbf{q}_{in} - q_{out} = \mathbf{1}_\alpha^{\mathrm{T}}\bar{\mathbf{q}}_{in} - \mathbf{1}_\alpha^{\mathrm{T}}\bar{\mathbf{q}}_{in} + \mathbf{1}_p^{\mathrm{T}}\mathbf{q}_{in} - q_{out} \\ &= \mathbf{1}_{\alpha+1}^{\mathrm{T}}\,\mathbf{q}_{in}^{\star} - q_{out} \\ \mathbf{q}_{in}^{\star} &\equiv \begin{bmatrix} \bar{\mathbf{q}}_{in} \\ q_{in,aug} \end{bmatrix}, \qquad q_{in,aug} \equiv \mathbf{1}_p^{\mathrm{T}}\mathbf{q}_{in} - \mathbf{1}_\alpha^{\mathrm{T}}\bar{\mathbf{q}}_{in}, \end{aligned} \tag{2.45}$$

where $\mathbf{q}_{in}^{\star}$ is the $(\alpha+1)$-dimensional inlet volumetric flowrate vector. For $q_{in,aug} = 0$, which is the case when $\mathbf{C}_{in}$ is constant, the addition of one $(\alpha+1)$ does not apply.

The $S \times (\alpha+1)$ constant inlet concentration matrix $\mathbf{C}_{in}^{\star}$ is defined as:

$$\mathbf{C}_{in}^{\star} \equiv \begin{bmatrix} \bar{\mathbf{C}}_{in} & \mathbf{0}_S \end{bmatrix}. \tag{2.46}$$

With definition (2.46):

$$\mathbf{C}_{in}(k)\,\mathbf{q}_{in}(k) = \bar{\mathbf{C}}_{in}\,\bar{\mathbf{q}}_{in}(k) = \mathbf{C}_{in}^{\star}\,\mathbf{q}_{in}^{\star}(k), \qquad \mathbf{1}_p^{\mathrm{T}}\mathbf{q}_{in}(k) = \mathbf{1}_{\alpha+1}^{\mathrm{T}}\mathbf{q}_{in}^{\star}(k). \tag{2.47}$$

The structure of (2.38) is conserved for varying inlet concentrations with the following redefinitions: $\mathbf{C}_{in} := \mathbf{C}_{in}^{\star}$, $\mathbf{q}_{in} := \mathbf{q}_{in}^{\star}$, and $p := \alpha+1$. Thus, the assumption that the inlet

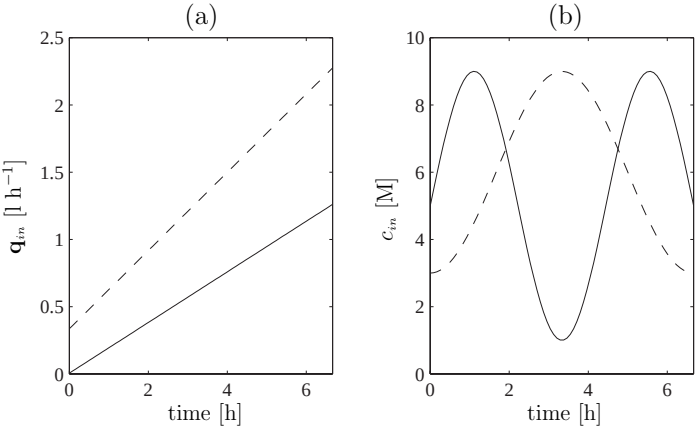

Figure 2.3. Example 2.4b: (a) flowrates $\mathbf{q}_{in}$ of the two inlets and (b) time-varying concentrations in Case C2. Legend for (a): (–) q_{in}^1, (– –) q_{in}^2. Legend for (b): (–) c_{in}^1, (– –) c_{in}^2.

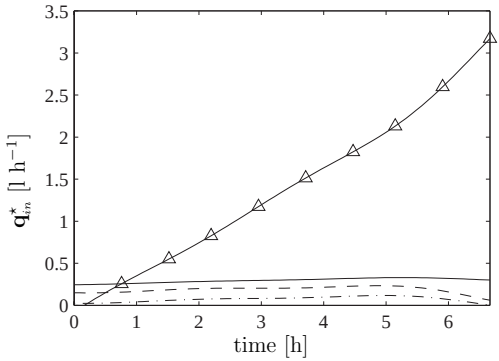

Figure 2.4. Example 2.4b: flowrates of the independent inlet streams $\mathbf{q}_{in}^{\star}$ for Case C2. Legend: (–) $q_{in}^{\star,1}$, (– –) $q_{in}^{\star,2}$, (–.) $q_{in}^{\star,3}$, (△) $q_{in,aug}$.

concentrations and the flowrates are linearly independent and the inlet concentrations are constant, is without loss of generality.

Since the decomposition in (2.44) is non-unique, non-negativity constraints can be imposed in the decomposition procedure to guarantee nonnegative entries in both $\bar{\mathbf{C}}_{in}$ and $\bar{\mathbf{\Omega}}_{in}$.

Example 2.4b (Varying inlet concentrations; cont'd from page 12**)**
Assume a constant-density semibatch reaction system (2.8) or (2.9) with $p = 2$ inlet streams. Figure 2.3a shows the corresponding inlet flowrates $\mathbf{q}_{in}$. Two cases are considered for the 4×2 inlet concentration matrix $\mathbf{C}_{in}$ [M]:

(a) Constant $\mathbf{C}_{in}$: $\mathbf{C}_{in,C1} = \begin{bmatrix} 8 & 3 \\ 3 & 7 \\ 2 & 6 \\ 1 & 0 \end{bmatrix}$,

(b) Case C1 but with varying concentrations of $\mathcal{X}_1$ in both inlets: $\mathbf{C}_{in,C2}(t) = \begin{bmatrix} c_{in}^1(t) & c_{in}^2(t) \\ 3 & 7 \\ 2 & 6 \\ 1 & 0 \end{bmatrix}$.

The concentration profiles of the time-varying inlet concentrations are given in Figure 2.3b.
For Case C1, $\alpha = 2 = p$, since $\mathbf{C}_{in,C1}$ is constant and linearly independent, and the flowrates are linearly independent. Thus, $q_{in,aug,C1} = \mathbf{1}_2^{\mathrm{T}}\mathbf{q}_{in,C1} - \mathbf{1}_2^{\mathrm{T}}\bar{\mathbf{q}}_{in,C1} = 0$, and the addition of one does not apply ($p := \alpha$).
For Case C2, $\alpha = 3$. Since α is a function of the number of species with time-varying inlet concentrations, α in Case C2 is the same in either of the two scenarios: time-varying inlet concentration of Species $\mathcal{X}_1$ varied in a single (not shown) or both inlet streams. Furthermore, since $q_{in,aug,C2} = \mathbf{1}_2^{\mathrm{T}}\mathbf{q}_{in,C2} - \mathbf{1}_3^{\mathrm{T}}\bar{\mathbf{q}}_{in,C2} \neq 0$, the number of independent inlet streams is $p := \alpha + 1 = 4$. Figure 2.4 shows the corresponding flowrates $\mathbf{q}_{in}^{\star}(k)$. Non-negativity constraints are imposed on (2.46) to guarantee nonnegative entries in both $\mathbf{C}_{in}^{\star}$ and $\mathbf{\Omega}_{in}^{\star}$. The corresponding matrix of independent inlet concentrations is obtained as:

$$\mathbf{C}_{in,C2}^{\star} = \begin{bmatrix} 118.53 & 149.20 & 24.50 & 0 \\ 73.11 & 105.2 & 9.56 & 0 \\ 54.50 & 89.94 & 5.84 & 0 \\ 14.30 & 0.50 & 4.11 & 0 \end{bmatrix}$$

2.4.3 Generalized inlets and outlets

In practical applications, three *variation modes* for a reaction system can usually be found:

(V1) inlet streams with each containing a particular or a mixture of species,
(V2) impulse additions with each containing a particular or a mixture of species,
(V3) multiple batch runs with different initial conditions.

All three variation modes can be cast in the framework of *generalized inlets and outlets*, as it will be shown in the following proposition. The proposition uses the principle that every initial condition can be considered as an impulse (Chen, 1984).

Proposition 2.12 *Variation modes V2–V3 for a reaction system can be cast in the framework of generalized inlets and outlets by properly redefining* $\mathbf{C}_{in}$, $\mathbf{q}_{in}$, *and* q_{out}.

Proof:
The two variation modes are analyzed separately, and redefinitions for $\mathbf{C}_{in}$, $\mathbf{q}_{in}$, and q_{out} are provided. For simplicity of notation, it will be assumed that all runs are in batch mode, and the initial concentrations $\mathbf{c}_0^1$ of the first batch run are the initial conditions for the reaction system with generalized inlets and outlets:

$$\mathbf{c}_0 := \mathbf{c}_0^1. \tag{2.48}$$

(V2) Impulse additions during the course of a reaction can be considered as a special case of inlet streams. Thus, the jth impulse addition with concentrations $\mathbf{c}_\Delta^j$ corresponds to the jth generalized inlet stream with inlet concentrations:

$$\mathbf{c}_{in}^j := \mathbf{c}_\Delta^j. \tag{2.49}$$

The corresponding (generalized) inlet volumetric flowrate, q_{in}^j, can be modeled as a Dirac impulse, δ, at time instant t^j:

$$q_{in}^j(t) := \delta(t - t^j)\, V_\Delta^j, \tag{2.50}$$

where V_Δ^j is the amount of volume added. Note that the generalized outlet volumetric flowrate q_{out} is 0.

(V3) Let $\mathbf{c}_0^j$ and V_0^j denote the initial concentrations and volume of the jth additional batch run, respectively. Then:

$$\mathbf{c}_{in}^j := \mathbf{c}_0^j. \tag{2.51}$$

The starting phase of the jth additional run has two steps: (i) emptying the reactor from the $(j-1)$st run with final concentrations, $\acute{\mathbf{c}}^{j-1}$, and volume, $\acute{V}^{j-1}$, at the final time instant $\acute{t}^{j-1}$ and (ii) filling the reactor with species having initial concentrations, $\mathbf{c}_0^j$, and volume, V_0^j. The two steps can be formalized in mathematical terms.

Let time t be monotonically increasing from batch to batch: $t \in]\sum_{i=1}^{j-1} \acute{t}^i, \sum_{i=1}^{j} \acute{t}^i]$. With this redefinition of the time axis, the emptying step involves a Dirac in the generalized outlet stream at time instant $t_0^j \equiv \sum_{i=1}^{j-1} \acute{t}^i$ with

$$q_{out}(t) := \delta(t - t_0^j)\, \acute{V}^{j-1}. \tag{2.52}$$

The filling step is equivalent to an impulse addition. Thus, from (2.50), the generalized inlet stream becomes:

$$q_{in}^j(t) := \delta(t - t_0^j)\, V_0^j, \tag{2.53}$$

□

Based on the above discussion, "inlets and outlets" and "generalized inlets and outlets" will be used in an equivalent manner in the remainder of this dissertation.

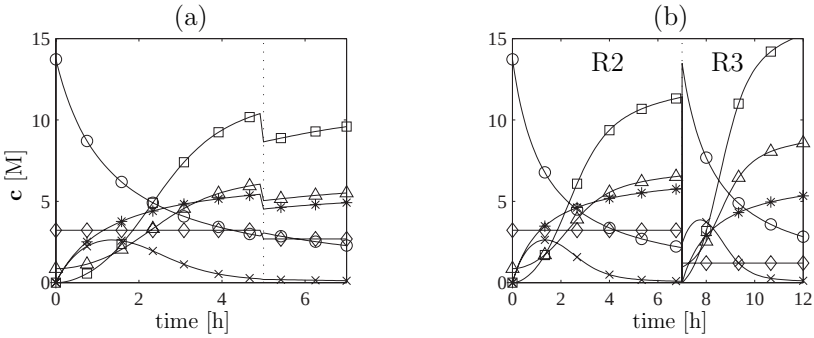

Figure 2.5. Example 2.13: The concentrations $\mathbf{c}$ of (a) Case C1 and (b) Case C2 (Run R2 and R3 for $t \in [0, 7]$h and $t \in]7, 12]$h, respectively). The legend is: $\mathcal{X}_1$ ($\circ$), $\mathcal{X}_2$ ($\times$), $\mathcal{X}_3$ ($\star$), $\mathcal{X}_4$ ($\triangle$), $\mathcal{X}_5$ ($\square$), $\mathcal{X}_6$ ($\diamond$).

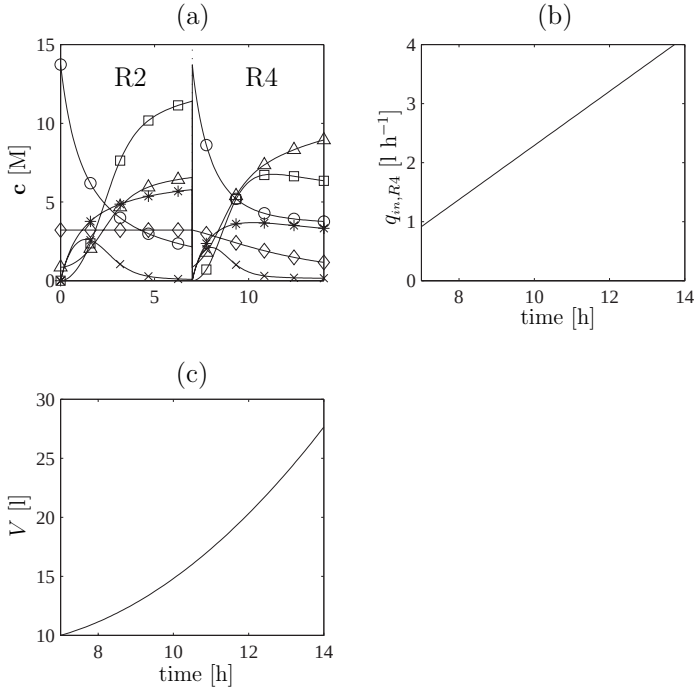

Figure 2.6. Example 2.13: (a) The concentrations $\mathbf{c}$, (b) the physical inlet volumetric flowrate $q_{in,R4}$ (Run R4), and (c) the volume V_{R4} (Run R4) of Case C3. For legend of $\mathbf{c}$, see Figure 2.5 (Runs R2 and R4 for $t \in [0, 7]$h and $t \in]7, 14]$h, respectively).

Example 2.13 (Generalized inlets and outlets)
The three variation modes are illustrated by means of an isothermal, constant-density reaction system (2.38), involving $R = 2$ independent reactions and conducted in a solvent $\mathcal{X}_6$ (modified example of Amrhein *et al.*, 1996):

$$\begin{aligned} 2\mathcal{X}_1 &\rightarrow \mathcal{X}_2 + \mathcal{X}_3 \\ \mathcal{X}_2 &\rightarrow \mathcal{X}_4 + 2\mathcal{X}_5 \end{aligned} \quad (2.54) \qquad \text{with} \qquad \mathbf{N} = \begin{bmatrix} -2 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 2 \end{bmatrix}$$

For the first reaction, power-law kinetics are assumed, i.e., $r_1(c_1) = \kappa_1 c_1^2$, while for the second, power-law kinetics with auto-catalyzation by $\mathcal{X}_4$, i.e., $r_2(c_2, c_4) = \kappa_2 c_2 c_4$. Let $\mathbf{c} = [c_1,\ c_2,\ c_3,\ c_4,\ c_5]^{\mathrm{T}}$.
Four runs are conducted: a batch run R1 with one impulse addition of $\mathcal{X}_1$ at $t = 5\,\mathrm{h}$ ($V_\Delta^1 = 2\,\ell$), two batch runs R2 and R3, and one semibatch run R4 with a single continuous inlet of a mixture of $\mathcal{X}_1$ and $\mathcal{X}_4$. The initial concentrations of Runs R1, R2, and R4 are the same (see Appendix A.2 for the parameters).
Three cases are studied. Cases C1, C2, and C3 consider Run R1, Runs R2 & R3, and Runs R2 & R4, respectively. Figures 2.5 and 2.6 present for Cases C1–C3 the time evolutions of the concentrations $\mathbf{c}$ and the (physical) inlet volumetric flowrate q_{in} and the volume V of Run R4.
The initial and inlet concentrations of the reaction system with generalized inlet and outlet streams for the three cases are:

$$\begin{aligned} \mathbf{c}_{0,C1} &= \mathbf{c}_{0,R1}, & \mathbf{C}_{in,C1} &= \mathbf{c}_{\Delta,R1}^{\mathrm{T}} \\ \mathbf{c}_{0,C2} &= \mathbf{c}_{0,R2}, & \mathbf{C}_{in,C2} &= \mathbf{c}_{0,R3}^{\mathrm{T}} \\ \mathbf{c}_{0,C3} &= \mathbf{c}_{0,R2}, & \mathbf{C}_{in,C3} &= \begin{bmatrix} \mathbf{c}_{in,R4}^{\mathrm{T}} \\ \mathbf{c}_{0,R4}^{\mathrm{T}} \end{bmatrix} \end{aligned}$$

Thus, for Cases C1 and C2, the concentrations of the generalized inlet correspond to the concentrations of the impulse addition and the initial conditions, respectively. For Case C3, two generalized inlets are required with corresponding inlet concentrations containing the initial and inlet concentrations of Run R4. The flowrates of the generalized inlets are given as:

$$\begin{aligned} &q_{in,C1}^1(t) = \delta(t - t^1)\, V_\Delta^1,\ t \in [0,\ 14]\,\mathrm{h} \quad \text{with} \quad t^1 = 5\,\mathrm{h},\ V_\Delta^1 = 2\,\ell \\ &q_{in,C2}^1(t) = \delta(t - t_0^2)\, V_0^2,\ t \in [0,\ 12]\,\mathrm{h} \quad \text{with} \quad t_0^2 = 7\,\mathrm{h},\ V_0^2 = 15\,\ell \\ &q_{in,C3}^1(t) \quad \text{see } q_{in,R4} \text{ in Figure 2.6d} \\ &q_{in,C3}^2(t) = \delta(t - t_0^2)\, V_0^2,\ t \in [0,\ 14]\,\mathrm{h} \quad \text{with} \quad t_0^2 = 7\,\mathrm{h},\ V_0^2 = 10\,\ell \end{aligned} \quad (2.55)$$

and the outlet flowrates as:

$$\begin{aligned} q_{out,C2}^1(t) &= \delta(t - t_0^2)\, \acute{V}^1,\ t \in [0,\ 12]\,\mathrm{h} \quad \text{with} \quad t_0^2 = 7\,\mathrm{h},\ \acute{V}^1 = 10\,\ell \\ q_{out,C3}^1(t) &= \delta(t - t_0^2)\, \acute{V}^1,\ t \in [0,\ 14]\,\mathrm{h} \quad \text{with} \quad t_0^2 = 7\,\mathrm{h},\ \acute{V}^1 = 10\,\ell \end{aligned}$$

For Cases C1–C3, according to (2.55), the generalized inlet volumetric flowrates are zero before the impulse addition, the additional batch run, and the additional semibatch run, respectively. An impulse can be observed in q_{in} of Cases C1–C2, and $q_{in,1}$ of Case C3.

2.5 Special cases

In this section, the dynamic reaction models is considered for the following special cases: CSTRs, semibatch and batch reaction systems, and reactions in quasi-equilibrium conditions. All special cases are based on the assumptions of constant density and temperature but could easily be extended to the cases of varying density and/or temperature.

2.5.1 CSTR, semibatch and batch reaction systems

2.5.1.1 CSTR

In a CSTR, $V(t) = V_0$ and, thus, (2.38) becomes:

$$\begin{aligned} \dot{\boldsymbol{n}} = \frac{d(V\,\mathbf{c})}{dt} = V\,\mathbf{N}^{\mathrm{T}}\,\mathbf{r}(\mathbf{c}) + \mathbf{C}_{in}\,\mathbf{q}_{in} - q_{out}\,\mathbf{c}, \qquad & \boldsymbol{n}(0) = V_0\mathbf{c}_0 = \boldsymbol{n}_0, \\ \dot{V} = 0 = \mathbf{1}_p^{\mathrm{T}}\,\mathbf{q}_{in} - q_{out}, \qquad & V(0) = V_0. \end{aligned} \tag{2.56}$$

From $\dot{V} = 0$, it follows that $q_{out} = \mathbf{1}_p^{\mathrm{T}}\mathbf{q}_{in}$. Thus, the number of linearly-independent inputs is p. Note that for consistency with the subsequent study, V is considered a state and the corresponding equation is retained.

2.5.1.2 Semibatch reaction systems

Owing to the absence of an outlet stream ($q_{out} = 0$), (2.38) becomes:

$$\begin{aligned} \dot{\boldsymbol{n}} = \frac{d(V\,\mathbf{c})}{dt} = V\,\mathbf{N}^{\mathrm{T}}\,\mathbf{r}(\mathbf{c}) + \mathbf{C}_{in}\,\mathbf{q}_{in}, \qquad & \boldsymbol{n}(0) = V_0\mathbf{c}_0 = \boldsymbol{n}_0, \\ \dot{V} = \mathbf{1}_p^{\mathrm{T}}\,\mathbf{q}_{in}, \qquad & V(0) = V_0. \end{aligned} \tag{2.57}$$

2.5.1.3 Batch reaction systems

Owing to the absence of inlet and outlet streams ($\mathbf{q}_{in} = 0$, $q_{out} = 0$), (2.38) turns into (2.20):

$$\begin{aligned} \dot{\boldsymbol{n}} = V_0\,\dot{\mathbf{c}} = V_0\,\mathbf{N}^{\mathrm{T}}\,\mathbf{r}(\mathbf{c}), \qquad & \boldsymbol{n}(0) = V_0\mathbf{c}_0 = \boldsymbol{n}_0, \\ \dot{V} = 0, \qquad & V(0) = V_0. \end{aligned} \tag{2.58}$$

Note that similarly to the CSTR, V is considered a state and the corresponding equation is retained. Furthermore, $\mathbf{c} = \boldsymbol{n}/V_0$, and thus, (2.20) is obtained.

2.5.2 Systems with reactions in quasi-equilibrium conditions

When reactions are in *quasi-equilibrium conditions* (Hill, 1977), the corresponding concentrations can be determined from algebraic equilibrium relationships. These relationships are usually derived from the mass action law. Important cases include fast

neutralization, redox, precipitation, and complexation reactions (Titration, 1954, 1957; Oehme and Richter, 1987).

First, the mass action law that determines the rates of the reactions in quasi-equilibrium is introduced.

Mass action law

Consider the situation of a system with R_a independent reactions in quasi-equilibrium conditions and the remaining $R_b = (R - R_a)$ independent reactions far from equilibrium. Assume, without loss of generality, that the first R_a independent reactions are in quasi-equilibrium conditions. This induces a partition in $\mathbf{N}$ and $\mathbf{r}$ and, thus, $\boldsymbol{n}$ in (2.38) becomes:

$$\dot{\boldsymbol{n}} = \frac{d(V\,\mathbf{c})}{dt} = V\,\begin{bmatrix}\mathbf{N}_a^{\mathrm{T}} & \mathbf{N}_b^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\mathbf{r}_a(\mathbf{c})\\ \mathbf{r}_b(\mathbf{c})\end{bmatrix} + \mathbf{C}_{in}\,\mathbf{q}_{in} - q_{out}\,\mathbf{c}, \quad \boldsymbol{n}(0) = V_0\mathbf{c}_0 = \boldsymbol{n}_0. \tag{2.59}$$

In titration applications (pH-control, etc.), the mass action law is often used to specify $\mathbf{r}_a$ in (2.59) (Izquierdo-Ridorsa *et al.*, 1997; Dyson *et al.*, 1997). Let K_i be the equilibrium constant of the ith reaction. Typically, the corresponding mass action law is given by:

$$\mathsf{k}_i(\mathbf{c}) = \mathsf{K}_i, \quad i = 1, \ldots, R_a, \tag{2.60}$$

where k_i is a function depending on $\mathbf{c}$. Let $\boldsymbol{k} = [\mathsf{k}_1, \ldots, \mathsf{k}_{R_a}]^{\mathrm{T}}$ be an R_a-dimensional vector function. Then, by differentiating (2.60) with respect to time:

$$\mathbf{0}_{R_a} = \frac{d\boldsymbol{k}(\mathbf{c})}{dt} = \boldsymbol{\mathcal{J}}_{\mathsf{k}}\frac{d\mathbf{c}}{dt}, \quad \boldsymbol{\mathcal{J}}_{\mathsf{k}} \equiv \frac{d\boldsymbol{k}}{d\mathbf{c}}, \tag{2.61}$$

where $\boldsymbol{\mathcal{J}}_{\mathsf{k}}$ is the $R_a \times S$ Jacobian matrix of $\boldsymbol{k}$ with respect to $\mathbf{c}$. Computing $\dot{\mathbf{c}}$ explicitly from (2.59) and multiplying by $\boldsymbol{\mathcal{J}}_{\mathsf{k}}$,

$$\mathbf{0}_{R_a} = \boldsymbol{\mathcal{J}}^{\star}_{\mathsf{k}}\mathbf{r}_a + \boldsymbol{\mathcal{J}}_{\mathsf{k}}\left[\mathbf{N}_b^{\mathrm{T}}\mathbf{r}_b + \boldsymbol{\beta}(V,\mathbf{c})\right] \text{ with } \boldsymbol{\beta}(V,\mathbf{c}) \equiv \mathbf{C}_{in}\,\mathbf{q}_{in}/V - \frac{\mathbf{1}_p^{\mathrm{T}}\mathbf{q}_{in}}{V}\,\mathbf{c},\ \boldsymbol{\mathcal{J}}^{\star}_{\mathsf{k}} \equiv \boldsymbol{\mathcal{J}}_{\mathsf{k}}\mathbf{N}_a^{\mathrm{T}}, \tag{2.62}$$

where $\boldsymbol{\beta}$ is an S-dimensional vector and $\boldsymbol{\mathcal{J}}^{\star}_{\mathsf{k}}$ an $R_a \times R_a$ matrix. Assuming the inverse of $\boldsymbol{\mathcal{J}}^{\star}_{\mathsf{k}}$ to exist, (2.62) can be solved for $\mathbf{r}_a$:

$$\mathbf{r}_a(\mathbf{c}, V) = -\left(\boldsymbol{\mathcal{J}}^{\star}_{\mathsf{k}}\right)^{-1}\boldsymbol{\mathcal{J}}_{\mathsf{k}}\left[\mathbf{N}_b^{\mathrm{T}}\,\mathbf{r}_b + \boldsymbol{\beta}(V,\mathbf{c})\right]. \tag{2.63}$$

For $R_a = R$, (2.63) becomes

$$\mathbf{r}(\mathbf{c}, V) = -\left(\boldsymbol{\mathcal{J}}^{\star}_{\mathsf{k}}\right)^{-1}\boldsymbol{\mathcal{J}}_{\mathsf{k}}\boldsymbol{\beta}(V,\mathbf{c}) \tag{2.64}$$

with $\boldsymbol{\mathcal{J}}^{\star}_{\mathsf{k}} = \boldsymbol{\mathcal{J}}_{\mathsf{k}}\mathbf{N}^{\mathrm{T}}$.

Note that the equilibrium constants K_i ($i = 1, \ldots, R_a$) do not appear explicitly in the mole balance equations (2.59). However, they influence them implicitly through the initial concentrations, since the initial concentrations of R_a out of the S species are determined by the mass action law (2.60).

Dynamic model

For simplicity of notation, it is assumed that all R independent reactions are in quasi-equilibrium conditions. For example, in the case of neutralization titrations (Amrhein, 1998), bases (acids) or mixtures of bases (acids) are continuously fed to a reaction system with an optional outlet. Since equilibrium is almost instantaneously reached with the additions, the reaction system does not depend on time anymore. Thus, the system equations (mole balances and continuity equations) can be reparameterized with respect to any other variable, s, (such as the pH), that is monotonically increasing. The monotonicity is guaranteed by:

$$\frac{ds}{dt} \geq 0, \quad s(0) = s_0, \tag{2.65}$$

where s_0 is the initial condition of s. Since $\dot{\boldsymbol{n}} = \frac{d\boldsymbol{n}}{ds}\dot{s}$ and $\dot{V} = \frac{dV}{ds}\dot{s}$, $\mathbf{c}$ and V in (2.38) are reparameterized in terms of the new independent variable s:

$$\begin{aligned} \frac{d\boldsymbol{n}}{ds} = \frac{d(V\,\mathbf{c})}{ds} &= V\,\mathbf{N}^{\mathrm{T}}\,\mathbf{r}(\mathbf{c})/\dot{s} + \mathbf{C}_{in}\,\mathbf{q}_{in}/\dot{s} - q_{out}\,\mathbf{c}/\dot{s}, && \boldsymbol{n}(0) = V_0\mathbf{c}_0 = \boldsymbol{n}_0, \\ \frac{dV}{ds} &= \mathbf{1}_p^{\mathrm{T}}\,\mathbf{q}_{in}/\dot{s} - q_{out}/\dot{s}, && V(0) = V_0, \end{aligned} \tag{2.66}$$

where V_0, $\mathbf{c}_0$, and $\boldsymbol{n}_0$ are the initial conditions at $s = s_0$.

3

Reaction and flow variants/invariants in dynamic models

In the previous chapter, dynamic models for reaction systems with inlet and outlet streams were derived in terms of independent reactions and inlets. These models include information regarding both the chemical reactions (stoichiometry and kinetics) and the operational mode of the reactor. For the analysis, it is important to distinguish between the states that depend on the reactions and/or evolve with time and those which do not. A transformation has been proposed by Waller and Mäkilä (1981) to separate the *reaction-variant* space (same dimension as the number of independent reactions) from the reaction-invariant space. For reaction systems without inlet nor outlet streams, the concentration vector can be reconstructed by integrating only the reaction variants.

For considerations of safety and productivity, however, the reactors used in the production of specialty chemicals are often of the semibatch type (systems with inlets), or run continuously (systems with inlet and outlets). The aim of this chapter is to extend the concept of invariance to include *flow invariants* of reaction systems with inlet and outlet streams. A *transformation to normal form* is proposed herein that performs a three-level decomposition of the states into: (i) the evolution of the independent reactions, (ii) the influence of the inlet and outlets streams, and (iii) the behavior of the *reaction and flow invariants* (see also Srinivasan *et al.*, 1998). No kinetic information (except for the assumption of independent kinetics) is necessary for the analysis.

The proposed transformation to normal form is important for the analysis of dynamic models, concentration measurements, and spectral measurements of reaction systems. It will be shown how this transformation helps determine the minimal dimensionality of dynamic models (i.e., the minimal number of differential equations) and analyze control-relevant properties such as state accessibility and feedback linearizability. The important results regarding state accessibility proposed by Bastin and Lévine (1993) can be readily derived from the transformation to normal form. Under certain conditions, the system can be proven to be feedback linearizable, and the linearizing control can be obtained easily.

For simplicity, reaction systems with constant temperature and density are considered first using the basic model derived in Section 2.3.4. Then, the assumptions of constant

temperature and density are relaxed, and indications for constructing the transformation to normal form are provided. Furthermore, the following special cases are considered: continuous stirred-tank reaction systems (CSTRs), semibatch and batch reaction systems, and systems with reactions in quasi-equilibrium conditions.

The chapter is organized as follows: Section 3.1 develops the transformation to normal form and presents a simulated example. The implications of this transformation are discussed in Section 3.2. Extensions to the basic model are provided in Section 3.3, and special cases are studied in Section 3.4.

3.1 Transformation of the basic dynamic model to normal form

The concept of reaction invariants is extended in this section to include flow invariants of chemical reaction systems with inlet and outlet streams. A transformation to normal form is introduced that allows the separation of: (i) the evolution of the independent reactions, (ii) the influence of the inlet and outlet streams, and (iii) the behavior of the reaction and flow invariants. Resulting from this transformation to normal form, model reduction, state accessibility, and feedback linearizability will be analyzed. More importantly, this transformation will be used in Section 4.1 to derive a model for molar concentration measurements.

The transformation to normal form for the basic model (2.38) is derived next.

3.1.1 Reaction invariants in the absence of inlet and outlet streams

The aim here is to transform the basic dynamic model (2.38) in a form suitable for decoupling reaction variants and invariants. As a first step towards understanding reaction invariants, a chemical reaction system without inlet nor outlet streams (batch reaction system of constant volume) is considered:

$$\begin{aligned} \dot{\boldsymbol{n}} &= V\,\mathbf{N}^{\mathrm{T}}\,\mathbf{r}_{\mathfrak{n}}(\boldsymbol{n}), & \boldsymbol{n}(0) &= \boldsymbol{n}_0, \\ \dot{V} &= 0, & V(0) &= V_0, \end{aligned} \tag{3.1}$$

where $\mathbf{r}_{\mathfrak{n}V}$ is the R-dimensional reaction rate vector, which is a function of $\boldsymbol{n}$ and V, and $\boldsymbol{n}_0$ the initial number of moles. Though there are $(S+1)$ differential equations in (3.1), it is shown below that the reaction system has intrinsically only R states or directions in which $\boldsymbol{n}$ evolves with time, since there are only R independent reactions. The other $(S+1-R)$ directions in $\boldsymbol{n}$ are unaffected by the reactions and are termed the *invariants of the reaction system* (Waller and Mäkilä, 1981).

The state space will be separated into: (i) an R-dimensional vector space that evolves due to the R independent reactions (the *reaction variants*), and (ii) the $(S+1-R)$-dimensional *reaction invariants* (one corresponding to the constant volume). The ideas of Moore–Penrose pseudo-inverse (denoted by superscript +; see Appendix B.4) and completion of space will be used for this purpose. Let $\boldsymbol{\mathcal{P}} \in \mathbb{R}^{S\times(S-R)}$ be a matrix with

orthonormal columns that satisfies: (i) $\text{rank}\,([\mathbf{N}^{\mathrm{T}}, \boldsymbol{\mathcal{P}}]) = S$, and (ii) $\mathbf{N}\,\boldsymbol{\mathcal{P}} = \mathbf{0}_{R\times(S-R)}$. Then, the *linear* transformation that separates the R-dimensional reaction variants $\boldsymbol{z}_1$ and the $(S+1-R)$-dimensional reaction invariants $\boldsymbol{z}_2$ and V in $\boldsymbol{n}$ and V is given by

$$\begin{bmatrix} \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \\ V \end{bmatrix} = \begin{bmatrix} \mathbf{N}^{+\mathrm{T}}\,\boldsymbol{n} \\ \boldsymbol{\mathcal{P}}^{\mathrm{T}}\,\boldsymbol{n} \\ V \end{bmatrix}. \tag{3.2}$$

Substituting (3.2) in (3.1) and using $\mathbf{r}_z(\boldsymbol{z}, V)$ for the reaction rates in the new coordinates gives:

$$\begin{aligned} \dot{\boldsymbol{z}}_1 &= V\,\mathbf{r}_z(\boldsymbol{z}, V), & \boldsymbol{z}_1(0) &= \mathbf{N}^{+\mathrm{T}}\boldsymbol{n}_0, \\ \dot{\boldsymbol{z}}_2 &= \mathbf{0}_{S-R}, & \boldsymbol{z}_2(0) &= \boldsymbol{\mathcal{P}}^{\mathrm{T}}\boldsymbol{n}_0 \\ \dot{V} &= 0, & V(0) &= V_0. \end{aligned} \tag{3.3}$$

$\boldsymbol{n}$ is reconstructed from $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ by

$$\boldsymbol{n} = \mathbf{N}^{\mathrm{T}}\mathbf{z}_1 + \boldsymbol{\mathcal{P}}\mathbf{z}_2. \tag{3.4}$$

In (3.4), the relation $(\mathbf{N}^{\mathrm{T}}\mathbf{N}^{+\mathrm{T}} + \boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{P}}^{\mathrm{T}} = \mathbf{I}_S)$ was used. Note that the term $\boldsymbol{\mathcal{P}}\mathbf{z}_2$ is constant, since it contains the reaction invariants.

In Perry and Green (1984), the extent of Reaction i, $x_{b,i}$, is defined as $\dot{x}_{b,i} \equiv V\,r_i$ with $x_{b,i}(0) = \mathbf{0}_R$. Thus, $\boldsymbol{z}_1$ equals $\mathbf{x}_b$ except for initial conditions.

Remark 3.1 (Construction of $\boldsymbol{\mathcal{P}}$)
The $(S-R)$ columns of $\boldsymbol{\mathcal{P}}$ span the null space of $\mathbf{N}$ (see Appendix B.1 for the definition of the null space). Therefore, one way to construct $\boldsymbol{\mathcal{P}}$ is by singular value decomposition (SVD) of $\mathbf{N}$ (see Appendix B.3). $\boldsymbol{\mathcal{P}}$ is the matrix of the right singular vectors corresponding to the zero singular values of $\mathbf{N}$.

3.1.2 Reaction invariants in the presence of inlet and outlet streams

Even when inlets and outlets are present, the reactions affect only an R-dimensional space. The reaction invariants can be separated using the transformation (3.2). Applying (3.2) to (2.38) with $\mathbf{c} = \boldsymbol{n}/V$ gives,

$$\begin{aligned} \dot{\boldsymbol{z}}_1 &= V\,\mathbf{r}_{\mathbf{z}}(\boldsymbol{z}, V) + \mathbf{N}^{+\mathrm{T}}\mathbf{C}_{in}\,\mathbf{q}_{in} - (q_{out}/V)\,\boldsymbol{z}_1, & \boldsymbol{z}_1(0) &= \mathbf{N}^{+\mathrm{T}}\boldsymbol{n}_0, \\ \dot{\boldsymbol{z}}_2 &= \boldsymbol{\mathcal{P}}^{\mathrm{T}}\,\mathbf{C}_{in}\,\mathbf{q}_{in} - (q_{out}/V)\,\boldsymbol{z}_2, & \boldsymbol{z}_2(0) &= \boldsymbol{\mathcal{P}}^{\mathrm{T}}\boldsymbol{n}_0, \\ \dot{V} &= \left(\mathbf{1}_p^{\mathrm{T}}\mathbf{q}_{in}\right) - q_{out}, & V(0) &= V_0. \end{aligned} \tag{3.5}$$

$\boldsymbol{n}$ is reconstructed from (3.4). Note that, though $\boldsymbol{z}_2$ and V are reaction invariants, they do not remain constant. Thus, the system evolution is no longer restricted to an R-dimensional space as previously; it evolves in a larger space. A similar conclusion was drawn by Fjeld *et al.* (1974).

In the absence of inlets and outlets, both isolating the R-dimensional manifold in which $\boldsymbol{n}$ evolves and spotting the $(S+1-R)$ reaction invariants are the same problem. This is because the variation in $\boldsymbol{n}$ is caused only through the R independent reactions. In contrast, in the presence of inlets and outlets, the variation in $\boldsymbol{n}$ and V is not only due to the reactions but also to flows (inlets and outlets). Thus, isolating the manifold in which $\boldsymbol{n}$ and V evolve means finding and eliminating those states which are invariant with respect to both reactions and flows.

3.1.3 Reaction and flow invariants

The key aim of this chapter is to provide a three-part decomposition of the system equations (2.38). The first part is that which is affected by the R independent reactions (reaction variants). Among the reaction invariants, the second part picks those states which evolve with the flows (reaction invariants *and* flow variants). The third part consists of those states which remain constant (reaction *and* flow invariants). The inlet and outlet streams will be treated separately, since the inlet streams enter linearly while the outlet streams nonlinearly in the system equations. Thus, in the absence of outlet streams, the *transformation to normal form* proposed to this effect is *linear* while, in the presence of outlet streams, it is *nonlinear*.

The dimension of the reaction-variant space is R. As it will be shown below, the dimension of the manifold in which the transformed states evolve is given by $\varsigma + 1$, where $\varsigma \equiv \operatorname{rank}([\mathbf{N}^{\mathrm{T}}, \mathbf{C}_{in}])$. This means that the dimension of the manifold increases with every inlet stream whose concentrations are independent of the rows of $\mathbf{N}$ and the remaining inlet concentrations. First, it is assumed that $\varsigma = R + p$. It will be shown that the dimension of the reaction-variant space is R, that of the reaction-invariant and flow-variant space is $p + 1$, and that of the reaction and flow-invariant space is $S - \varsigma$. Next, the transformation to normal form will be developed, and it will then be generalized to include the case of $\varsigma < R + p$.

In comparison with (3.2), the completion of the state space involving $\boldsymbol{n}$ and V takes a three-level structure. Let $\mathbf{L} \in \mathbb{R}^{S\times(\varsigma-R)}$ and $\mathbf{\Omega} \in \mathbb{R}^{S\times(S-\varsigma)}$ be matrices with orthonormal columns that satisfy the following conditions: (i) $\operatorname{rank}([\mathbf{N}^{\mathrm{T}}, \mathbf{L}, \mathbf{\Omega}]) = S$, (ii) $\mathbf{N}^{\mathrm{T}}$, $\mathbf{L}$ and $\mathbf{\Omega}$ are mutually orthogonal, (iii) $\mathbf{\Omega}^{\mathrm{T}}\mathbf{C}_{in} = \mathbf{0}_{S-\varsigma\times p}$, (iv) $\mathbf{L}^{\mathrm{T}}\mathbf{C}_{in}$ is invertible.

For the application of the three-part decomposition of $\boldsymbol{n}$, the transformation to normal form (3.2) must be changed as follows: (i) instead of $\boldsymbol{\mathcal{P}}$, an $S \times (\varsigma - R)$ matrix, $\mathbf{\mathcal{M}}$, must be chosen that renders $\mathbf{\mathcal{M}}^{\mathrm{T}}\mathbf{C}_{in} = \mathbf{I}_{(\varsigma-R)}$, which is achieved by choosing $\mathbf{\mathcal{M}} \equiv \mathbf{L}(\mathbf{C}_{in}^{\mathrm{T}}\,\mathbf{L})^{-1}$; (ii) $\mathbf{N}^{+\mathrm{T}}\mathbf{C}_{in}\,\mathbf{q}_{in}$ must be eliminated. For this, an additional projection matrix $(\mathbf{I}_S - \mathbf{C}_{in}\mathbf{\mathcal{M}}^{\mathrm{T}})$ is formulated that, by construction of $\mathbf{\mathcal{M}}$, satisfies $(\mathbf{I}_S - \mathbf{C}_{in}\mathbf{\mathcal{M}}^{\mathrm{T}})\,\mathbf{N}^{\mathrm{T}} = \mathbf{N}^{\mathrm{T}}$. With these modifications, the *linear* transformation of $\boldsymbol{n}$ reads:

$$\begin{bmatrix} \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \\ \boldsymbol{z}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{N}^{+\mathrm{T}}\,(\mathbf{I}_S - \mathbf{C}_{in}\mathbf{\mathcal{M}}^{\mathrm{T}})\,\boldsymbol{n} \\ \mathbf{\mathcal{M}}^{\mathrm{T}}\,\boldsymbol{n} \\ \mathbf{\Omega}^{\mathrm{T}}\boldsymbol{n} \end{bmatrix} \tag{3.6}$$

and

$$\begin{aligned}
\dot{\boldsymbol{z}}_1 &= V\,\mathbf{r}_{\mathrm{z}}(\boldsymbol{z},V) - (q_{out}/V)\,\boldsymbol{z}_1, & \boldsymbol{z}_1(0) &= \mathbf{N}^{+\mathrm{T}}\,(\mathbf{I}_S - \mathbf{C}_{in}\boldsymbol{\mathcal{M}}^{\mathrm{T}})\,\boldsymbol{n}_0,\\
\dot{\boldsymbol{z}}_2 &= \mathbf{q}_{in} - (q_{out}/V)\,\boldsymbol{z}_2, & \boldsymbol{z}_2(0) &= \boldsymbol{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}_0,\\
\dot{\boldsymbol{z}}_3 &= -(q_{out}/V)\,\boldsymbol{z}_3, & \boldsymbol{z}_3(0) &= \boldsymbol{\Omega}^{\mathrm{T}}\boldsymbol{n}_0,\\
\dot{V} &= \left(\mathbf{1}_p^{\mathrm{T}}\mathbf{q}_{in}\right) - (q_{out}/V)\,V, & V(0) &= V_0,
\end{aligned} \tag{3.7}$$

where $\boldsymbol{z}_1$, $\boldsymbol{z}_2$, and $\boldsymbol{z}_3$ are vectors of dimension R, p, and $(S-\varsigma)$, respectively. For the elimination of the term (q_{out}/V), a *nonlinear* transformation can be used:

$$\begin{bmatrix}\boldsymbol{x}\\ \lambda\end{bmatrix} = \begin{bmatrix}\boldsymbol{z}/(V - \mathbf{1}_p^{\mathrm{T}}\,\boldsymbol{z}_2)\\ V - \mathbf{1}_p^{\mathrm{T}}\,\boldsymbol{z}_2\end{bmatrix}, \tag{3.8}$$

where $\boldsymbol{x}$ and λ are the transformed states of dimension $(R+S+p-\varsigma)$ and 1, respectively. By introducing a nonzero scaling factor η, the following theorem that formulates the so-called *normal form* of (2.38) is proposed.

Theorem 3.2
Let $\varsigma = \operatorname{rank}([\mathbf{N}^{\mathrm{T}},\,\mathbf{C}_{in}]) = (R+p)$. *Then, a diffeomorphism* $\mathcal{T} : [\,\frac{\boldsymbol{n}}{V}\,] \leftrightarrow [\,\frac{\boldsymbol{x}}{\lambda}\,]$ *exists that transforms model* (2.38) *into:*

$$\begin{aligned}
\dot{\boldsymbol{x}}_1 &= \mathrm{h}(\boldsymbol{x}_2)\,\mathbf{r}_{\boldsymbol{\chi}}(\boldsymbol{x}), & \boldsymbol{x}_1(0) &= \eta\,\mathbf{N}^{+\mathrm{T}}\,(\mathbf{I}_S - \mathbf{C}_{in}\,\boldsymbol{\mathcal{M}}^{\mathrm{T}})\boldsymbol{n}_0,\\
\dot{\boldsymbol{x}}_2 &= \mathbf{q}_{in}/\lambda, & \boldsymbol{x}_2(0) &= \eta\,\boldsymbol{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}_0,\\
\dot{\boldsymbol{x}}_3 &= \mathbf{0}_{S-\varsigma}, & \boldsymbol{x}_3(0) &= \eta\,\boldsymbol{\Omega}^{\mathrm{T}}\,\boldsymbol{n}_0,\\
\dot{\lambda} &= -q_{out}/\mathrm{h}(\boldsymbol{x}_2), & \lambda(0) &= 1/\eta,
\end{aligned} \tag{3.9}$$

where

$$\mathrm{h}(\boldsymbol{x}_2) = \eta\,(V_0 - \mathbf{1}_p^{\mathrm{T}}\boldsymbol{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}_0) + \mathbf{1}_p^{\mathrm{T}}\boldsymbol{x}_2, \tag{3.10}$$

and $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, *and* $\boldsymbol{x}_3$ *are vectors of dimension* R, p, *and* $(S-\varsigma)$, *respectively,* η *a nonzero arbitrary constant,* $\mathbf{r}_{\boldsymbol{\chi}}$ *the* R*-dimensional reaction rate vector expressed in terms of* $\boldsymbol{x}$, $\boldsymbol{\mathcal{M}} = \mathbf{L}(\mathbf{C}_{in}^{\mathrm{T}}\,\mathbf{L})^{-1}$, $\mathbf{L} \in \mathbb{R}^{S\times(\varsigma-R)}$, *and* $\boldsymbol{\Omega} \in \mathbb{R}^{S\times(S-\varsigma)}$ *matrices with orthonormal columns that satisfy: (i)* $\operatorname{rank}([\mathbf{N}^{\mathrm{T}},\,\mathbf{L},\,\boldsymbol{\Omega}]) = S$, *(ii)* $\mathbf{N}^{\mathrm{T}}$, $\mathbf{L}$ *and* $\boldsymbol{\Omega}$ *are mutually orthogonal, (iii)* $\boldsymbol{\Omega}^{\mathrm{T}}\mathbf{C}_{in} = \mathbf{0}_{S-\varsigma\times p}$, *(iv)* $\mathbf{L}^{\mathrm{T}}\mathbf{C}_{in}$ *is invertible.*

The transformation to normal form $\mathcal{T}$ *is one-to-one and can be written as follows using* $\mathrm{g}(\boldsymbol{n},V) = \eta\,(V_0 - \mathbf{1}_p^{\mathrm{T}}\boldsymbol{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}_0)/(V - \mathbf{1}_p^{\mathrm{T}}\boldsymbol{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n})$:

$$\begin{bmatrix}\boldsymbol{n}\\ V\end{bmatrix} \to \begin{bmatrix}\boldsymbol{x}\\ \lambda\end{bmatrix} \quad : \quad \begin{bmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\\ \boldsymbol{x}_3\\ \lambda\end{bmatrix} = \begin{bmatrix}\mathrm{g}(\boldsymbol{n},V)\,\mathbf{N}^{+\mathrm{T}}\,(\mathbf{I}_S - \mathbf{C}_{in}\,\boldsymbol{\mathcal{M}}^{\mathrm{T}})\boldsymbol{n}\\ \mathrm{g}(\boldsymbol{n},V)\,\boldsymbol{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}\\ \mathrm{g}(\boldsymbol{n},V)\,\boldsymbol{\Omega}^{\mathrm{T}}\,\boldsymbol{n}\\ 1/\mathrm{g}(\boldsymbol{n},V)\end{bmatrix} \tag{3.11}$$

$$\begin{bmatrix}\boldsymbol{x}\\ \lambda\end{bmatrix} \to \begin{bmatrix}\boldsymbol{n}\\ V\end{bmatrix} \quad : \quad \begin{bmatrix}\boldsymbol{n}\\ V\end{bmatrix} = \begin{bmatrix}\lambda\,(\mathbf{N}^{\mathrm{T}}\,\boldsymbol{x}_1 + \mathbf{C}_{in}\,\boldsymbol{x}_2 + \boldsymbol{\Omega}\boldsymbol{x}_3)\\ \lambda\,\mathrm{h}(\boldsymbol{x}_2)\end{bmatrix}. \tag{3.12}$$

(See Appendix H.2 for proof)

The matrices $\mathbf{\Omega}$ and $\mathbf{L}$ required in Theorem 3.2 can be constructed similarly as in Remark 3.1 by noting that the columns of $\mathbf{\Omega}$ and $\mathbf{L}$ span the null spaces of $\begin{bmatrix} \mathbf{N} \\ \mathbf{C}_{in}^{\mathrm{T}} \end{bmatrix}$ and $\begin{bmatrix} \mathbf{N} \\ \mathbf{\Omega}^{\mathrm{T}} \end{bmatrix}$, respectively. Note that the concentrations can be reconstructed from (3.12) as:

$$\mathbf{c} = \boldsymbol{n}/V = (\mathbf{N}^{\mathrm{T}}\,\boldsymbol{x}_1 + \mathbf{C}_{in}\,\boldsymbol{x}_2 + \mathbf{\Omega}\boldsymbol{x}_3)/\mathsf{h}(\boldsymbol{x}_2). \tag{3.13}$$

The transformation to normal form is global in nature, since $V(t) - \mathbf{1}_p^{\mathrm{T}}\mathbf{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}(t) = 0$ for all t if and only if $V_0 - \mathbf{1}_p^{\mathrm{T}}\mathbf{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}_0 = 0$. If $V_0 - \mathbf{1}_p^{\mathrm{T}}\mathbf{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}_0 = 0$, then g is defined using de l'Hospital's rule (see proof of Theorem 3.2).

Theorem 3.2 can easily be extended to the case where $\varsigma < R + p$. For this, the matrix $\mathbf{\mathcal{M}}$ is redefined using the pseudo-inverse of $(\mathbf{C}_{in}^{\mathrm{T}}\,\mathbf{L})$ instead of the inverse such that $\mathbf{\mathcal{M}} := \mathbf{L}\,(\mathbf{C}_{in}^{\mathrm{T}}\,\mathbf{L})^{+}$. It will be shown in Subsection 3.1.4 that, in this case, the transformation to normal form is no longer one-to-one.

Corollary 3.3 *For $\varsigma < R + p$, the transformed model* (3.9) *represents* (2.38)*, where $\mathbf{\mathcal{M}} := \mathbf{L}\,(\mathbf{C}_{in}^{\mathrm{T}}\,\mathbf{L})^{+}$. The relation $[\,\frac{\boldsymbol{x}}{\lambda}\,] \to [\,\frac{\boldsymbol{n}}{V}\,]$ is defined by* (3.12)*. (See Appendix H.2 for proof)*

3.1.4 Discussion

The results obtained in Theorem 3.2 and Corollary 3.3 are summarized in Table 3.1. Though a rigorous interpretation of the results is difficult due to the nonlinear term in the transformation to normal form, an intuitive meaning for the various parts of $\boldsymbol{x}$ can be given. The R-dimensional reaction variants are described by $\boldsymbol{x}_1$, while $\boldsymbol{x}_2$ captures the p-dimensional flow part affected directly by the inlets. The $(S - \varsigma)$-dimensional reaction and flow invariants are $\boldsymbol{x}_3$, whereas λ is related to the outflow.

The third row depicts the differential equations that explain why the interpretation of the states outlined earlier is possible. The fourth and fifth rows summarize the reconstruction of the original states $\boldsymbol{n}$ and V from $\boldsymbol{x}$ and λ. With this transformation to normal form, $\mathbf{c}$ (sixth row) and $\boldsymbol{n}$ (third row) are decomposed along the directions determined by the columns of: (i) $\mathbf{N}^{\mathrm{T}}$, (ii) $\mathbf{C}_{in}$, and (iii) the reaction and flow-invariant matrix $\mathbf{\Omega}$. In contrast to $\boldsymbol{n}$, $\mathbf{c}$ is not directly influenced by the outflow through λ. Note that $\boldsymbol{x}_2$ and λ form the flow variant space. However, since λ enters nonlinearly in the reconstruction of $\boldsymbol{n}$ and V, it is given a special symbol. $\mathsf{h}(\boldsymbol{x}_2)$ describes the direct influence of the inlet streams on the volume. Alternatively, the volume is reconstructed by taking into account the outlet stream through λ.

In the last row, the dimensions of the various spaces are given. The transformed system is of dimension $(S + R + p - \varsigma + 1)$, whereas the original system has dimension $(S + 1)$. Thus, if $\varsigma < (R + p)$, there is an increase in dimension as a result of the transformation to normal form, and the transformation is not one-to-one. An example is given as illustration next.

Example 3.4 (Rank deficient $\begin{bmatrix} \mathbf{N}^{\mathrm{T}} & \mathbf{C}_{in} \end{bmatrix}$)

Consider a reaction system with two independent reactions ($R = 2$): $\mathcal{X}_1 \to \mathcal{X}_2$, $\mathcal{X}_2 +$

Table 3.1. Summary of the transformation to normal form.

Original space	Reaction variants	Reaction invariants *and* flow variants (Inlets)	Reaction and flow invariants	Reaction invariants *and* flow variants (Outlet)
	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_3$	λ
	$\dot{\boldsymbol{x}}_1 = \mathsf{h}(\boldsymbol{x}_2)\,\mathbf{r}_{\mathrm{X}}(\boldsymbol{x})$	$\dot{\boldsymbol{x}}_2 = \mathbf{q}_{in}/\lambda$	$\dot{\boldsymbol{x}}_3 = \mathbf{0}$	$\dot{\lambda} = -q_{out}/\mathsf{h}(\boldsymbol{x}_2)$
$\boldsymbol{n} =$	$(\mathbf{N}^{\mathrm{T}}\boldsymbol{x}_1$	$+\ \mathbf{C}_{in}\boldsymbol{x}_2$	$+\ \boldsymbol{\Omega}\boldsymbol{x}_3)$	$\cdot\ \lambda$
$V =$		$\mathsf{h}(\boldsymbol{x}_2)$		$\cdot\ \lambda$
$\mathbf{c} =$	$\dfrac{(\mathbf{N}^{\mathrm{T}}\boldsymbol{x}_1}{\mathsf{h}(\boldsymbol{x}_2)}$	$+\ \dfrac{\mathbf{C}_{in}\boldsymbol{x}_2}{\mathsf{h}(\boldsymbol{x}_2)}$	$+\ \dfrac{\boldsymbol{\Omega}\boldsymbol{x}_3)}{\mathsf{h}(\boldsymbol{x}_2)}$	
$\mathbb{R}^{S+1}$	$\mathbb{R}^R$	$\mathbb{R}^p$	$\mathbb{R}^{(S-\varsigma)}$	$\mathbb{R}^1$

$\mathcal{X}_3 \to \mathcal{X}_4$. When $\mathcal{X}_1$ and $\mathcal{X}_2$ are fed as two different inlets ($p = 2$), then

$$\begin{bmatrix}\mathbf{N}^{\mathrm{T}} & \mathbf{C}_{in}\end{bmatrix} = \begin{bmatrix} -1 & 0 & c_{in,1} & 0 \\ 1 & -1 & 0 & c_{in,2} \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

and $\varsigma = 3 < R + p = 4$. This is due to the fact that feeding $\mathcal{X}_1$ along with the reaction $\mathcal{X}_1 \to \mathcal{X}_2$ is equivalent to feeding $\mathcal{X}_2$. This redundancy is reflected in the loss of rank in $[\mathbf{N}^{\mathrm{T}}, \mathbf{C}_{in}]$.

Owing to redundancy, a clear separation of the influence of reactions and flows is no longer possible. One way to handle this is to explain the same effect by more than one state, thus, artificially increasing the dimension of the state space (see Table 3.1). Thus, among the $(p+1)$ states present in $\left[\begin{smallmatrix}\boldsymbol{x}_2\\ \lambda\end{smallmatrix}\right]$ and constituting the reaction-invariant and flow-variant space, at most $(\varsigma - R + 1)$ are linearly independent (i.e., original space $(S+1)$ — reaction variant space (R) — reaction and flow-invariant space $(S-\varsigma)$).

Example 2.5a (Simulated example; cont'd from page 12**)**
The dynamic model in normal form is illustrated on a semibatch reaction system ($R = 3$ and $S = 7$), namely the ethanolysis reaction system described in Example 2.5 on page 12 with $\mathbf{N} = \left[\begin{smallmatrix} -1 & -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 & 1 \end{smallmatrix}\right]$. Since $S = 7$, the dynamic model is of order 8 (including the volume as the last state). The reaction rates are assumed to follow the power law:

$$r_1 = \kappa_1 c_1 c_2, \qquad r_2 = \kappa_2 c_2 c_3, \qquad r_3 = \kappa_3 c_2 c_4 - \kappa_4 c_6 c_7.$$

The numerical values for κ_i $(i = 1, \ldots, 4)$ are [11.07, 8.93, 12.00, 8.02] $\mathrm{M}^{-1}\ \mathrm{h}^{-1}$.
Semibatch operation with $p = 2$ inlet streams and no outlet will now be studied with $\mathbf{C}_{in}$ chosen such that $\mathrm{rank}\,([\mathbf{N}^{\mathrm{T}}, \mathbf{C}_{in}]) = R + p$. In this way, the transformation to

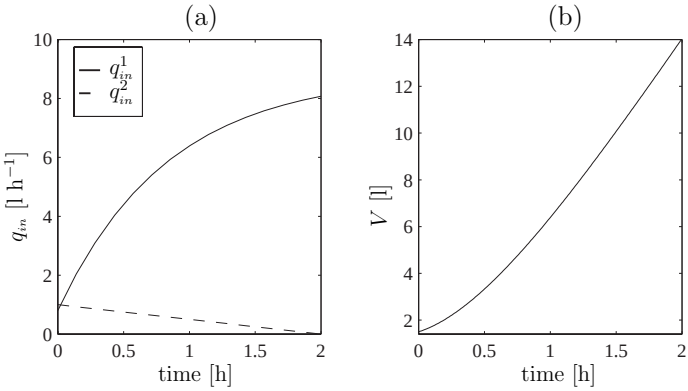

Figure 3.1. Time profiles (Example 2.5a) of: (a) the inlet volumetric flowrates $\mathbf{q}_{in}$, and (b) the volume V.

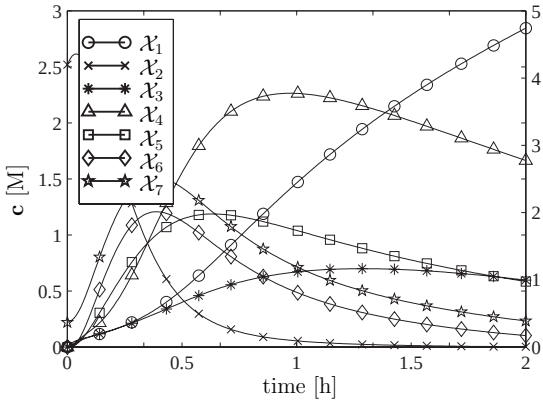

Figure 3.2. Time profiles of the concentrations **c** in Example 2.5a (right axis for c_2 and left axis for the other species).

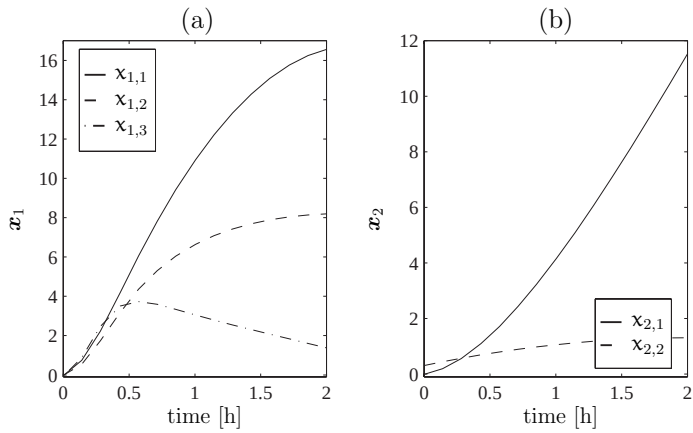

Figure 3.3. Time profiles (Example 2.5a) of: (a) $\boldsymbol{x}_1$, and (b) $\boldsymbol{x}_2$.

normal form (diffeomorphism) of Theorem 3.2 can be illustrated. Ethanol ($V_0 = 1.5$ l with 2 wt. % water) in 80 wt. % solvent (dimethyl sulfoxide) is initially placed in a vessel, i.e., $\mathbf{c}_0 = [0, 4.2, 0, 0, 0, 0, 0.22]^{\mathrm{T}}$ M. The first inlet stream contains phthaloyl dichloride, and the second an ethanol/water mixture, i.e., $\mathbf{c}_{in}^1 = [4.9, 0, 0, 0, 0, 0, 0]^{\mathrm{T}}$ M and $\mathbf{c}_{in}^2 = [0, 20, 0, 0, 0, 0, 1.5]^{\mathrm{T}}$ M. The profiles of the inlet volumetric flowrates and the volume are given in Figure 3.1, and those of the concentration in Figure 3.2.
Model (2.38) can be reduced to $(R + p) = 5$ differential equations as given in Theorem 3.2: $\dot{\lambda} = 0$, since $q_{out} = 0$, and there are $(S - R - p = 2)$ reaction and flow invariants. With the choice $\eta \equiv 1$, $\lambda(t) = \lambda_0 = 1$ and $\boldsymbol{x}_3(t) = \boldsymbol{x}_3(0) = V_0 \boldsymbol{\Omega}^{\mathrm{T}} \mathbf{c}_0$. From (3.9) and (3.12), the reduced model and the reconstructed concentrations $\mathbf{c}$ read:

$$\begin{aligned} \begin{bmatrix} \dot{\boldsymbol{x}}_1 \\ \dot{\boldsymbol{x}}_2 \end{bmatrix} &= \begin{bmatrix} \mathrm{h}(\boldsymbol{x}_2)\, \mathbf{r}_{\mathrm{X}}(\boldsymbol{x}) \\ \mathbf{q}_{in} \end{bmatrix}, \quad \boldsymbol{x}_1(0) = \begin{bmatrix} -0.034 \\ -0.023 \\ -0.065 \end{bmatrix}, \quad \boldsymbol{x}_2(0) = \begin{bmatrix} -0.007 \\ 0.209 \end{bmatrix}, \\ \mathbf{c} &= \left(\mathbf{N}^{\mathrm{T}} \boldsymbol{x}_1 + \mathbf{C}_{in}\, \boldsymbol{x}_2 + \boldsymbol{\Omega}\, \boldsymbol{x}_3\right) / \mathrm{h}(\boldsymbol{x}_2), \\ V &= \mathrm{h}(\boldsymbol{x}_2), \end{aligned} \tag{3.14}$$

where

$$\boldsymbol{\Omega} = \begin{bmatrix} 0 & -0.014 & 0.329 & -0.343 & 0.657 & -0.550 & 0.194 \\ 0 & 0.049 & 0.222 & -0.172 & 0.444 & 0.536 & -0.659 \end{bmatrix}^{\mathrm{T}}$$

and $\boldsymbol{x}_3 = [-0.028, 0.094]^{\mathrm{T}}$. The time profiles of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are shown in Figure 3.3. $\mathrm{x}_{1,1}$ and $\mathrm{x}_{1,2}$ are monotonically increasing, which is characteristic for irreversible reactions, while $\mathrm{x}_{1,3}$ has a increasing and decreasing part, which is characteristic for reversible reactions. $\mathrm{x}_{2,1}$ and $\mathrm{x}_{2,2}$ are monotonically increasing, since they are the integral of the (positive) inlet flowrates.

3.2 Some implications

The transformation to normal form of dynamic reaction models with *independent inlet and outlet streams* proposed in Section 3.1 clearly shows the evolution of $\boldsymbol{n}$ in three different spaces: the stoichiometric space ($\mathcal{S}_c(\mathbf{N}^{\mathrm{T}})$), the inlet space ($\mathcal{S}_c(\mathbf{C}_{in})$), and the invariant space ($\mathcal{S}_c(\boldsymbol{\Omega})$). Such a separation has immediate implications for model-order reduction, state accessibility, and linearization by feedback.

The concepts of model-order reduction, state accessibility, and linearization by feedback are studied in the light of reaction variants and invariants of chemical reaction systems under the usual assumptions (constant density and temperature), i.e., the basic model. It will be shown in Subsection 3.3 that the other models presented in Subsection 2.3.4 also have a normal form similar to that of the basic model. Thus, model-order reduction, state accessibility, and linearization by feedback can readily be extended.

3.2.1 Model-order reduction

From (3.9), it is clear that only $(R+p+1)$ differential equations need to be integrated to determine the trajectories. In addition, a set of $(S-\varsigma)$ constants must be computed from the initial conditions. For $S > (R+p)$, the analysis and design of controllers, observers and optimizers can be achieved with the reduced model of $(R+p+1)$ states if the initial conditions $\boldsymbol{n}_0$ and V_0 are known.

3.2.2 State accessibility

The notion of accessibility is one of the key concepts in nonlinear control theory, since it is closely related to the possibility of moving the states *via* the manipulated variables (inputs). Accessibility corresponds to the ability of reaching a subspace of the state space in final time. First, a definition for state accessibility is provided. Then, accessibility of chemical reaction systems is studied using the formulation of the dynamic model in normal form (3.9).

The set of reachable states at time T is introduced:

Definition 3.5 (Reachable set $\mathcal{R}(\boldsymbol{x}_0, T)$) *Let $\mathcal{R}(\boldsymbol{x}_0, T)$ be the reachable set from $\boldsymbol{x}_0$ at time $T > 0$ for which trajectories remain in the neighborhood $\mathcal{V}$ of $\boldsymbol{x}_0$ for $t \leq T$. In other words, there exists a set of control inputs that steers the system from the initial state $\boldsymbol{x}_0$ at time 0 to the final state $\mathbf{x}_f \in \mathcal{R}(\boldsymbol{x}_0, T)$ at time T such that the evolution of $\boldsymbol{x}(t)$ satisfies $\boldsymbol{x}(t) \in \mathcal{V}$, $0 \leq t \leq T$.*

Definition 3.6 (Local strong accessibility) *A system is said to be locally strongly accessible from a state $\boldsymbol{x}_0 \in \mathcal{X}$, where $\mathcal{X}$ denotes the space of states, if the reachable set $\mathcal{R}(\boldsymbol{x}_0, T)$ has a non-empty interior for all neighborhoods $\mathcal{V}$ of $\boldsymbol{x}_0$ and any $T > 0$ sufficiently small.*

Note that accessibility is by definition a local concept. To test on global accessibility, the local property must be validated for every initial point of the entire space.

It is obvious from the transformation to normal form in Theorem 3.2 that: (i) $\boldsymbol{x}_3$ is inaccessible, and (ii) $\boldsymbol{x}_2$ and λ are (locally strongly) accessible from the inlet and outlet streams. Moreover, a part of $\boldsymbol{x}_1$ can be inaccessible depending on the kinetics. Thus, the dimension of the inaccessible part is at least $(S - \varsigma)$, which leads to the following corollary.

Corollary 3.7 *The maximum dimension of the locally strongly accessible part is* $(\varsigma+1)$*, where* $\varsigma = \operatorname{rank}([\mathbf{N}^{\mathrm{T}}, \mathbf{C}_{in}])$.

Therefore, if local strong accessibility of the full state is required, the dimension of $\boldsymbol{x}_3$ should be zero (i.e., $\varsigma = S$) implying that the number of inlet streams should be at least $(S - R)$ (i.e., $p \geq S - R$).

Corollary 3.8 $\varsigma = S$ *is a necessary condition for local strong accessibility.*

These results have been reported by Bastin and Lévine (1993) for reaction systems with inlets normalized by the volume. This way, they were able to disregard the volume as an additional state, and the maximum dimension of the locally strongly accessible part was found to be ς. Here, the influence of the volume is considered explicitly. Furthermore, deriving the results from the transformation to normal form in Theorem 3.2 provides an alternate proof to Bastin and Lévine (1993).

Another interesting aspect of the transformation to normal form is that the dynamics of the inaccessible part can be transformed to $\dot{\boldsymbol{x}}_3 = \mathbf{0}_{S-\varsigma}$, which is marginally stable.

3.2.3 Linearization by feedback

Different approaches exist to linearize a nonlinear system by state feedback (Isidori, 1989; Nijmeijer and van der Schaft, 1990; Vidyasagar, 1993). A linearized system has the advantage that an external linear control loop can be designed using linear control theory (Chen, 1984). Here, input state feedback linearization or, briefly, feedback linearization (FL; diffeomorphism) of the dynamic model in normal form (3.9) is studied. First, the definition of feedback linearizability and, then, linearizing feedback for chemical reaction systems are presented.

Definition 3.9 (Feedback linearizability) *A system is feedback linearizable around* $\boldsymbol{x}_0$*, if a neighborhood of* $\boldsymbol{x}_0$ *exists for which a transformation* $\mathcal{T}$ *and a feedback law consisting of an* m*-dimensional vector field* $\mathbf{p}$*, an* $m \times m$ *invertible matrix* $\mathbf{Q}$*, and an* m*-dimensional external input vector* $\mathbf{v}$ *with*

$$\begin{aligned} \boldsymbol{z} &= \mathcal{T}(\boldsymbol{x}) \\ \mathbf{u} &= \boldsymbol{p}(\boldsymbol{x}) + \mathbf{Q}(\boldsymbol{x})\,\mathbf{v} \end{aligned} \tag{3.15}$$

transforms the state vector $\boldsymbol{x}$ *into a new* n_l*-dimensional state vector* $\boldsymbol{z}$ *that evolves according to the linear dynamic model:*

$$\dot{\boldsymbol{z}} = \boldsymbol{\mathcal{A}}\,\boldsymbol{z} + \boldsymbol{\mathcal{B}}\,\mathbf{v}, \tag{3.16}$$

where $\boldsymbol{\mathcal{A}}$ *and* $\boldsymbol{\mathcal{B}}$ *are matrices of dimension* $n_l \times n_l$ *and* $n_l \times m$*, respectively.*

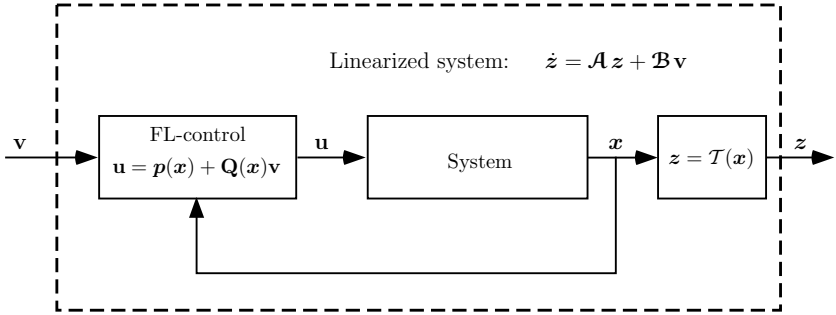

Figure 3.4. Feedback linearized (FL) system.

The concept of feedback linearization is illustrated in Figure 3.4. In practice, the nonlinear feedback law is usually difficult to obtain. For particular reaction systems, a linearizing feedback was found (e.g., Rothfuss *et al.*, 1996; Aguilar *et al.*, 1997; Valluri *et al.*, 1998). However, these approaches use the knowledge of kinetics structure. Here, under only minor assumptions on the kinetics that will be pointed later, the dynamic model in normal form is used for the construction of the linearizing feedback.

Consider the system (3.9) without the inaccessible part $\boldsymbol{x}_3$. The reduced system is given by:

$$\dot{\boldsymbol{x}} = \begin{bmatrix} \dot{\boldsymbol{x}}_1 \\ \hline \dot{\boldsymbol{x}}_2 \\ \hline \dot{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1(\boldsymbol{x}) \\ \hline \mathbf{0}_p \\ \hline 0 \end{bmatrix} + \left[\begin{array}{c|c} \multicolumn{2}{c}{\mathbf{0}_{R\times m}} \\ \hline (1/\lambda)\mathbf{I}_p & \mathbf{0}_p \\ \hline \mathbf{0}_p^{\mathrm{T}} & -1/\mathrm{h}(\boldsymbol{x}_2) \end{array}\right] \mathbf{u}, \quad \mathbf{u} \equiv \begin{bmatrix} \mathbf{q}_{in} \\ q_{out} \end{bmatrix}, \tag{3.17}$$

where $\mathbf{u}$ is the m-dimensional input vector, and m and n the dimensions of $\mathbf{u}$ and $\boldsymbol{x}$, respectively with, in this case, $m = (p+1)$ and $n = R + m$.

It will be shown that, for the case of $\varsigma = (R+p)$, (3.17) is feedback linearizable by: (i) constructing the diffeomorphism $\mathcal{T}$ and the linearizing feedback that consists of $\mathbf{p}(\boldsymbol{x})$ and $\boldsymbol{\mathfrak{Q}}(\boldsymbol{x})$, and (ii) assigning the new external inputs $\mathbf{v}$.

By differentiating $\dot{\boldsymbol{x}}_1$ with respect to time,

$$\ddot{\boldsymbol{x}}_1 = \frac{\partial \mathbf{f}_1}{\partial \boldsymbol{x}_1}\,\mathbf{f}_1 + \mathbf{J}\,\mathbf{u}, \qquad \mathbf{J} \equiv \left[(1/\lambda)\tfrac{\partial \mathbf{f}_1}{\partial \boldsymbol{x}_2} \quad (-1/\mathrm{h}(\boldsymbol{x}_2))\tfrac{\partial \mathbf{f}_1}{\partial \lambda}\right], \tag{3.18}$$

where $\mathbf{J}$ is an $R \times m$ matrix. Under the assumption that $\mathbf{J}_s \in \mathbb{R}^{R\times R}$ is invertible with $\mathbf{J} = [\mathbf{J}_s,\, \mathbf{J}_t]$, (3.18) can be solved for $\mathbf{u}$ and, furthermore, $\mathbf{p}(\boldsymbol{x})$, $\boldsymbol{\mathfrak{Q}}(\boldsymbol{x})$, and $\mathbf{v}$ can be constructed accordingly:

$$\mathbf{u} = \begin{bmatrix} -\mathbf{J}_s^{-1}\frac{\partial \mathbf{f}_1}{\partial \boldsymbol{x}_1}\,\mathbf{f}_1 \\ \mathbf{0}_{m-R} \end{bmatrix} + \begin{bmatrix} \mathbf{J}_s^{-1} & -\mathbf{J}_s^{-1}\mathbf{J}_t \\ \mathbf{0}_{m-R\times R} & \mathbf{I}_{m-R} \end{bmatrix} \begin{bmatrix} \ddot{\boldsymbol{x}}_1 \\ \dot{\boldsymbol{x}}_{2,v} \end{bmatrix} =: \mathbf{p}(\boldsymbol{x}) + \mathbf{Q}(\boldsymbol{x})\,\mathbf{v}, \tag{3.19}$$

where $\boldsymbol{x}_2 = \left[\begin{smallmatrix}\boldsymbol{x}_{2,u}\\ \boldsymbol{x}_{2,v}\end{smallmatrix}\right]$ with $\boldsymbol{x}_{2,u}$ and $\boldsymbol{x}_{2,v}$ being $(R+p-m)$- and $(m-R)$-dimensional vectors, respectively. From the choice of the new external inputs $\mathbf{v}$ in (3.19), the linearized system (3.16) is obtained with

$$\boldsymbol{z} \equiv \begin{bmatrix}\boldsymbol{x}_1\\ \dot{\boldsymbol{x}}_1\\ \boldsymbol{x}_{2,v}\end{bmatrix}, \quad \boldsymbol{\mathcal{A}} = \begin{bmatrix}\mathbf{0}_{R\times R} & \mathbf{I}_R & \mathbf{0}_{R\times m-R}\\ \mathbf{0}_{R\times R} & \mathbf{0}_{R\times R} & \mathbf{0}_{R\times m-R}\\ \mathbf{0}_{m-R\times R} & \mathbf{0}_{m-R\times R} & \mathbf{0}_{m-R\times m-R}\end{bmatrix}, \quad \boldsymbol{\mathcal{B}} = \begin{bmatrix}\mathbf{0}_{R\times R} & \mathbf{0}_{R\times m-R}\\ \mathbf{I}_R & \mathbf{0}_{R\times m-R}\\ \mathbf{0}_{m-R\times R} & \mathbf{I}_{m-R}\end{bmatrix}. \tag{3.20}$$

where $\boldsymbol{z}$ is an $(R+m)$-dimensional vector, and $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ matrices of dimension $n \times n$ and $n \times m$, respectively.

The condition of $\mathbf{J}_s$ being invertible is not very restrictive, especially when only independent reactions are considered. If only non-reacting species are added, it is clear that they will not affect the reaction rates, and $\mathbf{J}_s$ will not be invertible. Note that the condition of $\mathbf{J}_s$ being invertible implies that the number of inputs $m \geq R$. Thus, if different reacting species in at least R inlets are fed to the reaction system, the invertibility condition of $\mathbf{J}_s$ can be met easily. Since only R inputs are required to guarantee the invertibility condition of $\mathbf{J}_s$, the remaining $(m-R)$ inputs are superfluous and are, thus, not necessary to control the R reaction variants.

The invertibility condition is only a *sufficient* condition for feedback linearizability. An example illustrating the conservatism of this condition is given next. Consider the semibatch reaction system $\mathcal{X}_1 + \mathcal{X}_2 \rightarrow \mathcal{X}_3$ and $\mathcal{X}_1 + \mathcal{X}_3 \rightarrow \mathcal{X}_4$ with second-order kinetics and $\mathcal{X}_2$ being fed ($m = p = 1$, $R = 2$). Although the invertibility condition is not satisfied, since $m < R$, it can easily be verified that the system is feedback linearizable.

3.3 Extensions

In Section 3.1, the dynamic model in normal form (3.9) was derived for the cases of constant density and temperature. In this section, it is considered for the three extensions: (i) reaction systems with varying density, (ii) nonisothermal reaction systems, and (iii) reaction systems expressed in terms of weight fractions. Based on these extensions, the dynamic reaction model in normal form (3.9) of any combination of the (relaxed) assumptions could easily be derived and are, therefore, omitted here.

3.3.1 Reaction systems with varying density

Even in the case of varying density, $\boldsymbol{n}$ and V in (2.39) can still be transformed as in Theorem 3.2, if $\mathfrak{h}$ is redefined using the following algebra. Substituting $V = \mathfrak{h}\,\lambda$ in $\frac{d}{dt}(\rho V)$ of (2.30) and using (3.9) gives $\frac{d}{dt}(\rho\,\mathfrak{h}) = \mathbf{1}_p^{\mathrm{T}}\,\boldsymbol{\Phi}_{in}\,(\mathbf{q}_{in}/\lambda) = \mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Phi}_{in}\boldsymbol{x}_2$. After integrating both sides, the following redefinition of $\mathfrak{h}$ results:

$$\mathfrak{h}(\boldsymbol{x},\lambda) := (\eta^\star(\rho_0 V_0 - \mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Phi}_{in}\boldsymbol{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}_0) + \mathbf{1}_p^{\mathrm{T}}\,\boldsymbol{\Phi}_{in}\,\boldsymbol{x}_2)/\rho(\boldsymbol{x},\lambda), \tag{3.21}$$

where $\eta^\star$ is an arbitrary integration constant. Since from (3.11) and (3.12) $\mathfrak{g} = \mathfrak{h}(\boldsymbol{x},\lambda)/V$ and using (3.21), $\mathfrak{g}(\boldsymbol{n},V) := \eta^\star(\rho_0 V_0 - \mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Phi}_{in}\boldsymbol{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}_0)/(\rho V - \mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Phi}_{in}\boldsymbol{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n})$.

An important difference between the original definition of $\mathfrak{h}$ in (3.10) and the redefinition (3.21) is that, since ρ depends on both $\boldsymbol{x}$ and λ, the redefined $\mathfrak{h}$ depends not only on $\boldsymbol{x}_2$ but also on λ and the *entire* $\boldsymbol{x}$ vector.

Usually, the inlet mass flowrates $\boldsymbol{\nu}_{in} = \boldsymbol{\Phi}_{in}\mathbf{q}_{in}$ are known, since $\boldsymbol{\Phi}_{in}$ is assumed to be constant and available and the inlet volumetric flowrates $\mathbf{q}_{in}$ are available.

However, since the varying density ρ is usually unknown, the outlet mass flowrate $\nu_{out} = \rho\, q_{out}$ cannot be calculated from q_{out}, and conversely. For ν_{out} being known, $\dot{\lambda}$ in the transformed system (3.9) can be rewritten as:

$$\begin{aligned}\dot{\lambda} &= -q_{out}/\mathfrak{h}(\boldsymbol{x},\lambda)\\ &= -\nu_{out}/(\eta^{\star}(\rho_0 V_0 - \mathbf{1}_p^{\mathrm{T}}\boldsymbol{\Phi}_{in}\boldsymbol{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}_0) + \mathbf{1}_p^{\mathrm{T}}\,\boldsymbol{\Phi}_{in}\,\boldsymbol{x}_2).\end{aligned} \tag{3.22}$$

Thus, λ and also $\boldsymbol{x}_2$ (see 3.9), are independent of ρ, which implies that λ and $\boldsymbol{x}_2$ are reaction invariants.

In summary, the structure of the reaction system in normal form (3.9) is conserved for varying-density reaction systems with the redefinition of $\mathfrak{h}$ given in (3.21). With volumetric flowrates, the interpretation of the three-part decomposition discussed in Section 3.1.4 is not possible due to coupling through the density. Alternatively, if the mass flowrates are available, the interpretation of the three-part decomposition is still valid, since λ and $\boldsymbol{x}_2$ are independent of the density (see 3.22).

3.3.2 Nonisothermal reaction systems

First, the dynamic model in normal form for nonisothermal reaction systems will be derived. Then, state accessibility and linearization by feedback will be studied.

Non-negligible external heat flowrate ($\dot{Q}_{ext} \neq 0$)

Model (2.40) can be transformed to normal form (3.9) with the following redefinitions: $\boldsymbol{n} := \boldsymbol{n}_T$, $\mathbf{N} := \mathbf{N}_T$, $\mathbf{C}_{in} := \mathbf{C}_{T,in}$, $\mathbf{q}_{in} := \mathbf{q}_{T,in}$, $S := S_T$, and $p := p_T$. However, since $\mathfrak{h} = V/\lambda$ and, thus, depends only on the inlet volumetric flowrates:

$$\mathfrak{h}(\boldsymbol{x}_{2,n}) = \eta\,(V_0 - \mathbf{1}_p^{\mathrm{T}}\boldsymbol{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}_0) + \mathbf{1}_p^{\mathrm{T}}\boldsymbol{x}_{2,n}, \tag{3.23}$$

where $\boldsymbol{x}_2 = \left[\begin{smallmatrix}\boldsymbol{x}_{2,n}\\ \mathrm{x}_{2,T}\end{smallmatrix}\right]$ with $\boldsymbol{x}_{2,n}$ and $\mathrm{x}_{2,T}$ corresponding to $\mathbf{q}_{in}$ and $\dot{Q}_{ext}$, respectively.

From the dimensions of $\boldsymbol{x}_1$ (R) and $\boldsymbol{x}_3$ ($S_T - \varsigma$), it can be concluded that the number of reaction variants is the same as in the isothermal case (R), while the dimension of the reaction-invariant and flow-variant space is increased by one ($p_T + 1 = p + 2$). Note that $S_T = (S+1)$ and $\mathrm{rank}\,([\mathbf{N}^{\mathrm{T}}, \mathbf{C}_{T,in}]) = \mathrm{rank}\,([\mathbf{N}_T^{\mathrm{T}}, \mathbf{C}_{in}]) + 1$. Thus, since both the number of (pseudo-)species and ς are increased by one, the number of reaction and flow invariants ($S_T - \varsigma$) is the same as in the isothermal case. This can intuitively be explained by the temperature varying with the external heating independently of both the reactions and the flows. The dimension of the inaccessible part $\boldsymbol{x}_3$ remains the same as in the isothermal case, since the external heat flowrate $\dot{Q}_{ext}$ can be used as an additional input to control the additional state ($V\,T^{\dagger}$).

The dimension of the Jacobian matrix $\mathbf{J}$ (see 3.18) is $R \times (p+2)$, the extra column being due to the external heat flowrate $\dot{Q}_{ext}$ used as an additional input. Thus, a system that

is not feedback linearizable in the isothermal case (an invertible $\mathbf{J}_s$ does not exist) may become feedback linearizable in the nonisothermal case (an invertible $\mathbf{J}_s$ exists). In other words, the external heat flowrate can be used to control the reaction variants in addition to the state $V\,T^{\dagger}$.

Negligible external heat flowrate ($\dot{Q}_{ext} = 0$)
If $\operatorname{rank}([\mathbf{N}^{\mathrm{T}},\, \mathbf{C}_{in}]) = R + p$, then $\operatorname{rank}([\mathbf{N}_T^{\mathrm{T}}\, \mathbf{c}_{T,in}]) = R + p$. Since $(S_T - \varsigma) = (S + 1 - \varsigma)$, the number of reaction and flow invariants is increased by 1. This means that the additional state $(V\,T^{\dagger})$ is inaccessible in the absence of external heat flowrate. Intuitively, the temperature cannot vary independently of the concentrations without external heat flowrate.

3.3.3 Reaction systems expressed in weight fractions

Since (2.42) and (2.38) are structurally the same, the model in normal form (3.9) is also valid for reaction systems expressed in weight fractions by redefining the following quantities: $\boldsymbol{n} := \boldsymbol{m}$, $\mathbf{c} := \mathbf{w}$, $\mathbf{N} := \mathbf{N}_w$, $\mathbf{r} := \mathbf{r}_{\mathfrak{m}}$, $\mathbf{C}_{in} := \mathbf{W}_{in}$, $\mathbf{q}_{in} := \boldsymbol{\nu}_{in}$, $q_{out} := \nu_{out}$, $V := \mathfrak{m}$. However, $\boldsymbol{x}_1$ also directly depends on ρ:

$$\dot{\boldsymbol{x}}_1 := (\mathfrak{h}(\boldsymbol{x}_2)/\rho(\boldsymbol{x}, \lambda))\, \mathbf{N}_w^{\mathrm{T}} \mathbf{r}_x(\boldsymbol{x}). \tag{3.24}$$

3.4 Special cases

In this section, the dynamic model in normal form (3.9) is considered for the two special cases: (i) CSTRs, semibatch and batch reaction systems, and (ii) reactions in quasi-equilibrium conditions. These special cases are based on the assumptions of constant density and temperature, but could easily be extended to the cases of varying density and/or temperature.

3.4.1 CSTR, semibatch and batch reaction systems

In the derivation of the model in normal form (3.9), it was assumed that the inlets and the outlet are independent. In certain important special cases, however, q_{out} is not independent, for example, for (i) CSTRs ($q_{out} = \mathbf{1}_p^{\mathrm{T}} \mathbf{q}_{in}$), (ii) semibatch reaction systems ($q_{out} = 0$), and (iii) batch reaction systems ($q_{out} = 0$, $\mathbf{q}_{in} = 0$). Below, the dynamic models in normal form will be derived for these cases. Since the initial conditions of the dynamic models in normal form can easily be derived from (3.9), they are suppressed to simplify readability. Then, the dimension of the reaction-invariant and flow-variant space and state accessibility will be analyzed for CSTRs, semibatch and batch reaction systems.

CSTR

Since the volume remains constant, the nonlinear transformation (3.8) of V to λ is not required to get to normal form and, thus, a transformation similar to (3.6) can be used

to transform system (2.56):

$$\dot{\boldsymbol{x}}_1 = \mathsf{h}(\boldsymbol{x}_2)\,\mathbf{r}_{\mathsf{x}}(\boldsymbol{x}), \quad \dot{\boldsymbol{x}}_2 = \mathsf{h}(\boldsymbol{x}_2)\,\mathbf{q}_{in}/V, \quad \dot{\boldsymbol{x}}_3 = \mathbf{0}_{S-\varsigma}, \quad \dot{V} = 0. \tag{3.25}$$

The forward transformation (3.11) still holds. However, since the reconstruction of $\boldsymbol{n}$ in (3.12) depends on λ, which is undefined in (3.25), the reconstruction of $\boldsymbol{n}$ for CSTRs becomes:

$$\boldsymbol{n} = \frac{V}{\mathsf{h}(\boldsymbol{x}_2)}\,(\mathbf{N}^{\mathrm{T}}\,\boldsymbol{x}_1 + \mathbf{C}_{in}\,\boldsymbol{x}_2 + \boldsymbol{\Omega}\boldsymbol{x}_3). \tag{3.26}$$

Since $\dot{V} = 0$, the number of reaction and flow invariants is increased by one $(S-\varsigma+1)$.

Semibatch reaction systems

The model in normal form (3.9) simplifies, since $\dot{\lambda} = 0$ implies $\lambda(t) = \lambda_0$ and $V = \lambda_0\,\mathsf{h}(\boldsymbol{x}_2)$. Thus, for (2.57),

$$\dot{\boldsymbol{x}}_1 = \mathsf{h}(\boldsymbol{x}_2)\,\mathbf{r}_{\mathsf{x}}(\boldsymbol{x}), \quad \dot{\boldsymbol{x}}_2 = \mathbf{q}_{in}/\lambda, \quad \dot{\boldsymbol{x}}_3 = \mathbf{0}_{S-\varsigma}, \quad \dot{\lambda} = 0. \tag{3.27}$$

Note that the transformation given by (3.11)–(3.12) is linear, since $\mathsf{g}(\boldsymbol{n}, V) = 1/\lambda_0 = \not{c}$. Since $\dot{\lambda} = 0$, the number of reaction and flow invariants is increased by one $(S-\varsigma+1)$.

Batch reaction systems

The model in normal form (3.27) simplifies, since $V(t) = V_0$ and $\mathbf{q}_{in} = 0$ implies $\mathsf{h}(\boldsymbol{x}_2) = \eta\,(V_0 - \mathbf{1}_p^{\mathrm{T}}\mathfrak{M}^{\mathrm{T}}\boldsymbol{n}_0)$ and $\varsigma = R$. Thus, for (2.58),

$$\dot{\boldsymbol{x}}_1 = \eta\,(V_0 - \mathbf{1}_p^{\mathrm{T}}\mathfrak{M}^{\mathrm{T}}\boldsymbol{n}_0)\,\mathbf{r}_{\mathsf{x}}(\boldsymbol{x}), \quad \dot{\boldsymbol{x}}_3 = \mathbf{0}_{S-R}, \quad \dot{\lambda} = 0. \tag{3.28}$$

Since $\dot{\lambda} = 0$, the number of reaction and flow invariants is increased by one $(S-R+1)$.

Discussion

As discussed in Section 3.1.4, the reaction-invariant and flow-variant space has a maximum dimension of $(\varsigma - R + 1) \leq p + 1$. In the special cases of CSTRs, and semibatch and batch reaction systems, that space cannot have dimension larger than p, since there are only p independent flows. Combining these two elements, the dimension of the reaction-invariant and flow-variant space is determined as $\min(\varsigma - R + 1, p)$.

For the special cases considered above, the dimension of the reaction-invariant and flow-variant space $\min(\varsigma - R + 1, p)$ equals p for either of the following two cases: (i) $\varsigma = (R + p)$, and (ii) $\varsigma = R + p - 1$. This means that the number of accessible states is not reduced even when the concentrations of one inlet stream are dependent on the stoichiometry and the concentrations of the remaining inlet streams. Intuitively, this is due to the volume becoming accessible in the case of $\varsigma = (R + p - 1)$ as discussed below.

The addition of solvent is considered as an example for $\varsigma = R + p - 1$. The inlet concentration matrix has the form $\mathbf{C}_{in} := [\mathbf{c}_{in}^1, \ldots, \mathbf{c}_{in}^{(p-1)}, \mathbf{0}_S]$. The last column of $\mathbf{C}_{in}$ is zero, since the solvent is not one of S reacting species. If it is assumed that

rank $\left([\mathbf{N}^{\mathrm{T}},\, \mathbf{c}_{in}^{1},\, \ldots,\, \mathbf{c}_{in}^{(p-1)}\right)] = R + p - 1$, then rank $([\mathbf{N}^{\mathrm{T}},\, \mathbf{C}_{in}]) = R + p - 1$. Intuitively, when the outlet is not independent, the addition of solvent changes the volume, thereby diluting other species and affecting $\boldsymbol{n}$ and V. Note that, for the tangent-linearized model with no outlet stream and rank $\left([\mathbf{N}^{\mathrm{T}},\, \mathbf{c}_{in}^{1},\, \ldots,\, \mathbf{c}_{in}^{(p-1)}\right)] = R + p - 1$, local controllability of the volume by solvent addition has been proven by Dochain and Chen (1992). However, the work presented herein proposes *global* results for *nonlinear* chemical reaction systems.

3.4.2 Systems with reactions in quasi-equilibrium conditions

The reparameterization of systems with reactions in quasi-equilibrium conditions (see 2.66) leads to the following dynamic model in normal form:

$$\frac{d\boldsymbol{x}_1}{ds} = \mathrm{h}(\boldsymbol{x}_2)\, \mathbf{r}_{\mathrm{X}}(\boldsymbol{x})/\dot{s}, \quad \frac{d\boldsymbol{x}_2}{ds} = \mathbf{q}_{in}/(\lambda\, \dot{s}), \quad \frac{d\boldsymbol{x}_3}{ds} = \mathbf{0}_{S-\varsigma}, \quad \frac{d\lambda}{ds} = -q_{out}/(\mathrm{h}(\boldsymbol{x}_2)\, \dot{s}), \tag{3.29}$$

where the initial conditions of $\boldsymbol{x}$ and λ remain the same as in (3.9).

3.5 Summary

A generic nonlinear transformation to normal form for reaction systems with inlet and outlet streams was presented in this chapter. It leads to a normal form in which the reaction and flow invariants can be easily identified.

Based on the dynamic model in normal form, it was shown that order reduction of the dynamic model is possible under certain conditions, thereby enabling more efficient simulation and design of controllers, observers, and optimizers. State accessibility, a key concept in nonlinear control theory, was analyzed without the knowledge of kinetics. A linearizing feedback was constructed in a straightforward manner, which holds under minor assumptions on the kinetics.

The results obtained for constant-density, isothermal reaction systems (basic model) were extended to varying-density and non-isothermal situations, and to reaction systems expressed in weight fractions. Furthermore, special cases were studied such as CSTRs, semibatch and batch reaction systems. Systems with reactions in quasi-equilibrium conditions were also considered. Note that the results obtained are *global* in nature as they deal with the *full-fledged nonlinear* system and consider general expressions for the reaction rates (no explicit knowledge about the kinetic structure; this is also valid for the extensions and the special cases). In summary, major insight is gained by using the normal form which, in turn, can help perform simulation, analysis, control, and optimization more efficiently. It will be shown in the next two chapters that the proposed transformation to normal form is the basis for the analysis of concentration and spectral measurements from reaction systems.

4

Reaction and flow variants/invariants in the factorization of concentration data

Chemical analysis is a critically important enabling technology essential to every phase of chemical science, product and process development, and manufacturing control (American Chemical Society *et al.*, 1996). New knowledge and insight provided through chemical measurement greatly accelerate progress in chemical science, biotechnology, materials science, and process engineering by providing reliable data to evaluate chemical, biotechnological and pharmaceutical reaction processes.

In practice, two types of chemical measurements are commonly found: weight fraction and molar concentration. In most standard wet-chemical analyses, the composition is measured in weight fractions. Alternatively, many types of spectral measurements provide indirect information about molar concentrations (see Chapter 5).

The techniques discussed here are applicable to both molar concentration and weight fraction measurements, though only the former will be discussed in detail. Also, the case of unmeasured species is investigated. Such a situation can occur when some species either are difficult to measure or do not contain any additional information.

For the analysis of concentration data from reaction systems, a *factorization of concentration data* will be derived from the transformation to normal form proposed in the previous chapter. Concentration data will be factorized into three parts: (i) the reaction-variant part (related to the reactions), (ii) the reaction-invariant and flow-variant part (related to the inlet and outlet streams), and (iii) the reaction- and flow-invariant part (related to the initial concentrations and initial volume). Based on this factorization, it will be shown that the concentration data lie in a *hyperplane* of typically lower dimension than the number of species. Thus, *linear* statistical methods for the analysis of dynamic and/or static information are directly applicable.

In many practical applications, it is useful to separate the information related to the reaction-variant part from that related to the reaction-invariant part. This goal is achieved by *data pre-treatment to reaction-variant form* that subtracts the reaction-invariant part (the material exchange terms and the initial conditions) from measured concentrations, leaving the reaction-variant part. The resulting *factorization of concentration data in reaction-variant form* has the additional advantage that it separates

information related to the reaction dynamics (reaction variants) from static information (stoichiometry). Data in reaction-variant form also lie in a *hyperplane* of lower dimension than the number of species, and *linear* statistical methods are also applicable to such data. The two methods that are studied in this chapter are: (i) target factor analysis (TFA) for the determination of the stoichiometries (static information) and (ii) on-line reconstruction of concentration of unmeasured species *via* the reconstruction of the reaction variants (dynamic information).

The kinetic description still represents the main difficulty in modeling chemical reaction systems. For most industrially-relevant reactions, the kinetic parameters cannot be estimated reliably from theory and, thus, must be determined experimentally. This estimation requires the system stoichiometry and candidate kinetic structure(s) to be known. TFA has been used successfully with reaction data to determine, *without knowledge of reaction kinetics*, the number of independent reactions and the corresponding stoichiometries. Here, the applicability of TFA techniques to reaction data on the basis of concentration measurements is analyzed.

Often in practice, not all species of interest are measured. In cases where the reaction variants can be computed from the concentrations of a few measured species, it will be shown that the concentrations of the remaining species (of interest) can be reconstructed using the known reaction-invariant part. For both TFA and on-line state reconstruction, the rank of the concentration data and methods that change the dimension of the null space of selected matrices will play an important role.

For pedagogical reasons, data from a single isothermal constant-density reaction run are considered first. Then, extensions to data from nonisothermal, varying-density reaction systems and to process runs with different stoichiometries are provided. Furthermore, the following special cases are considered: CSTRs, semibatch reaction systems, systems with reactions in quasi-equilibrium conditions, reaction systems expressed in weight fractions, and non-reacting data with closure.

The chapter is organized as follows: Section 4.1 introduces (i) the factorization of concentration data, (ii) the experimental concentration matrix, (iii) data pre-treatment to reaction-variant form and the factorization of concentration data in reaction-variant form, and (iv) the factorization of concentration data for the case of unmeasured species. Section 4.2 investigates the rank and the dimension of the fundamental subspaces of the different data matrices, and it proposes methods to change the dimension of null spaces. Section 4.3 studies on-line state reconstruction and stoichiometric modeling using TFA. Extensions to the basic factorization are provided in Section 4.4, and special cases are studied in Section 4.5.

4.1 Factorization of concentration data

4.1.1 Derivation of the factorization of concentration data and extended stoichiometry

Let $\mathbf{c}(k)$ be the concentrations measured at the kth time instant. Using (3.13), it can be expressed as:

$$\mathbf{c}^{\mathrm{T}}(k) = \frac{1}{\mathrm{h}(\boldsymbol{x}_2(k))} \left[\begin{array}{c|c|c} \boldsymbol{x}_1^{\mathrm{T}}(k) & \boldsymbol{x}_2^{\mathrm{T}}(k) & 1 \end{array}\right] \left[\begin{array}{c} \mathbf{N} \\ \hline \mathbf{C}_{in}^{\mathrm{T}} \\ \hline \boldsymbol{x}_3^{\mathrm{T}}\boldsymbol{\Omega}^{\mathrm{T}} \end{array}\right], \tag{4.1}$$

where the matrix $\left[\begin{array}{c} \mathbf{N} \\ \mathbf{C}_{in}^{\mathrm{T}} \\ \boldsymbol{x}_3^{\mathrm{T}}\boldsymbol{\Omega}^{\mathrm{T}} \end{array}\right]$ is constructed such that (i) it contains only *constant* quantities and (ii) the vector $[\boldsymbol{x}_1^{\mathrm{T}}(k),\, \boldsymbol{x}_2^{\mathrm{T}}(k),\, 1]$ contains a minimum number of elements. In (4.1), $\mathbf{c}(k)$ depends *implicitly* on the initial concentrations $\mathbf{c}_0$ and volume V_0 through the initial conditions of $\boldsymbol{x}$ and λ in (3.9). Thus, the initial conditions $\mathbf{c}_0$ and V_0 need to be stored in the space spanned by the $(S-\varsigma)$ columns of $\boldsymbol{\Omega}$ (basis vectors).

A transformation is derived next that leads to reconstructed concentrations $\mathbf{c}(k)$ that depends *explicitly* on the initial conditions $\mathbf{c}_0$ and V_0. This way, the initial conditions $\mathbf{c}_0$ and V_0 are stored in the one-dimensional space spanned by $\mathbf{c}_0$ and V_0. This transformation is based on a difference formulation for $\boldsymbol{x}$ in (3.9) with respect to the corresponding initial conditions. The arbitrary nonzero constant η is chosen as $\eta = 1$:

$$\begin{aligned} \mathbf{x}(k) &\equiv \boldsymbol{x}_1(k) - V_0\, \mathbf{N}^{+\mathrm{T}} \left(\mathbf{I}_S - \mathbf{C}_{in}\boldsymbol{\mathcal{M}}^{\mathrm{T}}\right) \mathbf{c}_0, \\ \mathbf{z}(k) &\equiv \boldsymbol{x}_2(k) - V_0\, \boldsymbol{\mathcal{M}}^{\mathrm{T}}\, \mathbf{c}_0, \\ \mathbf{0}_{S-\varsigma} &= \boldsymbol{x}_3 - V_0\, \boldsymbol{\Omega}^{\mathrm{T}}\, \mathbf{c}_0, \end{aligned} \tag{4.2}$$

where $\mathbf{x}$ is the R-dimensional vector of reaction variants and $\mathbf{z}$ the p-dimensional vector of reaction invariants *and* flow variants. Note that $\eta = 1$ implies $\lambda(0) = 1$.

Theorem 4.1
Let Assumptions A1–7 in Appendix D be verified. The concentrations $\mathbf{c}(k)$ *from reaction systems described by* (2.38) *are factorized as*

$$\mathbf{c}^{\mathrm{T}}(k) = \frac{1}{\mathrm{h}(\mathbf{z}(k))} \left[\begin{array}{c|c|c} \mathbf{x}^{\mathrm{T}}(k) & \mathbf{z}^{\mathrm{T}}(k) & V_0 \end{array}\right] \left[\begin{array}{c} \mathbf{N} \\ \hline \mathbf{C}_{in}^{\mathrm{T}} \\ \hline \mathbf{c}_0^{\mathrm{T}} \end{array}\right] \equiv \mathbf{x}_e^{\mathrm{T}}\mathbf{N}_e \tag{4.3}$$

with $\mathrm{h}(\mathbf{z}(k)) \equiv V_0 + \mathbf{1}_p^{\mathrm{T}}\mathbf{z}(k)$,

where $\mathbf{x}_e$ *is an* $(R+p+1)$*-dimensional vector and* $\mathbf{N}_e$ *is the* extended stoichiometric matrix *of dimension* $(R+p+1)\times S$. (4.3) *is termed* the factorization of concentration data. *For* K *observations, the factorization of concentration data becomes:*

$$\mathbf{C} = \left\{\boldsymbol{\mathcal{H}}^{-1} \left[\begin{array}{c|c|c} \mathbf{X} & \mathbf{Z} & V_0\, \mathbf{1}_K \end{array}\right]\right\} \left[\begin{array}{c} \mathbf{N} \\ \hline \mathbf{C}_{in}^{\mathrm{T}} \\ \hline \mathbf{c}_0^{\mathrm{T}} \end{array}\right] = \mathbf{X}_e\, \mathbf{N}_e \tag{4.4}$$

with $\boldsymbol{\mathcal{H}} \equiv V_0\, \mathbf{I}_K + \mathrm{diag}\,(\mathbf{Z}\, \mathbf{1}_p)$

and $\mathbf{X}$ *being the* $K \times R$ *matrix of reaction variants,* $\mathbf{Z}$ *the* $K \times p$ *matrix of reaction invariants* and *flow variants,* $V_0\mathbf{1}_K$ *the* K*-dimensional vector of reaction and flow invariants,* $\mathbf{X}_e$ *an* $K \times (R+p+1)$ *matrix, and* $\mathcal{H}$ *a* K*-dimensional diagonal scaling matrix computed from* (4.5) *for* K *observations. (See Appendix H.3.1 for proof)*

The main advantage of deriving the factorization of concentration data (4.4) from the transformation to normal form (3.9) is the availability of a strong interpretation of the different terms in (4.3). This is in contrast to an alternate derivation of (4.3) given in Appendix E.1 which derives (4.3) directly from (2.38). Note that the derivation of (4.3) (and 4.4) is independent of the knowledge of kinetics. Furthermore, only in the special case of reaction systems with no outlet ($q_{out} = 0$, $\lambda(t) = 1$) is the reaction extent $x_{b,i}$ equal to the reaction variant x_i.

$V(k)$ and $\lambda(k)$ are related by (3.12): $V(k) = \mathfrak{h}(\mathbf{x}_2(k))\,\lambda(k)$ with $\mathfrak{h}(\mathbf{x}_2(k))$ defined as in (3.10). With the choice of $\eta = 1$, a similar relation for $\mathfrak{h}$ can be found with respect to the new coordinates ($\mathbf{x}$ and $\mathbf{z}$):

$$V(k) = \mathfrak{h}(\mathbf{z}(k))\,\lambda(k), \tag{4.5}$$

or for $\mathcal{H}$,

$$\mathcal{H} = \mathbf{V}\,\mathbf{\Lambda}^{-1}, \tag{4.6}$$

where $\mathbf{V} \equiv \operatorname{diag}(V(1), V(2), \ldots, V(K))$ and $\mathbf{\Lambda} \equiv \operatorname{diag}(\lambda(1), \lambda(2), \ldots, \lambda(K))$.

Since concentration data are generated from a reaction system, $\mathbf{x}$ and $\mathbf{z}$ in the factorization of concentration data (4.4) *implicitly* follow the following set of differential equations:

$$\begin{aligned} \dot{\mathbf{x}} &= \mathfrak{h}(\mathbf{z})\,\mathbf{r}(\mathbf{c}), & \mathbf{x}(0) &= \mathbf{0}_R, \\ \dot{\mathbf{z}} &= \mathbf{q}_{in}/\lambda, & \mathbf{z}(0) &= \mathbf{0}_p, \\ \dot{\lambda} &= -q_{out}/\mathfrak{h}(\mathbf{z}), & \lambda(0) &= 1, \end{aligned} \tag{4.7}$$

where (4.7) is obtained by differentiating (4.2) with respect to time, and transforming the initial conditions with $\eta = 1$.

Note that $\mathbf{x}$, $\mathbf{z}$, and $\mathbf{c}_0$ are quantities relative to the initial time t_0. Thus, $\mathbf{X}_e$ and $\mathbf{N}_e$ are defined with respect to t_0, and changing t_0 change both $\mathbf{X}_e$ and $\mathbf{N}_e$.

The results of Theorem 4.1 are summarized in Table 4.1. The factorization of concentration data (4.4) gives a complete description of the spaces in which the concentrations lie: the R-dimensional reaction-variant space ($\mathcal{S}_c(\mathbf{N}^{\mathrm{T}})$), the p-dimensional reaction-invariant and flow-variant space ($\mathcal{S}_c(\mathbf{C}_{in})$), and the one-dimensional reaction- and flow-invariant space ($\mathcal{S}_c(\mathbf{c}_0)$). Thus, the concentrations $\mathbf{c}(k)$ lie in a lower-dimensional *hyperplane* spanned by the rows of $\mathbf{N}_e$ ($\mathcal{S}_r(\mathbf{N}_e)$). For most varying-volume reaction systems, the hyperplane contains the origin, as it is illustrated by Figure 4.1 (see 4.5.2 for a special case where the origin is not contained).

The variability of $\mathbf{C}$ is typically determined by its rank (fourth row in Table 4.1). Since the three spaces are non-orthogonal, the rank of $\mathbf{C}$ can be less than $\min(R + p + 1, S)$ (see Section 4.2 below). Furthermore, the rows of $\mathcal{H}^{-1}\mathbf{X}$ are the coordinates in the

Table 4.1. Overview of the factorization of concentration data.

Original space	Reaction variants	Reaction invariants *and* flow variants (Inlets)	Reaction and flow invariants	Reaction invariants *and* flow variants (Outlet)
	$\mathbf{x}$	$\mathbf{z}$	V_0, $\mathbf{c}_0$	λ
$\mathbf{C} =$	$(\mathbf{X}\,\mathbf{N}$	$+\ \boldsymbol{\mathcal{H}}^{-1}\times \mathbf{Z}\,\mathbf{C}_{in}^{\mathrm{T}}$	$+\ V_0\mathbf{1}_K\,\mathbf{c}_0^{\mathrm{T}})$	
$\mathrm{rank}\,(\mathbf{C})$	R	$\min(p, S)$	1	
$\mathbf{V} =$		$\boldsymbol{\mathcal{H}}$		$\cdot\ \boldsymbol{\Lambda}$

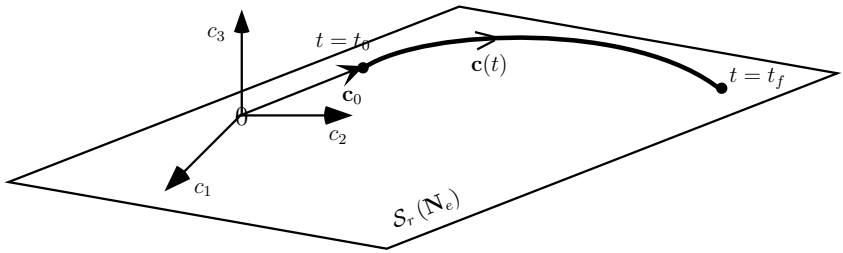

Figure 4.1. Geometric interpretation of the factorization of concentration data. Concentrations $\mathbf{c}$ evolve in a lower-dimensional hyperplane spanned by the rows of $\mathbf{N}_e$ ($\mathcal{S}_r(\mathbf{N}_e)$).

reaction-variant space, the rows of $\boldsymbol{\mathcal{H}}^{-1}\mathbf{Z}$ the coordinates in the reaction-invariant and flow-variant space, and the rows of $V_0\boldsymbol{\mathcal{H}}^{-1}\mathbf{1}_K$ the coordinates in the reaction- and flow-invariant space. Factorization (4.4) separates the dynamics ($\mathbf{X}_e$) from the static part ($\mathbf{N}_e$). However, experimental conditions influence directly both $\mathbf{X}_e$ and $\mathbf{N}_e$ (see also next subsection).

The dimensions of the reaction-variant space (R) and reaction-invariant and flow-variant space (p; see Table 3.1) correspond to their ranks. However, the dimension of the space in which the reaction and flow invariants $\boldsymbol{x}_3$ ($S - \varsigma$) lie differs from its rank (one), since the reaction and flow invariants are constant.

Example 2.13a and Example 2.2b below illustrate the different terms and the geometric interpretation of the factorization of concentration data by means of simulated reaction systems, respectively.

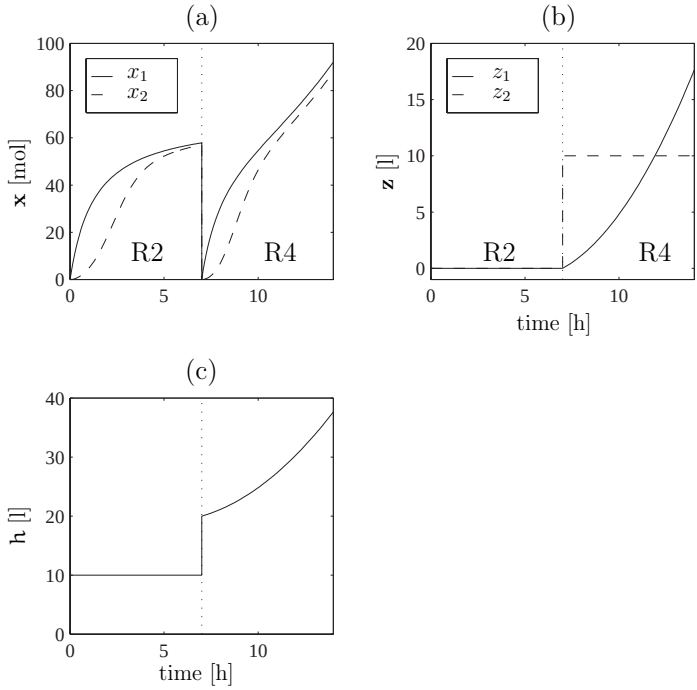

Figure 4.2. (a) The reaction-variant vector $\mathbf{x}$, (b) the reaction-invariant and flow-variant vector $\mathbf{z}$, and (c) the scaling factor $\mathtt{h}$ of Example 2.13a. For legends of $\mathbf{c}$ and $\mathbf{x}$, see Figure 2.5 (Runs R2 and R4 for $t \in [0,\ 7]$h and $t \in]7,\ 14]$h, respectively).

Example 2.13a (Extended stoichiometry; cont'd from page 31**)**
The different terms of the factorization of concentration data are illustrated by means of Case C3 of Example 2.13 ($R = 2$, $S = 5$). Figures 4.2 presents the time evolution of the reaction-variant vector $\mathbf{x}$, the reaction-invariant and flow-variant vector $\mathbf{z}$, and the scaling factor h.
The reaction variants $\mathbf{x}$ are strongly monotonically increasing functions, which is characteristic of irreversible reactions. $\mathbf{z}$ is zero for the batch run R2 and nonzero for the additional semibatch run R4. A step change can be observed in z_1 at the time the second run is considered. As expected, z_2 strongly monotonically increases due to continuous feeding.

Example 2.2b (Evolution of c in a hyperplane; cont'd from page 18**)**
Consider a constant-density esterification reaction with one inlet stream and one outlet stream ($S = 5$). The initial concentrations are $\mathbf{c}_0 = [2, 11, 0, 0.2, 0.01]^{\mathrm{T}}$ and the inlet concentrations $\mathbf{C}_{in} = 2\,\mathbf{c}_0$, i.e., the initial and inlet concentrations are linearly dependent. Figure 4.3 shows the evolution of the concentrations, the flowrates and the volume.
Owing to the linear dependence of the initial and inlet concentrations, $\mathcal{S}_r(\mathbf{N}_e) = \mathcal{S}_r\left(\begin{bmatrix}\mathbf{N}\\ \mathbf{c}_0^{\mathrm{T}}\end{bmatrix}\right)$. Thus, the concentrations evolve on a two-dimensional plane of the five-dimensional concentration space c_1–c_5. Figure 4.4 illustrates graphically this evolution in the three-dimensional space c_1–c_3.

4.1.2 Experimental concentration matrix

The evolution of the concentrations is determined by the R reaction variants and $(p+1)$ reaction invariants. The reaction variants are inherent to the chemical reaction systems, while the reaction invariants can be manipulated (or varied) externally by the choice of the experimental (or operational) conditions. These conditions are summarized in the $(p+1) \times S$ *experimental concentration matrix* $\mathbf{C}_x$ and the corresponding $K \times (p+1)$

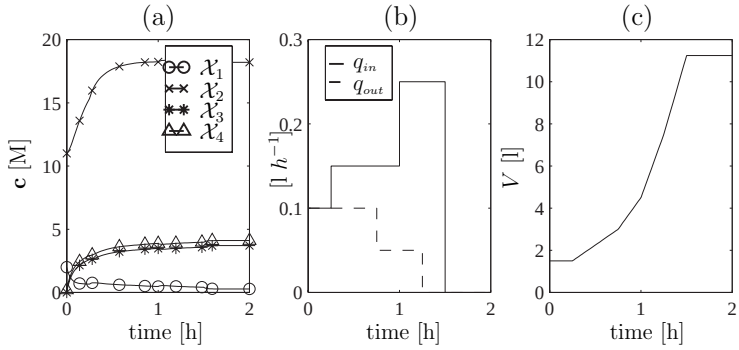

Figure 4.3. Example 2.2b: (a) The concentrations $\mathbf{c}$, (b) the flowrates q_{in} and q_{out}, and (c) the volume V.

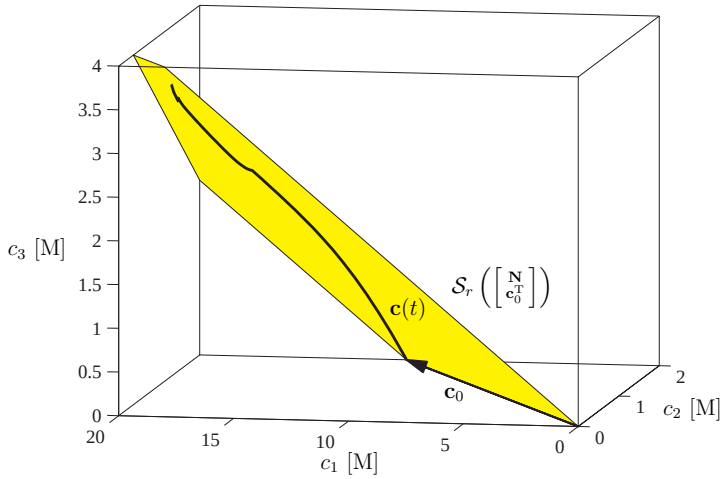

Figure 4.4. Example 2.2b: Geometric interpretation of the factorization of concentration data. Evolution of the concentrations of Species $\mathcal{X}_1$, $\mathcal{X}_2$, and $\mathcal{X}_3$ on a two-dimensional plane.

matrix $\mathbf{Z}_x$ including the reaction invariants:

$$\mathbf{C}_x \equiv \begin{bmatrix} \mathbf{C}_{in}^{\mathrm{T}} \\ \hline \mathbf{c}_0^{\mathrm{T}} \end{bmatrix}, \qquad \mathbf{Z}_x \equiv \left[\begin{array}{c|c} \mathbf{Z} & V_0 \mathbf{1}_K \end{array} \right]. \tag{4.8}$$

With these definitions,

$$\mathbf{C} = \mathbf{X}_e \mathbf{N}_e \quad \text{with} \quad \mathbf{N}_e = \begin{bmatrix} \mathbf{N} \\ \hline \mathbf{C}_x \end{bmatrix}, \quad \mathbf{X}_e = \mathcal{H}^{-1} \left[\begin{array}{c|c} \mathbf{X} & \mathbf{Z}_x \end{array} \right]. \tag{4.9}$$

Let S_x denote the number of species that can be varied *via* the initial concentrations $\mathbf{c}_0$ or the inlet concentrations $\mathbf{C}_{in}$ or, formulated in mathematical terms, have nonzero elements in $\mathbf{C}_x$. In practical applications, often only $S_x < S$ species are available for variation, as it will be discussed below.

4.1.3 Factorization of concentration data in reaction-variant form

In this subsection, concentration data is transformed to reaction-variant form in order to isolate the reaction-variant part. The resulting factorization of concentration data in reaction-variant form is also studied.

Theorem 4.2
Let Assumptions A1–7 in Appendix D be verified. If $\mathbf{C}_{in}$, V_0, $\mathbf{c}(k)$, $\mathbf{q}_{in}(k)$, *and* $q_{out}(k)$ *are known/measured, then an* S*-dimensional concentration vector in reaction-variant*

form (RV-concentration vector), $\mathbf{d}(k)$*, that relates directly to* $\mathbf{x}(k)$ *and* $\mathbf{N}$ *can be defined from* (4.3)*:*

$$\mathbf{d}(k) \equiv \mathrm{h}(\mathbf{z}(k))\,\mathbf{c}(k) - \mathbf{C}_{in}\,\mathbf{z}(k) - V_0\,\mathbf{c}_0 = \mathbf{N}^{\mathrm{T}}\,\mathbf{x}(k). \tag{4.10}$$

The operation to this effect is termed data pre-treatment to reaction-variant form. *For* K *observations, a* $K \times S$ *concentration matrix in reaction-variant form (RV-concentration matrix),* $\mathbf{D}$*, can be constructed from* (4.10) *that relates directly to* $\mathbf{X}$ *and* $\mathbf{N}$*:*

$$\mathbf{D} \equiv \mathcal{H}\mathbf{C} - \mathbf{Z}\,\mathbf{C}_{in}^{\mathrm{T}} - V_0\,\mathbf{1}_K\,\mathbf{c}_0^{\mathrm{T}} = \mathcal{H}\mathbf{C} - \mathbf{Z}_x\,\mathbf{C}_x = \mathbf{X}\,\mathbf{N}. \tag{4.11}$$

(4.11) *is termed* factorization of concentration data in reaction-variant form (factorization of RV-concentration data). *(See Appendix H.3.1 for proof)*

The factorization of RV-concentration data (4.11) is particularly useful as it enables the separation of information related to the reaction dynamics ($\mathbf{X}$) from structural information ($\mathbf{N}$). The experimental conditions such as the material exchange terms and the initial conditions affect the pre-treated data indirectly through the reaction variants. However, their *direct* contribution is removed and, thus, the rows of the RV-concentration matrix, $\mathbf{d}(k)$, lie in an R-dimensional *hyperplane* spanned by the rows of $\mathbf{N}$ ($\mathcal{S}_r(\mathbf{N})$; see Figure 4.5). In other words, removing the experimental conditions is equivalent to removing the reaction-invariant part, i.e., leaving only the reaction-variant part in $\mathbf{D}$.

The crucial point in data pre-treatment to reaction-variant form is that the volume must be known. If the density is constant, the volume can be reconstructed from (2.38) upon knowledge of V_0, $\mathbf{q}_{in}$, and q_{out}, which are known by assumption (see Theorem 4.2).

If data pre-treatment to reaction-variant form is not possible due to $\mathbf{q}_{in}$ and q_{out} being unknown (e.g., unknown disturbance such as leakage), the concentration matrix $\mathbf{C}$ must be used directly.

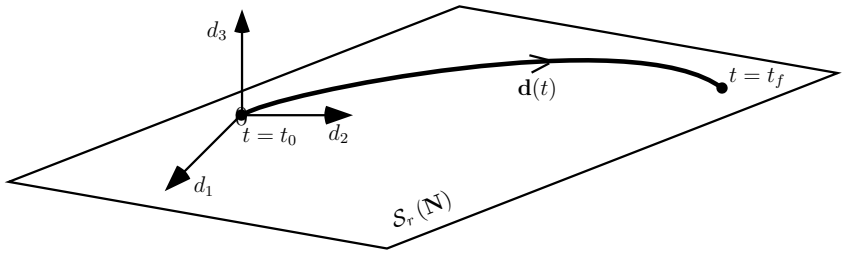

Figure 4.5. Geometric interpretation of the factorization of concentration data in reaction-variant form. RV-concentrations $\mathbf{d}$ evolve in an R-dimensional hyperplane spanned by the rows of $\mathbf{N}$ ($\mathcal{S}_r(\mathbf{N})$).

An alternate data pre-treatment that removes the reaction invariants is provided in Appendix E.2. It is based directly on (2.38). However, similarly to the alternative derivation of (4.3), it lacks the interpretation of the reaction variants.

Example 2.2c (Evolution of d in a hyperplane; cont'd from page 59**)**
Figure 4.6 shows the evolution of the RV-concentrations. Since there is only one single independent reaction, the RV-concentrations evolve on a line of the five-dimensional RV-concentration space d_1–d_5 (see Figure 4.7 in the three-dimensional space d_1–d_3).

4.1.4 Case of unmeasured species

First, some useful notation will be introduced. For both a particular species subset and its cardinality, the same notation is used. Thus, S_b has two meanings depending on the context: (i) "S_b species" means the set of species tagged by S_b, and (ii) "$a + S_b$" means the sum of a and the cardinality S_b of the particular species subset S_b.

An important case that will be dealt with herein is when only $S_m < S$ reacting species are measured because the remaining $(S - S_m)$ species are either difficult to measure or do not contain any additional information. The subscript m denotes that part of a quantity which corresponds to the S_m measured species (known part), and subscript u that part of a quantity which corresponds to the $S_u = (S - S_m)$ unmeasured species (unknown part). Without loss of generality, it is assumed that the first S_m species are measured. The known and unknown species also provide a partition in $\mathbf{C}$, $\mathbf{C}_{in}$, $\mathbf{c}_0$, $\mathbf{D}$, $\mathbf{N}$, $\mathbf{N}_e$, and $\mathbf{\Gamma}$, i.e., $\mathbf{C} = [\mathbf{C}_m, \mathbf{C}_u]$, $\mathbf{C}_{in} = \begin{bmatrix} \mathbf{C}_{in,m} \\ \mathbf{C}_{in,u} \end{bmatrix}$, $\mathbf{c}_0 = \begin{bmatrix} \mathbf{c}_{0,m} \\ \mathbf{c}_{0,u} \end{bmatrix}$, $\mathbf{D} = [\mathbf{D}_m, \mathbf{D}_u]$, $\mathbf{N} = [\mathbf{N}_m, \mathbf{N}_u]$, $\mathbf{N}_e = [\mathbf{N}_{e,m}, \mathbf{N}_{e,u}]$, and $\mathbf{\Gamma} = \begin{bmatrix} \mathbf{\Gamma}_m \\ \mathbf{\Gamma}_u \end{bmatrix}$. Also, note that

$$\mathbf{C}_m = \mathbf{X}_e \, \mathbf{N}_{e,m}, \qquad \mathbf{D}_m = \mathcal{H} \mathbf{C}_m - \mathbf{Z} \, \mathbf{C}_{in,m}^{\mathrm{T}} - V_0 \, \mathbf{1}_K \, \mathbf{c}_{0,m}^{\mathrm{T}} = \mathbf{X} \, \mathbf{N}_m. \tag{4.12}$$

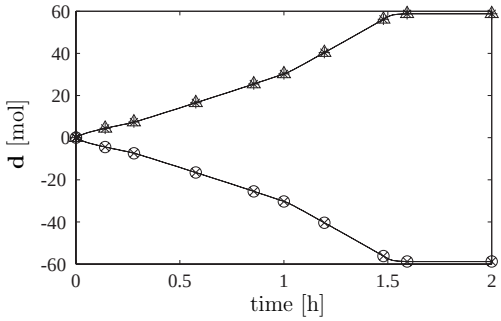

Figure 4.6. The RV-concentrations **d** of Example 2.2c. For legend, see Figure 4.3.

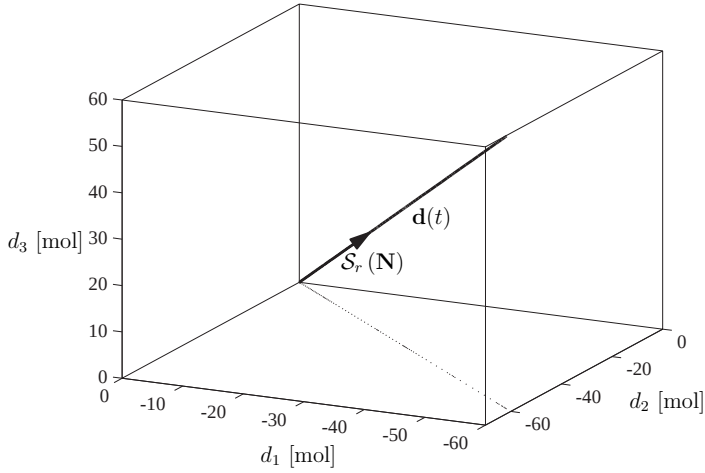

Figure 4.7. Example 2.2c: Geometric interpretation of the factorization of RV-concentration data. Evolution of the RV-concentrations of Species $\mathcal{X}_1$, $\mathcal{X}_2$, and $\mathcal{X}_3$ on a line in direction $\mathcal{S}_r(\mathbf{N})$.

4.2 Rank analysis of concentration data matrices

The rank and the dimension of the fundamental subspaces (column and row spaces, null space) of the concentration and RV-concentration data matrices are essential for the comprehension of the remainder of this work. In addition to the analysis of the rank and the dimension of the fundamental subspaces, methods that modify the dimension of the null space will be proposed.

The rank of a matrix $\mathbf{Y}$ determines the dimension of the fundamental subspaces in which $\mathbf{Y}$ lives or, equivalently, the variability (Lay, 1994, see also Appendix B.2 for the definition and some properties of the rank). The fundamental subspaces include the column and row spaces, and the null space (see Appendix B.1 for the definition of the fundamental subspaces of a matrix, and Appendices B.2.1, and B.3 for the dimension and the numerical construction of the null space, respectively).

Subsection 4.2.1 investigates the rank of various data matrices. Subsection 4.2.2 comments on the dimension of the corresponding null spaces (or, briefly, nullity) and on methods that change the nullity.

4.2.1 Rank of selected matrices

The ranks of $\mathbf{C}$, $\mathbf{X}_e$, $\mathbf{N}_e$, $\mathbf{D}$, $\mathbf{C}_m$, $\mathbf{D}_m$, and of the column-mean centered $\mathbf{C}$ denoted by $\bar{\mathbf{C}}_c$ (see Appendix F.1.2) are investigated next.

Let $\mathbf{Y}$ denote any of the four matrices: $\mathbf{C}$, $\mathbf{D}$, $\mathbf{C}_m$, $\mathbf{D}_m$. Then, the rank of $\mathbf{Y}$ determines

the dimension of the hyperplane in which the row vectors (e.g., concentrations, RV--concentrations) of $\mathbf{Y}$ evolve.

Rank of C

Theorem 4.3

Let Assumptions A1–7 in Appendix D be verified. Then, the rank of $\mathbf{C}$ *is bounded by:*

$$R + 1 \leq \operatorname{rank}(\mathbf{C}) \leq \min(R + p + 1, S). \tag{4.13}$$

(See Appendix H.3.2 for proof)

The lower bound of the rank of $\mathbf{C}$ is greater than R since each of the R independent reactions is assumed to be active on a time interval of the K observations (Assumption A7) and, according to Proposition 2.7d, the initial concentrations $\mathbf{c}_0$ do not lie in the row space of $\mathbf{N}$ and, thus, cause a rank increase of 1. The upper bound is determined by the smallest outer dimension of $\mathbf{C}$ (S, since $K > S$ by Assumption A6) and the inner dimension of $\mathbf{X}_e$ and $\mathbf{N}_e$ $(R + p + 1)$.

Conditions to guarantee $\operatorname{rank}(\mathbf{C}) = \min(R + p + 1, S)$ are given next as a corollary to Property B.10.

Corollary 4.4 *Let Assumptions A1–7 in Appendix D be verified. Then, the following properties hold:*

(a) *If* $\operatorname{rank}(\mathbf{X}_e) = (R + p + 1)$, *then* $\operatorname{rank}(\mathbf{C}) = \operatorname{rank}(\mathbf{N}_e)$.

(b) *If* $\operatorname{rank}(\mathbf{N}_e) = (R + p + 1)$, *then* $\operatorname{rank}(\mathbf{C}) = \operatorname{rank}(\mathbf{X}_e)$.

(c) *If* $p < S - R - 1$, *then* $\operatorname{rank}(\mathbf{C}) < S$.

If the rank of either $\mathbf{X}_e$ or $\mathbf{N}_e$ is full, then according to Corollary 4.4a–b, the rank of $\mathbf{C}$ is determined by the rank of the (possibly rank-deficient) matrix $\mathbf{N}_e$ or $\mathbf{X}_e$, respectively.

For the case $\operatorname{rank}(\mathbf{C}) = (R + p + 1)$, the variability or, equivalently, the rank of the concentration matrix $\mathbf{C}$ is determined by the dimensions of the following three parts: the reaction-variant part (dimension R), the reaction-invariant and flow-variant part (dimension p), and the reaction and flow-invariant part (dimension 1). Note that $\operatorname{rank}(\mathbf{N}_e) = S$ implies $p \geq S - R - 1$. However, if $p < S - R - 1$, then $\mathbf{C}$ is rank deficient (see Corollary 4.4c).

Rank of $\mathbf{X}_e$

The rank of $\mathbf{X}$ is full due to the definition of the independent reactions (see Definition 2.8). Furthermore, since $\mathbf{1}_K \notin \mathcal{S}_c(\mathbf{X})$, $\operatorname{rank}([\mathbf{X}, \mathbf{1}_K]) = (R + 1)$. Thus, the rank of $\mathbf{X}_e$ is always full for batch runs. The transformation proposed in Subsection 2.4.2 guarantees full rank of the *modified* inlet flowrates and, thus, $\mathbf{Z}$ is also of full rank. Therefore, the only situation of rank-deficient $\mathbf{X}_e$ is when some columns of $\mathbf{Z}$ lie in the column space of $[\mathbf{X}, \mathbf{1}_K]$. This situation can occur, for example, when the concentrations of some of the species are kept constant by feedback control with $\mathbf{q}_{in}$ being the manipulated variables (see Example 2.2d).

Experimental conditions are formulated next that guarantees full rank of $\mathbf{X}_e$.

Proposition 4.5 *Let Assumptions A1–7 in Appendix D be verified. If the p generalized inlets are of the type*

(a) additional batch runs,
(b) impulse additions at different time instants, or
(c) inlet flowrates with abrupt changes at different time instants,

then $\operatorname{rank}(\mathbf{X}_e) = R + p + 1$. *(See Appendix H.3.2 for proof)*

Proposition 4.5 provides three conditions to achieve full rank of $\mathbf{X}_e$. These conditions can easily be guaranteed by appropriate experimental planning. Changing each inlet flowrate at different time instants guarantees linear independence of the flowrates. Furthermore, the abrupt changes guarantee independence of the columns of $\mathbf{Z}$ and $\mathbf{X}$, since the reaction variants cannot exhibit a discontinuity. Note that for additional runs with inlet streams, the corresponding inlet flowrates must appropriately be chosen as indicated in Proposition 4.5.

Example 2.2d (Feedback control of $\mathcal{X}_1$; cont'd from page 18)
Consider a constant-density semibatch esterification reaction with $\mathcal{X}_1$ being fed at concentration $c_{in,1} > 0$ ($R = 1$, $p = 1$). The initial concentrations are $\mathbf{c}_0 = [c_{1,0},\, c_{2,0},\, 0,\, 0,\, c_{5,0}]^{\mathrm{T}}$ with $c_{1,0},\, c_{2,0},\, c_{5,0} > 0$. The concentration of $\mathcal{X}_1$ is kept constant by feedback control at a given setpoint $c_{1,s} = c_{1,0}$ using q_{in} as the manipulated variable. From (2.38), it follows that

$$0 \stackrel{!}{=} \dot{c}_{1,s} = -r + (c_{in,1} - c_{1,s})\, q_{in}/V,$$

and a feedback law for q_{in} can be constructed:

$$q_{in} = \alpha\, V\, r \tag{4.14}$$

with $\alpha = 1/(c_{in,1} - c_{1,s})$. Thus, $\mathcal{X}_1$ must be fed at a flowrate proportional to the reaction rate. Since there is no outlet, $\lambda(t) = 1$. From (4.14) and the definition of x and z, it follows that $z(k) = \alpha\, x(k)$ for all k, or, $\mathbf{z} = \alpha\, \mathbf{x}$. Since $\operatorname{rank}(\mathbf{N}_e) = R + p + 1 = 3$, by invoking Corollary 4.4b, $\operatorname{rank}(\mathbf{C}) = \operatorname{rank}(\mathbf{X}_e)$. Since the first element of $\mathbf{x}$ is 0, it follows that $\mathbf{1}_K \notin \mathcal{S}_c(\mathbf{x})$ and, thus, $\operatorname{rank}([\mathbf{x},\, \mathbf{1}_K]) = R + 1 = 2$. Since $\mathbf{X}_e = \mathbf{V}^{-1}[\mathbf{x},\, \mathbf{z},\, V_0\mathbf{1}_K] = \mathbf{V}^{-1}[\mathbf{x},\, \alpha\,\mathbf{x},\, V_0\mathbf{1}_K]$, it is concluded that $\operatorname{rank}(\mathbf{C}) = \operatorname{rank}(\mathbf{X}_e) = 2 < (R + p + 1) = 3$.

Rank of $\mathbf{N}_e$

The rank of $\mathbf{N}$ is full due to the definition of the independent reactions (see Definition 2.8). Furthermore, since according to Proposition 2.7d $\mathbf{c}_0 \notin \mathcal{S}_r(\mathbf{N})$, $\operatorname{rank}\left(\begin{bmatrix}\mathbf{N}\\ \mathbf{c}_0^{\mathrm{T}}\end{bmatrix}\right) = (R+1)$. Thus, the rank of $\mathbf{N}_e$ is always full for batch runs. An example of rank-deficient $\mathbf{N}_e$ is when, for reaction systems with $p < (S - R - 1)$ inlet streams, one or several rows of $\mathbf{C}_{in}$ lie in the row space of $\begin{bmatrix}\mathbf{N}\\ \mathbf{c}_0^{\mathrm{T}}\end{bmatrix}$.

To guarantee full rank of $\mathbf{N}_e$, it is useful to have $\operatorname{rank}(\mathbf{C}_x) = (p + 1)$. This is a practical assumption, since $\mathbf{C}_{in}$ and $\mathbf{c}_0$ are usually known. For computing the rank of $\mathbf{N}_e$ of reaction systems with inlet streams, $\mathbf{N}$ must be known. However, this is often *not* the case in practice. Thus, for reaction systems with inlet streams, the knowledge of $\mathbf{N}$ is often the bottleneck in the computation of the rank of $\mathbf{N}_e$ (or $\mathbf{C}$).

Rank of $\bar{\mathbf{C}}_c$

In practical applications, the column-mean-centered or time-differentiated $\mathbf{C}$ is often used. As can be seen from Property B.7, column-mean centering or differentiation with respect to time cause a rank drop by one if and only if its column space contains the $\mathbf{1}_K$-vector.

Rank of D

Proposition 4.6 *Let Assumptions A1–7 in Appendix D be verified. Then,*

$$\operatorname{rank}(\mathbf{D}) = R. \tag{4.15}$$

(See Appendix H.3.2 for proof)

It can be seen from (4.15) that the rank of $\mathbf{D}$ is determined by the number of independent reactions R, since the effects of the reaction invariants have been removed. Thus, in contrast to $\mathbf{C}$, the rank of $\mathbf{D}$ can be determined exactly. Furthermore, $\mathbf{D}$ is always rank deficient, since $R < S$ (see Proposition 2.7a).

Rank of $\mathbf{C}_m$ and $\mathbf{D}_m$

When the concentrations of only $S_m < S$ species are measured, then the rank of $\mathbf{D}_m$ may be less than the *total* number of independent reactions R. The rank of $\mathbf{D}_m$ determines the number of *observed* independent reactions, R_o.

When $R_o < R$, then $(R - R_o)$ rows of $\mathbf{N}_m$ are linear combinations (denoted by the $(R - R_o) \times R_o$ matrix $\mathbf{\Gamma}_r$) of the remaining R_m ones (denoted by the $R_o \times S_m$ matrix $\mathbf{N}_o$):

$$\mathbf{N}_m = \left[\begin{array}{c} \mathbf{N}_o \\ \hline \mathbf{\Gamma}_r \mathbf{N}_o \end{array}\right], \tag{4.16}$$

or

$$\mathbf{N}_m = \mathbf{N}\,\mathcal{P}_m = \mathcal{P}_N\,\mathbf{N}_o, \qquad \mathcal{P}_m = \begin{bmatrix} \mathbf{I}_{S_m} \\ \mathbf{0}_{S-S_m \times S_m} \end{bmatrix}, \quad \mathcal{P}_N \equiv \begin{bmatrix} \mathbf{I}_{R_o} \\ \mathbf{\Gamma}_r \end{bmatrix}, \tag{4.17}$$

where $\mathcal{P}_m$ and $\mathcal{P}_N$ are matrices of dimension $S \times S_m$ and $R \times R_o$, respectively. Using $\mathcal{P}_N$, the $K \times R_o$ matrix of *observed* reaction variants $\mathbf{X}_o$ is defined as:

$$\mathbf{X}_o \equiv \mathbf{X}\,\mathcal{P}_N = \left[\begin{array}{c|c} \mathbf{X}_1 & \mathbf{X}_2 \end{array}\right] \begin{bmatrix} \mathbf{I}_{R_o} \\ \mathbf{\Gamma}_r \end{bmatrix} = \mathbf{X}_1 + \mathbf{\Gamma}_r \mathbf{X}_2, \tag{4.18}$$

where $\mathbf{X}_1$ and $\mathbf{X}_2$ are submatrices of $\mathbf{X}$ of dimension $K \times R_o$ and $K \times (R - R_o)$, respectively. Thus, $\mathbf{D}_m$ (see 4.12) can be rewritten as:

$$\mathbf{D}_m = \mathbf{X}\,\mathbf{N}_m = \mathbf{X}\,\mathcal{P}_N\,\mathbf{N}_o = \mathbf{X}_o\,\mathbf{N}_o. \tag{4.19}$$

Equivalently, $\mathbf{C}_m$ becomes:

$$\mathbf{C}_m = \mathbf{X}_e\,\mathbf{N}_{e,m} = \mathbf{X}_{e,o}\,\mathbf{N}_{e,o}, \qquad \mathbf{X}_{e,o} \equiv \mathcal{H}^{-1}\begin{bmatrix} \mathbf{X}_o & \mathbf{Z}_x \end{bmatrix}, \quad \mathbf{N}_{e,o} \equiv \begin{bmatrix} \mathbf{N}_o \\ \mathbf{C}_{x,m} \end{bmatrix}, \tag{4.20}$$

where $\mathbf{X}_{e,o}$ and $\mathbf{N}_{e,o}$ are matrices of dimension $K \times (R_o + p + 1)$ and $(R_o + p + 1) \times S_m$, respectively. Thus, the rank of $\mathbf{C}_m$ may also be less than that of $\mathbf{C}$.

Proposition 4.7 *Let Assumptions A1–7 in Appendix D be verified. Let the number of* observed *independent reactions be defined by* $R_o \equiv \operatorname{rank}(\mathbf{D}_m)$. *Then,*

$$R_o \leq \operatorname{rank}(\mathbf{C}_m) \leq \min(R_o + p + 1, S_m). \tag{4.21}$$

(See Appendix H.3.2 for proof)

In contrast to $\mathbf{C}$ ($S_m = S$), $\mathbf{c}_{0,m}$ for $S_m < S$ can lie in $\mathbf{N}_m$ and, thus, the rank of $\mathbf{C}_m$ can equal that of $\mathbf{D}_m$ (lower bound on the rank of $\mathbf{C}_m$). The upper bound is determined by the smallest outer dimension of $\mathbf{C}_m$ (S_m) and the inner dimension of $\mathbf{X}_{e,o}$ and $\mathbf{N}_{e,o}$ ($R_o + p + 1$).

Example 4.8 ($R_o < R$)
Consider a batch reaction system with $S = 4$ reacting species and $R = 2$ independent reactions (Amrhein *et al.*, 1996):

$$2\,\mathcal{X}_1 \xrightarrow{\kappa_1} \mathcal{X}_2 \xrightarrow{\kappa_2} \mathcal{X}_3 + 2\,\mathcal{X}_4.$$

Let $\mathbf{c}_0 = [c_{1,0},\, 0,\, c_{3,0},\, 0]^{\mathrm{T}}$ with $c_{1,0} \neq 0$, $c_{3,0} \neq 0$.
Two cases are considered: (C1) all species are measured ($\mathbf{c}_{m,C1} = \mathbf{c}$), (C2) Species $\mathcal{X}_3$ and $\mathcal{X}_4$ are measured ($\mathbf{c}_{m,C2} = [c_3,\, c_4]^{\mathrm{T}}$).
In Case C1, according to Corollary 4.4b, $\operatorname{rank}(\mathbf{C}_{m,C1}) = R + 1 = 3$ (see also Subsection 4.5.2 below). In Case C2, $\mathbf{N}_{e,m,C2}$ becomes

$$\mathbf{N}_{e,m,C2} = \begin{bmatrix} \mathbf{N}_{m,C2} \\ \hline \mathbf{c}_{0,m,C2}^{\mathrm{T}} \end{bmatrix} = \left[\begin{array}{cc} 0 & 0 \\ 1 & 2 \\ \hline c_{3,0} & 0 \end{array}\right]. \tag{4.22}$$

From (4.22), it can be seen that the first reaction cannot be observed, since none of the two species $\mathcal{X}_1$ and $\mathcal{X}_2$ involved in the first reactions are measured. Thus, $R_o = 1 < R$ and $\operatorname{rank}(\mathbf{N}_{e,m,C2}) = 2$.

4.2.2 Nullity of selected matrices

The nullities (i.e., the dimension of the null space) of $\mathbf{C}$, $\mathbf{D}$, $\mathbf{C}_m$ and $\mathbf{D}_m$ are studied next. Furthermore, nullity-reducing and nullity-increasing operations are proposed.

4.2.2.1 Formulation

To set the stage, certain useful mathematical results are listed next, which are then discussed.

Proposition 4.9 *Let Assumptions A1–7 in Appendix D be verified. Then, for* $S_m \leq S$ *measured species, the nullities of* $\mathbf{C}_m$ *and* $\mathbf{D}_m$ *are given by:*

$$\begin{aligned} \max(S_m - R_o - p - 1, 0) \leq \dim(\mathcal{N}(\mathbf{C}_m)) = S_m - \operatorname{rank}(\mathbf{C}_m) \leq S_m - R_o, \\ \dim(\mathcal{N}(\mathbf{D}_m)) = S_m - \operatorname{rank}(\mathbf{D}_m) = S_m - R_o. \end{aligned} \tag{4.23}$$

If $S_m = S$, then

$$\begin{aligned} \max(S - R - p - 1, 0) \leq \dim(\mathcal{N}(\mathbf{C})) \leq S - R - 1, \\ \dim(\mathcal{N}(\mathbf{D})) = S - R. \end{aligned} \tag{4.24}$$

(See Appendix H.3.2 for proof)

The nullity of a matrix is determined by its column dimension and rank. When the concentrations of only $S_m < S$ species are measured and the rank of $\mathbf{D}_m$ is $R_o = R$, then the null space of $\mathbf{D}_m$ is smaller than that of $\mathbf{D}$. For $R_o < R$ (see Subsection 4.2.1), however, the nullities of $\mathbf{D}$ and $\mathbf{D}_m$ can also be equal. Similar statements also hold for $\mathbf{C}_m$ and $\mathbf{C}$.

Let $\mathbf{Y}$ represent any of the four data matrices considered: $\mathbf{C}$, $\mathbf{D}$, $\mathbf{C}_m$, or $\mathbf{D}_m$. *Nullity-changing operations* are defined next.

Definition 4.10 (Nullity-changing operations) *Let d_- and d_+ denote the nullities of $\mathbf{Y}_-$ ($K \times S_-$) before and $\mathbf{Y}_+$ ($K \times S_+$) after the nullity-changing operation, i.e.,*

$$d_- \equiv \dim(\mathcal{N}(\mathbf{Y}_-)) = S_- - \operatorname{rank}(\mathbf{Y}_-), \quad d_+ \equiv \dim(\mathcal{N}(\mathbf{Y}_+)) = S_+ - \operatorname{rank}(\mathbf{Y}_+). \tag{4.25}$$

The change in nullity is defined as:

$$\Delta d(\mathbf{Y}_-, \mathbf{Y}_+) \equiv d_+ - d_-. \tag{4.26}$$

If $\Delta d < 0$ ($\Delta d > 0$), then the operation is said to be nullity reducing *(*increasing*).*

4.2.2.2 Nullity-reducing operations

Experimental rank-increasing methods

Proposition 4.11 *Let Assumptions A1–7 in Appendix D be verified. Let $S_- = S_+ = S$ and $R_- = R_+ = R$, where the subscripts $-$ and $+$ denote a quantity before and after variation using any of Variation modes V1–V3 on page 28, respectively. Then, the following properties hold for the nullity change Δd:*

(a) *For $(p_+ - p_-)$ additional generalized inlets, $\Delta d \leq 0$.*

(b) *If, in addition to (a), $\operatorname{rank}(\mathbf{C}_-) = (R + p_- + 1)$ and $\operatorname{rank}(\mathbf{C}_+) = (R + p_+ + 1)$, then $\Delta d = p_- - p_+ < 0$.*

(c) *If $\operatorname{rank}(\mathbf{C}_+) = S$, then $d_+ = 0$ and $\Delta d = -d_-$.*

(See Appendix H.3.2 for proof)

Proposition 4.11 says that, since Variation modes V1–V3 on page 28 are cast in the generalized inlets, they are equivalent with respect to a rank increase of $\mathbf{C}$. Proposition 4.11a shows that additional generalized inlets can increase but never decrease the rank. If $\operatorname{rank}(\mathbf{C}_-) = (R + p_- + 1)$ and $\operatorname{rank}(\mathbf{C}_+) = (R + p_+ + 1)$, then each additional

generalized inlet increases the rank by one (see Proposition 4.11b). In some cases, the additional generalized inlets increase the rank of $\mathbf{C}_+$ to S, in which case, the nullity of $\mathbf{C}_+$ is zero (see also Subsection 4.2.2).

Note that a rank increase is possible only up to the point where $\operatorname{rank}(\mathbf{C}_+) = S$. If additional variations are performed, no further rank increase can be observed ($\Delta d = 0$).

Maximal attainable rank

In practical applications, often only S_x species are available for independent variation of the initial and inlet concentrations. Thus, the rank of $\mathbf{C}$ can at most $(R + S_x)$ or, in other words, the minimal nullity of $\mathbf{C}$ is $S - R - S_x$. To maximally increase the rank of $\mathbf{C}$ with a minimal set of generalized inlets, the ranks of $\mathbf{X}_e$ and $\mathbf{N}_e$ must be full. Furthermore, to guarantee $\mathbf{C}$ to be of full rank S, at least $(S - R)$ species should be available for independent variation or, in other words, the number of species unavailable for variation should not exceed R.

Proposition 4.5 proposes methods to guarantee full rank of $\mathbf{X}_e$. To guarantee full rank of $\mathbf{N}_e$, $\mathbf{N}$ must usually be known for the specification $\mathbf{C}_x$. However, if all S species are available for independent variation ($S_x = S$), then $\operatorname{rank}(\mathbf{C}_x) = S$ and, thus, $\operatorname{rank}(\mathbf{N}_e) = S$. In such a case, R generalized inlets are superfluous, but with the advantage that the knowledge of $\mathbf{N}$ is not required.

Example 2.13b (Variation modes, linear dependence in $\mathbf{N}_e$, column-mean centering; cont'd from page 31**)**
The experimental rank-increasing operations are illustrated by means of the reaction system explained in Example 2.13 ($R = 2$, $S = 6$). It is assumed that $S_- = S_+ = S$. The first reaction is assumed to be of either type (i) power law with $r_1^{(1)}(c_1) = \kappa_1 c_1^2$, or (ii) inhibition with $r_1^{(2)}(c_1, c_5) = \frac{k_{11} c_1^2}{1 + k_{12} c_5}$.
Six runs are conducted: the first four runs are the same as in Example 2.13, and Runs R5 and R6 are semibatch runs with one inlet stream. Key difference between

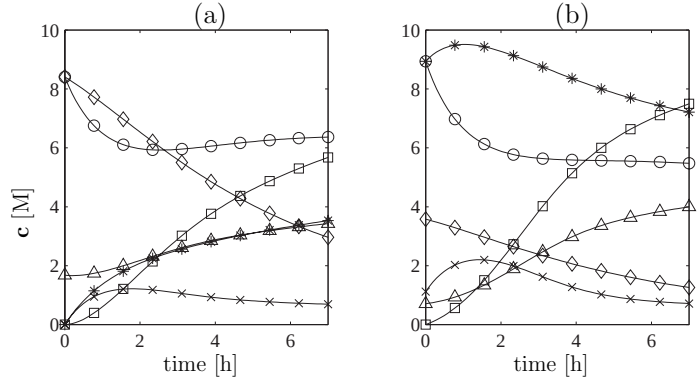

Figure 4.8. The concentrations $\mathbf{c}$ [M] of: (a) Run R5 and (b) Run R6 of Example 2.13b. For legend of $\mathbf{c}$, see Figure 2.5 on page 30.

R5 and the other runs is the kinetics of the first reaction. The initial concentrations of R6 involve Species $\mathcal{X}_3$ that is not initially present or added in the other runs (see Appendix A.2 for the parameters). Figure 4.8 shows the concentration profiles of Runs R5 and R6. The feed and volume profiles of Runs R5 and R6 are the same as for Run R4 (see Figure 4.2d and e).

Five cases are considered. Case C1: Run R2; Case C2: (a) Run R1, (b) Runs R2 & R3, (c) Run R4; Case C3: (a) Runs R1 & R4, (b) Runs R1 & R3 & R4; Case C4: Runs R1 & R4 & R5; Case C5: Runs R1 & R3 & R4 & R6.

Table 4.2 summarizes the ranks of $\mathbf{C}$, $\mathbf{X}_e$, and $\mathbf{N}_e$ for Cases C1–C5. For all the cases, $\text{rank}\,(\mathbf{X}_e) = (R + p + 1)$, since the conditions specified in Proposition 4.5 are verified. Thus, by invoking Corollary 4.4a, the rank of $\mathbf{C}$ is determined by the rank of $\mathbf{N}_e$.

$\mathbf{C}$ of a single batch run R2 (Case C1) is rank deficient with $\text{rank}\,(\mathbf{C}) = \text{rank}\,(\mathbf{N}_e) = R + 1 = 3$. In Case 2, variations such as a single impulse addition, appending an additional batch run, or a single continuous feed increase the rank by one with respect to batch run R2 ($\Delta d = -1$), since the corresponding concentrations are linearly independent of the initial concentrations of the (first) batch run R2. Case C3 illustrates stepwise rank increase from 4 (Case 2) to 5 (Case 3a; $\Delta d = -1$) and from 5 to full rank $S = 6$ (Case 3b; $\Delta d = -1$).

Rank deficiency in $\mathbf{N}_e$ occurs for the following cases: C2b, C3a–b, C4, C5. The reasons are various. Cases C2b, C3a–C3b are rank deficient, since the initial concentrations are the same for Runs R1–4.

For Case C4, although the initial and inlet concentrations, and the kinetics of Run R5 are different from Runs R1 and R4 (Case C3a), no rank increase can be observed when appending Run R5 to Runs R1 & R4 ($\Delta d = 0$). This is because (i) the initial concentrations, the impulse addition concentrations, and the inlet concentrations of Runs R1 and R4 already span the entire subspace for Species $\mathcal{X}_1$, $\mathcal{X}_4$, and $\mathcal{X}_6$, i.e., $\text{rank}\,(\mathbf{C}_{x,C3a}) = 3 = S_x$, and (ii) the initial and inlet concentrations of Run R5 only involve these three species and, therefore, $\mathbf{c}_{0,R5},\ \mathbf{c}_{in,R5} \in \mathcal{S}_c\,(\mathbf{C}_{x,C3a})$.

For Case C5, appending Run R6 to Runs R1 & R4 & R5 (Case C3b) does not increase the rank ($\Delta d = 0$), since the rank of $\mathbf{C}_{C3b}$ is already full. Thus, although the initial concentrations of R6 involve Species $\mathcal{X}_3$, which is not present in $\mathbf{C}_{x,C3b}$, $\mathbf{c}_{0,R6} \in \mathbf{N}_{e,C3b}$, i.e., the subspace for $\mathcal{X}_3$ is spanned by the rows of $\mathbf{C}_{x,C3b}$ *and* $\mathbf{N}$.

Table 4.2. Computed ranks of $\mathbf{C}$, $\mathbf{X}_e$, and $\mathbf{N}_e$ for Cases C1–C5 ($R = 2$, $S = 6$). The number of (generalized) inlets p is also listed. R1: batch + one impulse addition; R2: batch; R3: batch; R4: semibatch with one inlet stream; R5: semibatch with one inlet stream and different kinetics; R6: semibatch with one inlet stream and different kinetics. See Table A.4 for initial and (generalized) inlet concentrations.

Case	p	Rank of			Case	p	Rank of		
		$\mathbf{C}$	$\mathbf{X}_e$	$\mathbf{N}_e$			$\mathbf{C}$	$\mathbf{X}_e$	$\mathbf{N}_e$
C1 (R2)	0	3	3	3	C3a (R1,R4)	3	5	6	5
C2a (R1)	1	4	4	4	C3b (R1,R3,R4)	4	6	7	6
C2b (R2,R3)	1	4	4	4	C4 (R1,R4,R5)	5	5	8	5
C2c (R4)	1	4	4	4	C5 (R1,R3,R4,R6)	6	6	9	6

Example 2.2e (Linear dependence in $\mathbf{X}_e$; cont'd from page 65)
Two runs are conducted for the constant-density esterification reaction. Run R1 is a batch reaction with $\mathbf{c}_0 = [c_{1,0},\, c_{2,0},\, 0,\, 0,\, c_{5,0}]^{\mathrm{T}}$ with $c_{1,0},\, c_{2,0},\, c_{5,0} > 0$ and Run R2 the semibatch reaction under feedback control as described in Example 2.2d.
For both runs, the rank of $\mathbf{N}_e$ is full, i.e., $\operatorname{rank}(\mathbf{N}_{e,R1}) = R+1 = 2$ and $\operatorname{rank}(\mathbf{N}_{e,R2}) = R+p+1 = 3$. Thus, by invoking Corollary 4.4b it follows that $\operatorname{rank}(\mathbf{C}) = \operatorname{rank}(\mathbf{X}_e)$, i.e., the rank of $\mathbf{C}$ is determined by the rank of $\mathbf{X}_e$.
In Example 2.2d it was already shown that $\operatorname{rank}(\mathbf{X}_{e,R1}) = \operatorname{rank}([\mathbf{x},\, \mathbf{1}_K]) = R+1 = 2$. Since $\operatorname{rank}(\mathbf{X}_{e,R2}) = 2$ (see Example 2.2d), it follows that $\operatorname{rank}(\mathbf{C}_{R1}) = \operatorname{rank}(\mathbf{C}_{R2})$. Thus, if the additional feed is constraint by the feedback law (4.14), then the rank does not increase with respect to the batch Run R1 ($\Delta d = 0$).

Measuring fewer species

The nullity of the concentration matrix can be affected by measuring fewer species.

Proposition 4.12 *Let Assumptions A1–7 in Appendix D be verified, and the subscripts − and + denote a quantity related to more and fewer measurements, respectively. If* $S_- = S$ *and* $\operatorname{rank}(\mathbf{C}_-) = R+p+1$*, then measuring less species* ($S_+ = S_m < S$) *leads to the nullity reduction* Δd *of the corresponding concentration matrices* $\mathbf{C}_-$ *and* $\mathbf{C}_+$*:*

$$\max(S_m - S + R - R_o, R + p + 1 - S) \leq \Delta d \leq S_m - S + R - R_o + p + 1. \tag{4.27}$$

(See Appendix H.3.2 for proof)

Measuring fewer species can reduce but never increase the nullity. The same results also hold for $\mathbf{D}_-$ and $\mathbf{D}_+$ with the redefinition $(p+1) := 0$.

4.2.2.3 Nullity-increasing operations

The pre-treatment of $\mathbf{C}$ and $\mathbf{C}_m$ to reaction-variant form $\mathbf{D}$ (see 4.11) and $\mathbf{D}_m$ (see 4.12), respectively, is a nullity-increasing operation, as shown next.

Proposition 4.13 *Let Assumptions A1–7 in Appendix D be verified. Then,*

$$\begin{aligned} 1 &\leq \Delta d(\mathbf{C}, \mathbf{D}) \leq \min(p+1, S-R), \\ 0 &\leq \Delta d(\mathbf{C}_m, \mathbf{D}_m) \leq \min(p+1, S_m - R_o). \end{aligned} \tag{4.28}$$

(See Appendix H.3.2 for proof)

When subtracting the reaction-invariant part from $\mathbf{C}$, Proposition 4.13 shows that the nullity increases at least by 1 but maximally by $\min(p+1, S-R)$. Alternatively, when subtracting the reaction-invariant part from $\mathbf{C}_m$, the nullity stays can increase by as much as $\min(p+1, S_m - R_o)$.

4.3 Some implications

4.3.1 On-line state reconstruction

In many practical situations, the concentrations of some species of interest are difficult to measure on-line (e.g. biomass or glucose in biotechnological reactions). Consequently, it is desirable to reconstruct them, without knowledge of kinetics, from the concentration of the species that are routinely measured on-line (e.g. CO_2 or O_2 in biotechnological reactions). In this subsection, the well-known asymptotic observer of Bastin and Dochain (1990) is analyzed in the light of reaction and flow variants/invariants, and certain useful results regarding the observation error are provided.

It is shown below that if the concentrations, $\mathbf{c}_m$, of $S_m \geq R$ species are measured, the reaction variants $\mathbf{x}$ can be estimated, without knowledge of kinetics, under certain minor assumptions. Upon estimating $\mathbf{x}$, the concentration vector $\mathbf{c}_u$ of the remaining $S_u = (S - S_m)$ species can be reconstructed from (4.3).

Proposition 4.14 *Let Assumptions A1–7 in Appendix D be verified. Let the concentrations $\mathbf{c}_m$ of $S_m \geq R$ species be measured, and $\mathbf{N}_m$ be the $R \times S_m$ submatrix of $\mathbf{N}$ corresponding to these S_m species. Given $\mathbf{N}$, $\mathbf{C}_{in}$, $\mathbf{q}_{in}$, q_{out}, $\mathbf{c}_0$, and V_0, if $\mathbf{N}_m^{\mathrm{T}}$ has a unique left pseudo-inverse, the reaction variants $\mathbf{x}$ can be reconstructed without knowledge of reaction kinetics using*

$$\hat{\mathbf{x}} = \mathbf{N}_m^{+\mathrm{T}} \left(\mathrm{h}(\mathbf{z})\, \mathbf{c}_m - \mathbf{C}_{in,m}\mathbf{z} - V_0 \mathbf{c}_{m,0} \right), \tag{4.29}$$

where the subscript m denotes a quantity corresponding to the S_m measured species. From (4.29), the concentrations of the remaining $(S - S_m)$ species, $\mathbf{c}_u$, can be reconstructed using

$$\hat{\mathbf{c}}_u = \left(\mathbf{N}_u^{\mathrm{T}} \hat{\mathbf{x}} + \mathbf{C}_{in,u}\mathbf{z} + V_0 \mathbf{c}_{u,0} \right) / \mathrm{h}(\mathbf{z}), \tag{4.30}$$

where the subscript u denotes a quantity corresponding to the S_u unmeasured species.

Let only an estimate of the initial concentrations of the S_u species be available, $\hat{\mathbf{c}}_{u,0}$. Then the estimation error $\boldsymbol{\varepsilon}_u \equiv \hat{\mathbf{c}}_u - \mathbf{c}_u$, with $\hat{\mathbf{c}}_u$ being the estimated concentrations, is given by

$$\varepsilon_u^i(k) = V_0 \varepsilon_{u,0}^i / \mathrm{h}(\mathbf{z}(k)), \qquad i = 1, \ldots, S_u, \tag{4.31}$$

where ε_u^i is the ith element of $\boldsymbol{\varepsilon}_u$. If the inlets are present at least intermittently, then the estimation error asymptotically converges to zero. (See Appendix H.3.3 for proof)

Since the reconstructed reaction variants $\hat{\mathbf{x}}$ in (4.29) are independent of the (unmeasured) initial concentrations $\mathbf{c}_{u,0}$, the RV-concentration vector $\mathbf{d}_u$ of the S_u unmeasured species are, in principle, correctly reconstructed:

$$\mathbf{d}_u = \mathbf{N}_u^{\mathrm{T}} \mathbf{x}. \tag{4.32}$$

However, the concentrations $\mathbf{c}_u$ can exhibit an estimation error due to inaccurate estimate of the initial concentrations $\mathbf{c}_{u,0}$. Fortunately, asymptotic convergence to the

true concentrations can be guaranteed. The asymptotic convergence can intuitively be explained from (4.31). The numerator term is constant while the denominator monotonically increases if the inlets are present at least intermittently. $\mathrm{h}(\mathbf{z})$ tends to infinity with time, thereby pushing the error to zero. Since $\hat{\mathbf{c}}_u^i$ does not depend on $\hat{\mathbf{c}}_{u,0}^j$ $(i \neq j)$, an initial concentration error of any species does not propagate on the concentration estimates of the remaining species.

It can be seen from (4.31) that the error $\boldsymbol{\varepsilon}_u$ tends to zero only due $\mathrm{h}(\mathbf{z})$ going to infinity. Thus, since $\mathrm{h} = V/\lambda$, either the volume goes to infinity or λ goes to zero. Both imply that infinite dilution is necessary to make the estimation independent of the initial concentration estimates.

Note that, under the assumption that the concentrations of more R species are measured, the procedure presented here only provides an estimate at the kth time instant from a measurement at the kth time instant. If it is desirable to estimate concentrations for all time instants (past, present, future) from past and present measurements, then observer techniques relying on the knowledge of reaction kinetics should be used (Soroush, 1997). In some cases, these techniques work also when the concentrations of less R species are measured.

Example 2.5b (Ethanolysis reaction system; cont'd from page 12**)**
It is assumed that the concentrations of Species $\mathcal{X}_1$, $\mathcal{X}_4$, and $\mathcal{X}_7$ are measured with additive zero-mean Gaussian noise with standard deviations 2.8%, 3.3%, and 1.5%, respectively, and the volume is reconstructed perfectly from the continuity equation. Furthermore, it is assumed that the initial concentrations of all the species are perfectly known except that of the (unmeasured) Species $\mathcal{X}_2$: $\hat{c}_{2,0} = 3.9$ M (true $c_{2,0} = 4.2$). Figure 4.9 shows the reconstructed concentrations of the unmeasured species, $\hat{\mathbf{c}}_u = [\hat{c}_2, \hat{c}_3, \hat{c}_5, \hat{c}_6]^{\mathrm{T}}$. Since the pseudo-inverse of $\mathbf{N}_m^{\mathrm{T}} = \begin{bmatrix} -1 & 0 & 0 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$ is unique and the initial concentrations are correct, the concentrations of Species $\mathcal{X}_3$, $\mathcal{X}_5$, and $\mathcal{X}_6$ are recon-

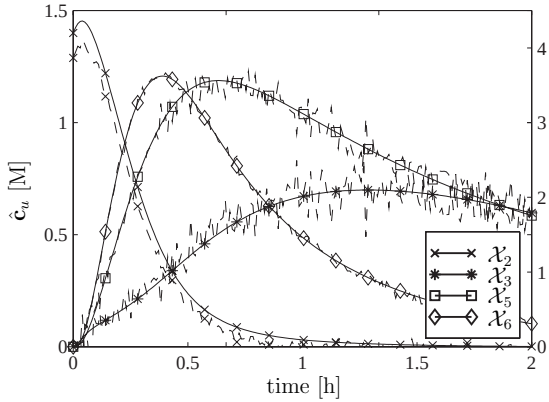

Figure 4.9. Example 2.5b: concentrations of the S_u unmeasured species: (–) true, (– –) reconstructed. Right axis for $\mathcal{X}_2$ and left axis for the remaining species.

structed without offset.
Since there is no outlet, $\mathrm{h}(\mathbf{z}) = V$. Thus, the asymptotic convergence to the true concentrations observed for Species $\mathcal{X}_2$ is only due to the increase in volume (see Figure 3.1b).

4.3.2 Modeling of the stoichiometries using target factor analysis (TFA)

The kinetic description still represents the main difficulty in modeling chemical reaction systems. For most industrially-relevant reactions, the kinetic parameters cannot be estimated reliably from theory and, thus, must be determined experimentally. This estimation requires the system stoichiometry *and* candidate kinetic structures to be available. In practice, however, it would be preferable to identify the reaction stoichiometries from the available data independently of reaction kinetics. In other words, instead of fitting a global model (stoichiometric and kinetic) to measured data, it would be better to proceed in two steps: (i) determine the reaction stoichiometries and reaction variants from measured data without knowledge of kinetics, and (ii) determine the kinetic structure for each reaction individually from the corresponding reaction variant computed in Step i. For Step ii, the interested reader is referred to the literature on kinetic modeling (Himmelblau *et al.*, 1967; Hill, 1977; Perry and Green, 1984).

Target factor analysis (TFA) is a useful multivariate analysis tool to determine, without knowledge of reaction kinetics, the number of independent reactions and the corresponding stoichiometries. It was initially used by Bonvin and Rippin (1990) to determine the number of independent reactions and the corresponding stoichiometries from composition measurements. Composition data can be collected from various experimental setups (isothermal or nonisothermal, isobaric or non-isobaric, batch or semibatch reactors, reactors with or without an outlet stream, etc.) so long as the reaction-variant part in the measured data can be isolated from the reaction-invariant part (material exchange terms and initial conditions).

The identification of stoichiometric models using factor-analytical techniques requires the composition of all the species to be measured and expressed in mass units (Hamer, 1989) or number of moles (Bonvin and Rippin, 1990). In many cases, concentrations of only a limited number of species are available. Also, concentration and weight fraction measurements are those which are found commonly in practice. To convert concentration (weight fraction) measurements into number of moles (masses), knowledge of reactor volume (mass) is required. In many practical situations, the volume is unknown, since it depends on the inlet and outlet streams and the density. This represents a bottleneck in applying existing TFA techniques to measured concentration data.

Herein, the applicability of TFA to concentration data from reaction systems for determining the number of independent reactions and the corresponding stoichiometries and reaction variants is analyzed. Data in reaction-variant form $\mathbf{D}$ will be used to enable direct application of existing TFA techniques. In addition, TFA is generalized to the situation where concentration measurements must be used directly, i.e., data pre-treatment to reaction-variant form is not possible.

Existing TFA techniques rely on the fact that the measured concentration vectors must lie in the row space of the stoichiometric matrix (the reaction-variant space) or, in other words, the null space of the measured data matrix should coincide with that of the stoichiometric matrix $\mathbf{N}$. The concentration matrix $\mathbf{C}$ has a row space, the space spanned by the reaction variants *and* invariants, larger than that of $\mathbf{N}$. Data pre-treatment to reaction-variant form eliminates the reaction-invariant part so that the rows of $\mathbf{D}$ and of $\mathbf{N}$ span the same space. Thus, the null space of $\mathbf{D}$ and $\mathbf{N}$ coincide, and unambiguous determination of the number of independent reactions and target testing are possible.

A TFA methodology is proposed to handle concentration measurements $\mathbf{C}$ directly for which the nullity is smaller than that of $\mathbf{N}$. In such a case, for unambiguous target testing, conditions related to the reaction invariance matrix $\mathbf{\Gamma}^{\mathrm{T}}$ (see Subsection 2.2.2) are proposed. Once a stoichiometric model has been determined, a major advantage of this methodology is that the volume can be reconstructed from concentration measurements without knowledge of density and reaction kinetics.

In many practical situations, only some of the reacting species can be measured (partial concentration measurements; $\mathbf{C}_m$ or $\mathbf{D}_m$). A nullity problem, similar to that faced while handling concentration measurements $\mathbf{C}$ directly, is encountered. Consequently, for unambiguous target testing, conditions related to the reaction-invariance matrix will also be proposed.

To stay in the spirit of this dissertation, all results are first presented for chemical reaction systems with constant density and temperature. It will be shown in Subsections 4.4.1 and 4.4.2 that, since the structures of the factorization of concentration data for reaction systems with varying density and temperature remain the same, the results of this subsection can readily be extended to the cases of varying density and temperature.

Subsection 4.3.2.1 briefly reviews some basics of the TFA procedure applied to reaction data. In Subsection 4.3.2.2 and 4.3.2.3, TFA results with pre-treated and raw measured concentration data are presented, respectively. Subsection 4.3.2.4 discusses the results obtained. In Subsections 4.4.1, 4.4.3, and 4.5.2, TFA is considered for reaction systems with varying density, weight fraction measurements, and batch reaction systems, respectively. Subsection 4.3.2.5 concludes this subsection by illustrating the TFA procedure using batch and semibatch reaction systems with varying volume and density.

4.3.2.1 Preliminaries

A brief review of TFA is provided next. For more details, see Appendix C.3. Let $\mathbf{Y}$ be an $K \times S$ noise-free matrix of rank A that is described by a linear factorization, i.e., it can be described by two physical matrices $\mathfrak{X}$ and $\mathfrak{N}$ of dimension $K \times A$ and $A \times S$, respectively:

$$\mathbf{Y} = \mathfrak{X}\,\mathfrak{N}. \tag{4.33}$$

TFA consists of two steps: (i) determination of the abstract matrices $\mathfrak{X}_a$ and $\mathfrak{N}_a$ that have the same column and row spaces as $\mathfrak{X}$ and $\mathfrak{N}$, respectively, and (ii) rotation of

$\mathfrak{X}_a$ and $\mathfrak{N}_a$ into physical ones (estimates of $\mathfrak{X}$ and $\mathfrak{N}$). For the first step, principal components analysis is usually used (PCA; see Appendix C.1). In TFA, the second step is performed by individual testing of targets $\mathbf{n}_{tar}$ of dimension S (i.e., possible rows of $\mathfrak{N}$ that are assumed to be available from prior knowledge) on the row space of $\mathbf{Y}$. Once A targets have been accepted (appended in $\mathfrak{N}$), $\mathfrak{X}$ can be determined as $\mathbf{Y}\mathfrak{N}^+$.

Here, it is desired to identify the stoichiometric matrix $\mathbf{N}$ from $\mathbf{Y}$, where $\mathbf{Y}$ typically represents $\mathbf{D}$ or $\mathbf{C}$. Thus, $\mathfrak{X}$ and $\mathfrak{N}$ correspond to either $\mathbf{X}$ and $\mathbf{N}$ or $\mathbf{X}_e$ and $\mathbf{N}_e$. Stoichiometric targets $\mathbf{n}_{tar}$ (i.e., possible rows of $\mathbf{N}$ that are assumed to be available from prior knowledge) are used for target testing. Bonvin and Rippin (1990) applied TFA to the RV-concentration matrix $\mathbf{D}$ (see 4.11). However, if data pre-treatment to reaction-variant form is not possible, target testing must be applied directly to $\mathbf{C}$ (see 4.4).

Remark 4.15 (Discrimination of reaction pathways)
TFA enables to identify only R independent stoichiometries $\mathbf{N}$ from $\mathbf{Y}$. Thus, it is only possible to discriminate between candidate reaction pathways whose row spaces are different. In many practical situations, however, candidate reaction pathways have the same row space and, therefore, cannot be discriminated on the basis of the data $\mathbf{Y}$ only. Consider, for example, seven candidate reaction pathways involving three species:

(1) $\mathcal{X}_1 \rightarrow \mathcal{X}_2$, $\mathcal{X}_1 \rightarrow \mathcal{X}_3$: $\mathbf{n}_{tar,11}^{\mathrm{T}} = \begin{bmatrix}-1 & 1 & 0\end{bmatrix}$, $\mathbf{n}_{tar,12}^{\mathrm{T}} = \begin{bmatrix}-1 & 0 & 1\end{bmatrix}$

(2) $\mathcal{X}_1 \rightarrow \mathcal{X}_2 \rightarrow \mathcal{X}_3$: $\mathbf{n}_{tar,21} = \mathbf{n}_{tar,11}$, $\mathbf{n}_{tar,22}^{\mathrm{T}} = \begin{bmatrix}0 & -1 & 1\end{bmatrix}$

(3) $\mathcal{X}_1 \rightarrow \mathcal{X}_3 \rightarrow \mathcal{X}_2$: $\mathbf{n}_{tar,31} = \mathbf{n}_{tar,12}$, $\mathbf{n}_{tar,32}^{\mathrm{T}} = \begin{bmatrix}0 & 1 & -1\end{bmatrix}$

(4) $\mathcal{X}_1 \rightarrow \mathcal{X}_2 \rightarrow \mathcal{X}_3$, $\mathcal{X}_1 \rightarrow \mathcal{X}_3$: $\mathbf{n}_{tar,41} = \mathbf{n}_{tar,11}$, $\mathbf{n}_{tar,42} = \mathbf{n}_{tar,22}$, $\mathbf{n}_{tar,43} = \mathbf{n}_{tar,12}$

(5) $\mathcal{X}_1 \rightarrow \mathcal{X}_2 \rightarrow \mathcal{X}_3 \rightarrow \mathcal{X}_1$: $\mathbf{n}_{tar,51} = \mathbf{n}_{tar,11}$, $\mathbf{n}_{tar,52} = \mathbf{n}_{tar,22}$, $\mathbf{n}_{tar,53}^{\mathrm{T}} = \begin{bmatrix}1 & 0 & -1\end{bmatrix}$

(6) $\mathcal{X}_1 \rightarrow \mathcal{X}_2$, $2\mathcal{X}_2 \rightarrow \mathcal{X}_1 + \mathcal{X}_3$: $\mathbf{n}_{tar,61} = \mathbf{n}_{tar,11}$, $\mathbf{n}_{tar,62}^{\mathrm{T}} = \begin{bmatrix}1 & -2 & 1\end{bmatrix}$

(7) $\mathcal{X}_1 \rightarrow \mathcal{X}_3$, $\mathcal{X}_1 + \mathcal{X}_3 \rightarrow 2\mathcal{X}_2$: $\mathbf{n}_{tar,71} = \mathbf{n}_{tar,12}$, $\mathbf{n}_{tar,72}^{\mathrm{T}} = \begin{bmatrix}-1 & 2 & -1\end{bmatrix}$.

A matrix of independent reactions describing all seven pathways is

$$\mathbf{N} = \begin{bmatrix}-1 & 1 & 0\\ -1 & 0 & 1\end{bmatrix}.$$

It can easily be verified that the stoichiometric targets of the seven physically-different reactions $\mathbf{n}_{tar,11}$, $\mathbf{n}_{tar,12}$, $\mathbf{n}_{tar,22}$, $\mathbf{n}_{tar,32}$, $\mathbf{n}_{tar,53}$, $\mathbf{n}_{tar,62}$, and $\mathbf{n}_{tar,72}$ are accepted by TFA using $\mathbf{N}$ or, equivalently, $\mathbf{Y}$. One way to discriminate between the various reaction pathways is to fit a kinetic model to the reaction data $\mathbf{Y}$. However, this step is beyond the scope of this dissertation.

4.3.2.2 TFA applied to concentration data in reaction-variant form

All species measured

If the data are brought in reaction-variant form (4.11) so that $\mathcal{S}_r(\mathbf{D}) = \mathcal{S}_r(\mathbf{N})$, target testing can be unambiguously performed. The result to this effect is stated below:

Proposition 4.16 *Let Assumptions A1–7 in Appendix D be verified. Let* $\mathbf{C}_{in}$, V_0, $\mathbf{c}(k)$, $\mathbf{q}_{in}(k)$, *and* $q_{out}(k)$ *of reaction systems described by* (2.38) *be known/measured for all* k. *Then, the following properties hold:*

(a) The number of independent reactions can be determined from $\mathbf{D}$ *in* (4.11) *as*

$$R = \operatorname{rank}(\mathbf{D}). \tag{4.34}$$

(b) $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$ *iff* $\varepsilon_p(\mathbf{n}_{tar}, \mathbf{D}) = 0$, *where* $\varepsilon_p(\mathbf{n}_{tar}, \mathbf{D})$ *is the least-squares error of* $\mathbf{n}_{tar}$ *on the row space of* $\mathbf{D}$ *(see C.15).*

(c) Once $\mathbf{N}$ *has been determined (e.g., after having accepted* R *stoichiometric targets),* $\mathbf{X}$ *and the vector of reaction rates* $\mathbf{r}(k)$ *can be reconstructed.*

(See Appendix H.3.4 for proof)

Unmeasured species

The above proposition uses data pre-treatment to reaction-variant form so that existing TFA techniques can be applied when the concentrations of all species are measured. In the presence of unmeasured species, however, unambiguous target acceptance is no longer possible using Proposition 4.16. Though true targets will not be rejected, false targets can be accepted, which is illustrated next: Let $\mathbf{n}_{m,tar}$ represent the part of the target that corresponds to the measured species. Although $\mathbf{n}_{m,tar} \notin \mathcal{S}_r(\mathbf{N}_m)$ implies $\mathbf{n}_{tar} \notin \mathcal{S}_r(\mathbf{N})$, $\mathbf{n}_{m,tar} \in \mathcal{S}_r(\mathbf{N}_m)$ does not imply $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$. To get necessary and sufficient conditions for the acceptance of the targets, reaction-invariant relationships contained in $\mathbf{\Gamma}$ are used.

Theorem 4.17
Let Assumptions A1–7 in Appendix D be verified. Let $\mathbf{C}_{in,m}$, V_0, $\mathbf{c}_m(k)$, $\mathbf{q}_{in}(k)$, *and* $q_{out}(k)$ *of reaction systems described by* (2.38) *be known/measured for all* k. *If* $\operatorname{rank}(\mathbf{\Gamma}_u) \geq S - \max(S_m, R)$, *then* $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$ *iff* $\varepsilon_p(\mathbf{n}_{m,tar}, \mathbf{D}_m) = 0$ *and* $\mathbf{n}_{tar}^{\mathrm{T}} \mathbf{\Gamma} = \mathbf{0}_N^{\mathrm{T}}$. *(See Appendix H.3.4 for proof)*

The assumption that $\mathbf{\Gamma}$ is known is not very restrictive when all S species can be chemically identified. Biotechnological reaction systems represent an important exception since, very often, the biomass composition is only approximately known.

Theorem 4.17 indicates that either (i) as many relationships from the reaction-invariance matrix are needed as there are unmeasured species, or (ii) $\operatorname{rank}(\mathbf{\Gamma}_u) = \operatorname{rank}(\mathbf{\Gamma}) = S - R$. In the latter case, target testing can be performed directly on the reaction-invariance matrix $\mathbf{\Gamma}$, without the need for measured data, since the columns of $\mathbf{\Gamma}$ span the null space of $\mathbf{N}$ (see Proposition 2.7b).

4.3.2.3 TFA applied to concentration data

All species measured

In Subsection 4.3.2.2, TFA of concentration data in reaction-variant form was discussed. The crucial point in data pre-treatment to reaction-variant form is that $\mathbf{q}_{in}$ and q_{out} (and V_0) need to be known. Furthermore, in the case of varying-density, the density must be known. However, if these quantities are unknown, TFA must be applied to the measured concentrations $\mathbf{C}$ (see 4.4) rather than to the *unknown* $\mathbf{D}$ (see 4.11).

When TFA is applied to $\mathbf{C}$, unambiguous target acceptance is no longer possible. This is due to the rank of $\mathbf{C}$ being larger than R (the row space of $\mathbf{C}$ is larger than that of $\mathbf{N}$ or, equivalently, the null space of $\mathbf{C}$ is smaller than that of $\mathbf{N}$). The situation is similar to that of unmeasured species discussed above: although true targets will not be rejected, false targets can be accepted. To resolve this ambiguity and provide necessary and sufficient conditions, it is again proposed to use the reaction-invariant relationships contained in $\mathbf{\Gamma}$.

Theorem 4.18
Let Assumptions A1–7 in Appendix D be verified and $\mathbf{c}(k)$ *of reaction systems described by* (2.38) *be measured for all* k. *If* $\mathbf{X}_e$ *is full rank, and* $\operatorname{rank}(\mathbf{C}_x\,\mathbf{\Gamma}) = p+1$, *then the following properties hold:*

(a) *The number of independent reactions can be determined from* $\mathbf{C}$ *as*

$$R = \operatorname{rank}(\mathbf{C}) - p - 1. \tag{4.35}$$

(b) $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$ *iff* $\varepsilon_p(\mathbf{n}_{tar}, \mathbf{C}) = 0$ *and* $\mathbf{n}_{tar}^{\mathrm{T}}\,\mathbf{\Gamma} = \mathbf{0}_N^{\mathrm{T}}$.

(c) *Given* $\mathbf{C}_{in}$, V_0, *and* $q_{out}(k)$ *for all* k, *once* $\mathbf{N}$ *has been determined (e.g. after having accepted* R *stoichiometric targets), the time profiles* $V(k)$, $\mathbf{z}(k)$ *(and* $\mathbf{q}_{in}(k)$*),* $\mathbf{x}(k)$, *and* $\mathbf{r}(k)$ *can be reconstructed.*

(See Appendix H.3.4 for proof)

Theorem 4.18 states that, by using the knowledge of the reaction-invariance matrix, unambiguous target testing is possible when concentration data are used directly. In particular, if either $\varepsilon_p(\mathbf{n}_{tar}, \mathbf{C}) \neq 0$ or $\mathbf{n}_{tar}^{\mathrm{T}}\,\mathbf{\Gamma} \neq \mathbf{0}_N^{\mathrm{T}}$, then $\mathbf{n}_{tar} \notin \mathcal{S}_r(\mathbf{N})$.

The two assumptions of Theorem 4.18, $\operatorname{rank}(\mathbf{C}_x\mathbf{\Gamma}) = (p+1)$ and $\mathbf{X}_e$ of full rank, should be viewed as experimental planning conditions rather than restrictions. Given $(p+1)$ linearly-independent columns of $\mathbf{\Gamma}$, it is always possible to plan the experiments (choice of $\mathbf{C}_x$) such that $\operatorname{rank}(\mathbf{C}_x\mathbf{\Gamma}) = p+1$. Also, Proposition 4.5 gives methods to experimentally guarantee full rank of $\mathbf{X}_e$.

Unmeasured species

A link between the use of $\mathbf{D}_m$ and that of $\mathbf{C}$ in TFA is apparent. One reaction-invariant relationship is required for every unmeasured species and for every inlet stream present. Also, the presence of the initial concentrations in $\mathbf{C}$ requires an invariant relationship.

In the extreme case where $S_u = (S - R)$ or $(p + 1) = (S - R)$, target testing relies exclusively on the reaction-invariant relationships and disregards the data.

For the case of unmeasured species, target testing on $\mathbf{C}_m$ is proposed that combines the two concepts developed for $\mathbf{D}_m$ (Theorem 4.17) and $\mathbf{C}$ (Theorem 4.18).

Theorem 4.19
Let Assumptions A1–7 in Appendix D be verified and $\mathbf{c}_m(k)$ *of reaction systems described by* (2.38) *be measured for all* k. *If* $\mathbf{X}_e$ *is full rank and* $\operatorname{rank}\left(\left[\begin{smallmatrix}\mathbf{C}_{x,m}\boldsymbol{\Gamma}_m\\ \boldsymbol{\Gamma}_u\end{smallmatrix}\right]\right) = S - max(S_m - p - 1, R)$, *then* $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$ *iff* $\varepsilon_p(\mathbf{n}_{m,tar}, \mathbf{C}_m) = 0$ *and* $\mathbf{n}_{tar}^{\mathrm{T}}\boldsymbol{\Gamma} = \mathbf{0}_N^{\mathrm{T}}$. *(See Appendix H.3.4 for proof)*

4.3.2.4 Discussion

The results for target testing can be summarized as follows. For unambiguous target testing, the nullity of the data matrix should be $(S - R)$, i.e., it should equal that of the stoichiometric matrix. However, the measured concentration matrix $\mathbf{C}$ has a null space that is smaller than $(S-R)$ due to the explicit presence of the reaction invariants (the effects of inlet streams and initial conditions).

If the flowrates and the initial volume are known *and* all species are measured, concentration data can be pre-treated so that the nullity is increased to $(S - R)$ (see also Proposition 4.13). For all the other cases, ambiguity arises in the testing procedure when considering only the projection error of the targets on the available data. To resolve the ambiguity, a subspace of the theoretical stoichiometric space has to be known. The size of this subspace depends on the number of measured species S_m and the number of reaction invariants:

$$\operatorname{rank}(\boldsymbol{\Gamma}) \geq S - \max(S_m - \alpha, R),$$

where $\alpha = \operatorname{rank}(\mathbf{C}_x) = (p + 1)$ for $\mathbf{C}$ and $\mathbf{C}_m$, while $\alpha = 0$ for $\mathbf{D}$ and $\mathbf{D}_m$.

According to Proposition 4.13, pre-treatment of $\mathbf{C}_m$ to reaction-variant form can improve the discrimination power with respect to stoichiometric targets. However, contrary to pre-treating $\mathbf{C}$, nullity increase cannot be guaranteed when pre-treating $\mathbf{C}_m$, since the number of observed independent reactions R_o can be smaller than the total number of independent reactions R. Note that the number of reaction-invariant relationships necessary for unambiguous target testing is independent of R_o, since the goal of TFA is to determine the *total* number of independent reactions R.

An advantage of using $\mathbf{C}$ is that, once the extended stoichiometric space has been determined, the (possibly varying) volume can be reconstructed.

4.3.2.5 Simulated example

The theoretical results presented above are illustrated on the reaction system described in Example 4.8 on page 67 ($S = 4$, $R = 2$). Simulations are conducted for the case of varying density and temperature. The first reaction follows the power law, while the second reaction is autocatalyzed by $\mathcal{X}_3$. Thus,

$$r_1(c_1, T) = \kappa_1(T)\, c_1^2, \qquad r_2(c_2, c_3, T) = \kappa_2(T)\, c_2\, c_3$$

with κ_i following the Arrhenius law $\kappa_i(T) = \kappa_{i,0} \exp(-E_i^*/T)$. The total density is expressed as:

$$\rho(\mathbf{w}, T) = \sum_{i=1}^{S} w_i \, \bar{\rho}_i(T), \quad \bar{\rho}_i(T) = \alpha_i \, (T - \beta_i)^2 + \gamma_i, \; \forall i = 1, \ldots, 4, \quad T_{min} < T < T_{max},$$

with $\bar{\rho}_i$ being the pure component density of Species i dependent on temperature T in the interval $[T_{min}, T_{max}]$ and the coefficients α_i, β_i, and γ_i. A temperature profile $T(t)$ (see Figure 4.10a) is imposed. The numerical values of the parameters are given in Appendix A.3. The reaction-invariance matrix $\mathbf{\Gamma}^{\mathrm{T}}$ is

$$\mathbf{\Gamma}^{\mathrm{T}} = \begin{bmatrix} \mathbf{\Gamma}_1^{\mathrm{T}} \\ \mathbf{\Gamma}_2^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} 50 & 100 & 75 & 12.5 \\ 122 & 244 & 66 & 89 \end{bmatrix}, \tag{4.36}$$

where the first row of $\mathbf{\Gamma}^{\mathrm{T}}$ is the vector of molecular weights.

Two operating modes are studied: batch (b), and semibatch (sb). The initial concentrations and the concentrations of the inlet stream are $\mathbf{c}_0^{\mathrm{T}} = [19, 0, 0.505, 0]$ M and $\mathbf{c}_{in}^{\mathrm{T}} = [16.15, 0, 0, 0]$ M, respectively. In semibatch mode, reactant $\mathcal{X}_1$ is fed at 373 K with the inlet rate shown in Figure 4.10b. The volume, density, and concentration profiles of the reaction system in batch and in semibatch modes are shown in Figure 4.10 and Figure 4.11, respectively.

$\mathbf{C}$ and $\mathbf{D}$ are constructed according to (4.4) and (4.11), respectively, with $K = 100$ observations equidistantly spaced. For the purpose of illustration, $\mathbf{c}(k)$ is corrupted by

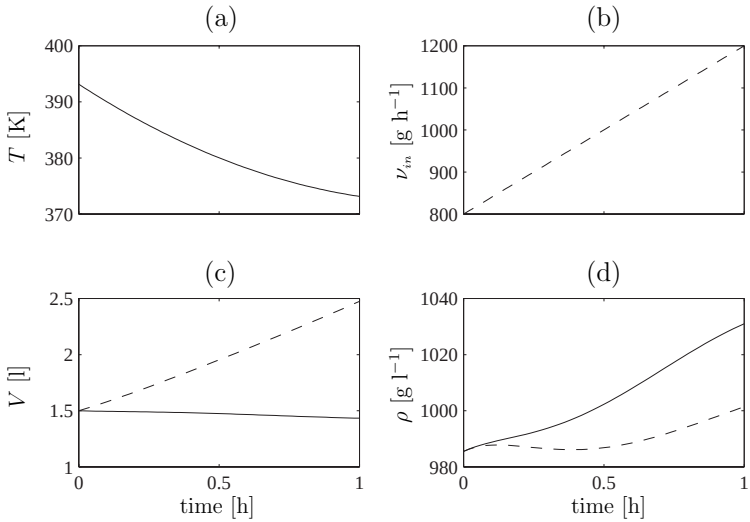

Figure 4.10. Simulated TFA example: (a) Temperature T, (b) inlet mass flowrate ν_{in}, (c) volume V, and (d) density ρ for the batch (−) and semibatch (−−) reaction systems.

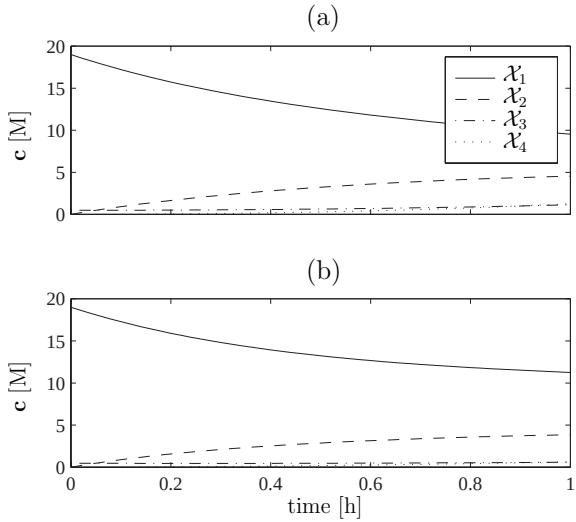

Figure 4.11. concentrations **c** for the simulated TFA example: (a) the batch and (b) semibatch reaction systems.

Table 4.3. Projection errors of true (t) and false (f) reaction targets ($\mathbf{n}_{tar,t}^{\mathrm{T}} = [-2, 1, 0, 0]$, $\mathbf{n}_{tar,f}^{\mathrm{T}} = [1, 1, -2, 0]$) in batch mode for different noise levels.

	ε_p of $\mathbf{n}_{tar,t}$ on		ε_p of $\mathbf{n}_{tar,f}$ on	
σ	**C**	**D**	**C**	**D**
0	0	0	1.836	2.009
0.05	0.008	0.012	1.837	2.012
0.1	0.018	0.031	1.842	2.020

Table 4.4. Projection errors of true (t) and false (f) reaction targets ($\mathbf{n}_{tar,t}^{\mathrm{T}} = [-2, 1, 0, 0]$, $\mathbf{n}_{tar,f}^{\mathrm{T}} = [1, 1, -2, 0]$) in semibatch mode for different noise levels.

	ε_p of $\mathbf{n}_{tar,t}$ on		ε_p of $\mathbf{n}_{tar,f}$ on	
σ	**C**	**D**	**C**	**D**
0	0	0	0	2.009
0.05	0	0.012	0	2.012
0.1	0	0.032	0	2.021

Table 4.5. Verification of $\mathbf{n}_{tar,i}^{\mathrm{T}}\boldsymbol{\Gamma} = \mathbf{0}_2^{\mathrm{T}}$ for true ($i = t$) and false ($i = f$) reaction targets ($\mathbf{n}_{tar,t}^{\mathrm{T}} = [-2, 1, 0, 0]$, $\mathbf{n}_{tar,f}^{\mathrm{T}} = [1, 1, -2, 0]$)

i	$\mathbf{n}_{tar,i}^{\mathrm{T}}\boldsymbol{\Gamma}_1$	$\mathbf{n}_{tar,i}^{\mathrm{T}}\boldsymbol{\Gamma}_2$
t	0	0
f	0	234

Table 4.6. Projection errors of true (t) and false (f) reaction targets ($\mathbf{n}_{tar,t}^{\mathrm{T}} = [-2, 1, 0, 0]$, $\mathbf{n}_{tar,f}^{\mathrm{T}} = [1, 1, -2, 0]$) in semibatch mode for different noise levels and unmeasured species.

	ε_p of $\mathbf{n}_{tar,t}$ on		ε_p of $\mathbf{n}_{tar,f}$ on	
σ	$\mathbf{D}_{m,C1}$	$\mathbf{D}_{m,C2}$	$\mathbf{D}_{m,C1}$	$\mathbf{D}_{m,C2}$
0	0	0	0	1.789
0.05	0	0.006	0	1.789
0.1	0	0.017	0	1.791

multiplicative zero-mean Gaussian noise with 3 different standard deviations, σ: 0, 0.05, 0.1. Since the sample covariances of $\mathfrak{N}_a$ used to compute ε_p in (C.15) depend strongly on the particular noise added (Jackson, 1991), extreme values for ε_p can occur. To avoid this, the ε_p values for 1000 calibration sets corrupted by different noise sequences are considered (Monte–Carlo simulation). The average of ε_p for each noise level is calculated. For a more detailed discussion on the impact of noise on TFA, the reader is referred to Bonvin and Rippin (1990); Malinowski (1991) and Appendix C.4.1.

Two stoichiometric targets are tested using $\mathbf{C}$ and $\mathbf{D}$ obtained from both the batch and the semibatch reaction systems: a true stoichiometric target $\mathbf{n}_{tar,t}^{\mathrm{T}} = [-2, 1, 0, 0]$, which corresponds to the stoichiometry of the first reaction, and a false stoichiometric target $\mathbf{n}_{tar,f}^{\mathrm{T}} = [1, 1, -2, 0]$. As can be seen in Tables 4.3 and 4.4, the true stoichiometric target is always accepted, since $\varepsilon_p \approx 0$ for both modes and, for the reaction in semibatch mode, the conditions $\mathbf{n}_{tar,t}^{\mathrm{T}}\boldsymbol{\Gamma} = \mathbf{0}_2^{\mathrm{T}}$ and $\mathrm{rank}\,(\mathbf{C}_x\boldsymbol{\Gamma}) = p + 1 = 2$ are fulfilled. However, when only the projection error on the data matrices is considered, then for $\mathbf{C}$ from the reaction in semibatch mode, the false stoichiometric target is (erroneously) accepted. Yet, Table 4.5 shows that it is rejected by the condition $\mathbf{n}_{tar}^{\mathrm{T}}\boldsymbol{\Gamma} = \mathbf{0}_2^{\mathrm{T}}$: though $\boldsymbol{\Gamma}_1$ cannot detect the false target ($\mathbf{n}_{tar,f}^{\mathrm{T}}\boldsymbol{\Gamma}_1 = 0$; $\mathbf{n}_{tar,f}$ was chosen to that effect to stress the need for *two* linearly-independent columns of $\boldsymbol{\Gamma}$), $\boldsymbol{\Gamma}_2$ can ($\mathbf{n}_{tar,f}^{\mathrm{T}}\boldsymbol{\Gamma}_2 \neq 0$).

The pseudo-ranks of $\mathbf{C}$ and $\mathbf{D}$, based on the reduced F-test with 95% confidence limit (see C.32), for the simulated batch and semibatch reaction systems correspond to the ranks predicted by equations (4.35) and (4.34): (i) the rank of $\mathbf{C}$ equals $(R+1) = 3$ for the batch mode and $(R+p+1) = 4$ for the semibatch mode, and (ii) $\mathrm{rank}\,(\mathbf{D}) = R = 2$ for both the batch and semibatch modes.

Two cases are considered regarding the problem of unmeasured species: in Case C1,

the concentrations of only Species $\mathcal{X}_1$ and $\mathcal{X}_3$ are measured from the reaction system in semibatch mode, while in Case C2, Species $\mathcal{X}_1$, $\mathcal{X}_3$, and $\mathcal{X}_4$ are measured for the same mode. As can be seen in Table 4.6, the true stoichiometric target is always accepted based on the projection error alone since $\varepsilon_p(\mathbf{n}_{m,t,tar}, \mathbf{D}_m) \approx 0$ for both cases. Furthermore, the false stoichiometric target is rejected on $\mathbf{D}_m$ in Case C2. However, similarly to $\mathbf{C}$ in semibatch mode, the false stoichiometric target is (erroneously) accepted by $\mathbf{D}_m$ in Case C1.

In Case C1, $\mathbf{D}_m$ is useless for target testing. For unambiguous target testing, Theorem 4.17 requires $\operatorname{rank}(\mathbf{\Gamma}_u) = S - \max(S_m, R) = 2$, which implies $\operatorname{rank}(\mathbf{\Gamma}) \geq 2$. Thus, at least two linearly-independent reaction-invariant relationships must be available, which is the case with (4.36). Table 4.5 shows that the false target is rejected by the condition $\mathbf{n}_{tar}^{\mathrm{T}}\mathbf{\Gamma} = \mathbf{0}_2^{\mathrm{T}}$.

In Case C2, the false target is rejected by $\mathbf{D}_m$. Thus, the reaction-invariant relationships of Theorem 4.17 remain to be verified for the true target: $\operatorname{rank}(\mathbf{\Gamma}_u)$ must be 1, which implies $\operatorname{rank}(\mathbf{\Gamma}) \geq 1$. Thus, at least one reaction-invariant relationship must be available, either $\mathbf{\Gamma}_1$ or $\mathbf{\Gamma}_2$. The condition $\mathbf{n}_{tar,t}^{\mathrm{T}}\mathbf{\Gamma}_i = 0$ is fulfilled for $i = 1, 2$ and, thus, the true target is (correctly) accepted.

Note that two linearly-independent columns of $\mathbf{\Gamma}$ are assumed to be known to enable TFA using $\mathbf{C}$ in semibatch mode ($p + 1 = 2$) or $\mathbf{D}_m$ in Case C1 ($S_u = 2$). In this example, however, $(S - R)$ also equals 2. Thus, target testing can unambiguously be performed on $\mathbf{\Gamma}$ without knowledge of $\mathbf{C}$ or $\mathbf{D}_m$, i.e., from $\mathbf{n}_{tar,t}^{\mathrm{T}}\mathbf{\Gamma} = \mathbf{0}_2^{\mathrm{T}}$ and $\mathbf{n}_{tar,f}^{\mathrm{T}}\mathbf{\Gamma} \neq \mathbf{0}_2^{\mathrm{T}}$ it can be directly concluded that $\mathbf{n}_{tar,t}$ and $\mathbf{n}_{tar,f}$ is a true and a false stoichiometry, respectively.

It is interesting to note that the projection errors on $\mathbf{C}$ in semibatch mode are identically 0 also in the noisy case, since $\mathbf{C}$ is of full rank ($R + p + 1 = S = 4$). Therefore, any S-dimensional stoichiometric target will be accepted with projection error being 0 when considering only the projection on $\mathbf{C}$. A similar remark also applies to $\mathbf{D}_m$ in Case C1, since $\mathbf{D}_m$ is of full rank $S_m = 2$.

4.4 Extensions

In Section 4.1, the factorization of concentration data (4.4) and (4.11) were derived for the cases of constant density and temperature, and process runs with the same stoichiometries. In the following subsections, each assumption is relaxed separately. Based on these extensions, the factorization of concentration data from reaction systems for which any combination of the assumptions is relaxed can easily be derived and are, therefore, omitted here. However, reaction systems expressed in weight fractions is studied, and extensions of the factorization of concentration data to include temperature or calorimetric measurements are introduced.

4.4.1 Reaction systems with varying density

The elements of $\mathcal{H}$ are now determined by:

$$\mathrm{h}(\mathbf{x}, \mathbf{z}, \lambda) := (V_0\rho_0 + \mathbf{1}_p^{\mathrm{T}}\mathbf{\Phi}_{in}\mathbf{z})/\rho(\mathbf{x}, \mathbf{z}, \lambda). \tag{4.37}$$

This expression was obtained by choosing $\eta = 1$ and substituting $\mathbf{z}$ from (4.2) into (3.21).

4.4.2 Nonisothermal reaction systems

Since the temperature only appears in $\mathbf{r}(\mathbf{c}, T)$ which, in turn, appears in the (unknown) vector $\mathbf{x}$, the factorization of concentration data from exothermic/endothermic reaction systems conserve the structures of (4.4) and (4.11).

4.4.3 Reaction systems expressed in weight fractions

Since the dynamic models for weight fractions (see 2.42) and concentrations (see 2.38) are structurally similar, the results obtained herein for concentrations can be extended directly to weight fraction measurements.

With the quantities defined in Subsection 2.4.1, the factorization of weight fraction data become:

$$\mathbf{W} = \left\{ \mathcal{H}_m^{-1} \left[\begin{array}{c|c|c} \mathbf{X} & \mathbf{Z} & \mathfrak{m}_0 \mathbf{1}_K \end{array} \right] \right\} \left[\begin{array}{c} \mathbf{N}_w \\ \hline \mathbf{W}_{in}^{\mathrm{T}} \\ \hline \mathbf{w}_0^{\mathrm{T}} \end{array} \right] =: \mathbf{X}_e \, \mathbf{N}_e, \qquad \mathcal{H}_m \equiv \mathbf{M} \, \mathbf{\Lambda}^{-1}, \tag{4.38}$$

$$\mathbf{D}_w \equiv \mathcal{H}_m \mathbf{W} \, \mathbf{M}_w^{-1} - \mathbf{Z} \, \mathbf{W}_{in}^{\mathrm{T}} \mathbf{M}_w^{-1} - \mathfrak{m}_0 \, \mathbf{1}_K \mathbf{w}_0^{\mathrm{T}} \mathbf{M}_w^{-1} = \mathbf{X} \, \mathbf{N}, \tag{4.39}$$

where $\mathbf{W}$ is the $K \times S$ weight fraction matrix, $\mathbf{M} = \mathrm{diag}\,(\mathfrak{m}(1), \mathfrak{m}(2), \ldots, \mathfrak{m}(K))$, and $\mathcal{H}_m$ a diagonal scaling matrix of dimension K. For the construction of $\mathbf{D}$, it was assumed that $\mathfrak{m}_0$, $\mathbf{M}_w$, $\mathbf{W}_{in}$, $\mathbf{w}_0$, $\mathbf{w}(k)$, $\boldsymbol{\nu}_{in}(k)$, and $\nu_{out}(k)$ are either known or measured for K observations.

Note that the extended stoichiometric matrices for weight fractions and concentrations are different and, thus, also the row spaces. However, the row spaces of $\mathbf{D}$ and $\mathbf{D}_w$ are the same (space spanned by the rows of $\mathbf{N}$). Note that, in contrast to the computation of $\mathbf{D}$ (4.11), the volume V need not to be known for the computation of $\mathbf{D}_w$ (4.39). This might represent a decisive advantage of weight fraction data over concentration data in chemical reaction systems with varying density. The following corollary results from Proposition 4.16 for weight fraction measurements in reaction-variant form (4.39).

Corollary 4.20 *Let Assumptions A1–7 in Appendix D be verified. Let* $\mathbf{W}_{in}$, $\mathfrak{m}(0)$, $\mathbf{M}_w$, $\mathbf{w}(k)$, $\boldsymbol{\nu}_{in}(k)$, *and* $\nu_{out}(k)$ *of reaction systems described by* (2.42) *be known/measured for all* k. *Then, the following properties hold:*

(a) *The number of independent reactions can be determined from* $\mathbf{D}_w$ *in* (4.39) *as*

$$R = \mathrm{rank}\,(\mathbf{D}_w)\,. \tag{4.40}$$

(b) $\mathbf{n}_{tar} \in \mathcal{S}_r\,(\mathbf{N})$ *iff* $\varepsilon_p(\mathbf{n}_{tar}, \mathbf{D}_w) = 0$.

(c) *Once* $\mathbf{N}$ *has been determined (e.g., after having accepted* R *stoichiometric targets), the reaction-variant matrix* $\mathbf{X}$ *can be reconstructed. If, furthermore, the*

volume $V(k)$ is known for all k, then the vector of reaction rates $\mathbf{r}(k)$ can be reconstructed.

If the mass is unknown, a corollary of Theorem 4.18 for weight fraction measurements $\mathbf{W}$ is stated below.

Corollary 4.21 *Let Assumptions A1–7 in Appendix D be verified. If* $\operatorname{rank}(\mathbf{W}_x \boldsymbol{\Gamma}) = (p+1)$ *with* $\mathbf{W}_x = \begin{bmatrix} \mathbf{W}_{in}^{\mathrm{T}} \\ \mathbf{w}_0^{\mathrm{T}} \end{bmatrix}$ *and* $\mathbf{X}_e$ *is full rank, then the following properties hold:*

(a) The number of independent reactions can be determined from $\mathbf{W}$ *as*

$$R = \operatorname{rank}(\mathbf{W}) - p - 1. \tag{4.41}$$

(b) $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$ *iff* $\varepsilon_p(\mathbf{n}_{tar}, \mathbf{W}) = 0$ *and* $\mathbf{n}_{tar}^{\mathrm{T}} \boldsymbol{\Gamma} = \mathbf{0}_N^{\mathrm{T}}$.

(c) Given $\mathbf{W}_{in}$, $\mathfrak{m}_0$, *and* $\nu_{out}(k)$ *for all k, once* $\mathbf{N}$ *has been determined (e.g. after having accepted R stoichiometric targets), the time profiles* $\mathbf{z}(k)$ *(and* $\boldsymbol{\nu}_{in}(k)$*) and* $\mathbf{x}(k)$ *can be reconstructed. In addition, if the volume $V(k)$ is known, the reaction rates* $\mathbf{r}(k)$ *can be computed.*

A notable difference in using $\mathbf{C}$ and $\mathbf{W}$ is that, though the volume can be reconstructed from $\mathbf{C}$ according to Theorem 4.18c, it cannot be obtained from $\mathbf{W}$. Not knowing the volume prevents the computation of the reaction rates $\mathbf{r}(k)$ from $\mathbf{W}$.

4.4.4 Process runs with different stoichiometries

In the case of various process runs, the S species can undergo chemical reactions exhibiting different linearly-independent stoichiometries, i.e., the row spaces of the various stoichiometric matrices can be distinct.

Factorization of concentration data

Let K^j denote the number of observations of the jth process run ($j = 1, \ldots, B$). Then,

$$\mathbf{C} := \begin{bmatrix} \mathbf{C}^1 \\ \ldots \\ \mathbf{C}^B \end{bmatrix} = \mathbf{X}_e \mathbf{N}_e = \mathcal{H}^{-1} \left[\begin{array}{c|c} \mathbf{X} & \mathbf{Z}_x \end{array}\right] \begin{bmatrix} \mathbf{N} \\ \hline \mathbf{C}_x \end{bmatrix}, \qquad \mathbf{D} := \begin{bmatrix} \mathbf{D}^1 \\ \ldots \\ \mathbf{D}^B \end{bmatrix} = \mathbf{X}\,\mathbf{N} \tag{4.42}$$

$$\mathbf{N} := \begin{bmatrix} \mathbf{N}^1 \\ \ldots \\ \mathbf{N}^B \end{bmatrix}, \ \mathbf{C}_x := \begin{bmatrix} \mathbf{C}_{in}^{\mathrm{T}} \\ \mathbf{C}_0^{\mathrm{T}} \end{bmatrix}, \ \mathbf{X} := \begin{bmatrix} \mathbf{X}^1 & \cdots & \mathbf{0}_{K^1 \times R^B} \\ \ldots & \ldots & \ldots \\ \mathbf{0}_{K^B \times R^1} & \cdots & \mathbf{X}^B \end{bmatrix}, \tag{4.43}$$

$$\mathbf{Z}_x := \left[\begin{array}{ccc|ccc} \mathbf{Z}^1 & \cdots & \mathbf{0}_{K^1 \times p^B} & V_0^1 \mathbf{1}_{K^1} & \cdots & \mathbf{0}_{K^1} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \mathbf{0}_{K^B \times R^1} & \cdots & \mathbf{Z}^B & \mathbf{0}_{K^B} & \cdots & V_0^B \mathbf{1}_{K^B} \end{array}\right], \ \mathbf{C}_{in} := \left[\mathbf{C}_{in}^1 \ \cdots \ \mathbf{C}_{in}^B\right],$$

$$\mathcal{H} := \begin{bmatrix} \mathbf{V}^1(\mathbf{\Lambda}^1)^{-1} & \cdots & \mathbf{0}_{K^B \times K^B} \\ \cdots & \cdots & \cdots \\ \mathbf{0}_{K^1 \times K^1} & \cdots & \mathbf{V}^B(\mathbf{\Lambda}^B)^{-1} \end{bmatrix}, \quad \mathbf{C}_0 \equiv \begin{bmatrix} \mathbf{c}_0^1 & \cdots & \mathbf{c}_0^B \end{bmatrix}, \tag{4.44}$$

with $R = \sum_{j=1}^{B} R^j$, $p = \sum_{j=1}^{B} p^j$, $K = \sum_{j=1}^{B} K^j$, $\mathbf{C}$ $(K \times S)$, $\mathbf{D}$ $(K \times S)$, $\mathbf{X}_e$ $(K \times R + p + 1)$, $\mathbf{X}$ $(K \times R)$, $\mathbf{Z}_x$ $(K \times (p+1))$, $\mathbf{N}_e$ $(R + p + 1 \times S)$, $\mathbf{N}$ $(R \times S)$, $\mathbf{C}_x$ $((p+1) \times S)$, $\bar{\mathbf{V}}$ $(K \times K)$, $\mathbf{V}_0$ $(K \times B)$, $\mathbf{C}_{in}$ $(S \times p)$, and $\mathbf{C}_0$ $(S \times B)$.

To stay consistent with the definition of R being the number of independent reactions, it is assumed that $\text{rank}\,(\mathbf{N}) = R$, i.e., the stoichiometries of the different reaction systems are mutually linearly independent.

Note that the structures of the factorization of concentration data (4.4) and (4.11) are conserved by (4.42).

Example 2.4c (Multiple runs and dependent kinetics; cont'd from page 19**)**

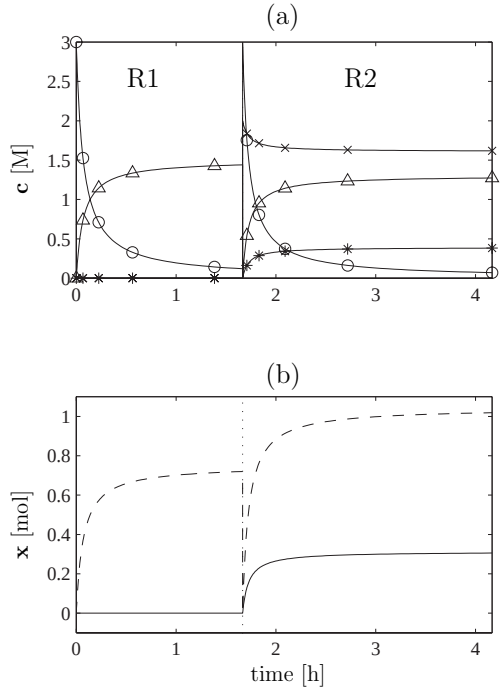

Figure 4.12. (a) The concentrations $\mathbf{c}$, and (b) the reaction variants $\mathbf{x}$ of Example 2.4c. The legend for (a) is: $\mathcal{X}_1$ (∘), $\mathcal{X}_2$ (×), $\mathcal{X}_3$ (∗), $\mathcal{X}_4$ (△); and the legend for (b): first reaction (–), second reaction (– –). (Run R1 and R2 for $t \in [0,\ 1.67]$h and $t \in]1.67,\ 4.17]$h, respectively).

Consider the irreversible parallel reaction system (2.8) with linearly-dependent kinetic rates for $c_2(t) > 0$:

$$r_1(t) = \begin{cases} 0.3\, r_2(t) & c_2(t) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Two batch runs R1 and R2 are conducted with $\mathbf{c}_{0,R1} = [3,\ 0,\ 0,\ 0]^{\mathrm{T}}$ M and $\mathbf{c}_{0,R2} = [3,\ 2,\ 0,\ 0]^{\mathrm{T}}$ M. $\mathbf{c}_{0,R1}$ has been chosen such that only the second reaction is active in the first process run, whereas the initial concentrations $\mathbf{c}_{0,R2}$ activates both reactions in the second process run (see Appendix A.4 for the parameters). Figure 4.12 presents the concentrations $\mathbf{c}$, and the matrix of reaction variants $\mathbf{x}$ of the two runs. It can be seen that for Run R1, the variant of the first reaction is zero, as expected, since the first reaction is not activated. Furthermore, it can be seen that for Run R2, the reaction variants are proportional (linearly dependent).
Three cases are studied. Cases C1, C2, and C3 consider Run R1, Run R2, and Runs R1 & R2. For Cases C1 and C2, according to Cases C1–C2 in Example 2.4b on page 28:

$$\mathbf{N}_{C1} = \begin{bmatrix} -2 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{N}_{C2} = \begin{bmatrix} -2.3 & -0.3 & 0.3 & 1 \end{bmatrix}.$$

Note that due to the linearly-dependent kinetics, the number of independent reactions in Case C2 is only $R = 1$. Thus, Cases C1 and C2 can be considered as reaction systems with two different linearly-independent stoichiometries. When appending Runs R1 and R2 (Case C3), then according to (4.43), one way to construct $\mathbf{N}$ is to append the stoichiometries of $\mathbf{N}_{C1}$ and $\mathbf{N}_{C2}$. Here, however, the augmented stoichiometric matrix can also be chosen to be the stoichiometric matrix of the physical reactions:

$$\mathbf{N}_{C3} = \begin{bmatrix} -1 & -1 & 1 & 0 \\ -2 & 0 & 0 & 1 \end{bmatrix} \quad \text{with} \quad \mathbf{C}_{in,C3} = \mathbf{c}_{0,R1}^{\mathrm{T}}.$$

Rank-increasing operation

Let $S = S_+ = S_-$. By appending concentration data from a reaction system with p^j inlet streams and $\mathbf{N}^j$ to concentration data whose rows of the stoichiometric matrix $\mathbf{N}_-$ do not lie in the subspace spanned by the rows of $\mathbf{N}^j$, the stoichiometric matrix of the augmented data matrix becomes $R_+ = R_- + R^j$. Since (i) the initial and inlet concentrations and the stoichiometries of the jth run may be linearly dependent on the rows of $\mathbf{N}_{e,-}$, and (ii) a variation can never reduce the rank, then only an interval for the rank increase Δd can be given: $0 \leq \Delta d \leq (R^j + p^j + 1)$. This is illustrated in the example below.

Example 2.4d (Multiple runs and dependent kinetics; cont'd from pageon the facing page**)**
The numbers of independent reactions of Run R1 and R2 (Cases C1 and C2 in Example 2.4c, respectively) are $R^1 = R^2 = 1$. Furthermore, the rank of $\mathbf{C}$ of Run R1 (Case C1) is given by $\operatorname{rank}(\mathbf{C}_{C1}) = \operatorname{rank}(\mathbf{N}_{e,C1}) = 2$, since the rank of $\mathbf{X}_{e,C1}$ is full.
Consider the case of appending Runs R1 and R2 (Case C3 in Example 2.4c). Then, $\mathbf{N}_- = \mathbf{N}_{C1}$ and, since the row spaces of $\mathbf{N}_{C1}$ and $\mathbf{N}_{C2}$ are distinct, $\mathbf{N}_+ = \begin{bmatrix} \mathbf{N}_{C1} \\ \mathbf{N}_{C2} \end{bmatrix}$. Since the initial condition of Run R2 is linearly independent, $\operatorname{rank}(\mathbf{C}_+) = \operatorname{rank}(\mathbf{N}_{e,+}) = 4$ and, thus, a rank increase occurs with $\Delta d = R^2 + 1 = 2$.

4.4.5 Factorization of extended concentration data

Extensions of the factorization of concentration data to include temperature or calorimetric measurements are studied. Furthermore, for $S_m < S$ measured species, it will be shown that temperature/calorimetric measurements can contain (valuable) additional information to the concentration measurements.

Factorization of concentration data

Calorimetric measurements describe the *total* heat production or consumption due to chemical reactions. They represent one-channel responses constructed from various measured quantities such as reactor temperature, inlet temperatures, inlet and outlet flowrates, and external power.

Bonvin and Rippin (1990) have proposed to augment composition change data from constant-volume, constant-density reaction systems with calorimetric measurements. Here, the methodology is extended to chemical reaction systems with inlet and outlet streams. If calorimetric measurements cannot be constructed, measurements of the reactor temperature from nonisothermal reaction systems can be used instead. It is shown below how the factorization of concentration data (4.4) and (4.11) must be modified to include temperature or calorimetric measurements.

Similarly to the factorization of concentration data (4.4) for the isothermal case, one can write:

$$\mathbf{C}_T \equiv \left[\mathbf{C} \quad \mathfrak{T}^\dagger\right] = \left\{\mathfrak{H}^{-1}\left[\begin{array}{c|c|c} \mathbf{X} & \mathbf{Z} & V_0\,\mathbf{1}_K \end{array}\right]\right\}\left[\begin{array}{c} \mathbf{N}_T \\ \hline \mathbf{C}_{T,in}^{\mathrm{T}} \\ \hline \mathbf{c}_{T,0} \end{array}\right] \equiv \mathbf{X}_{e,T}\,\mathbf{N}_{e,T}, \tag{4.45}$$

where $\mathfrak{T}^\dagger$ is the K-dimensional modified temperature vector (see 2.36), the dimensions of $\mathbf{C}_T$, $\mathbf{Z}$, $\mathbf{N}_{e,T}$, and $\mathbf{X}_{e,T}$ are $K \times S_T$, $K \times p_T$, $(R+p_T+1) \times S_T$, and $K \times (R+p_T+1)$, respectively. The elements of λ and the rows of $\mathbf{X}$ and $\mathbf{Z}$ are described by:

$$\begin{aligned} \dot{\mathbf{x}} &:= \mathfrak{h}(\mathbf{z}_n)\,\mathbf{r}(\mathbf{c}), & \mathbf{x}(0) &= \mathbf{0}_R, \\ \dot{\mathbf{z}} &:= \mathbf{q}_{T,in}/\lambda, & \mathbf{z}(0) &= \mathbf{0}_{p_T}, \\ \dot{\lambda} &= -q_{out}/\mathfrak{h}(\mathbf{z}_n), & \lambda(0) &= 1, \end{aligned} \tag{4.46}$$

where $\mathbf{z} = \left[\begin{smallmatrix} \mathbf{z}_n \\ z_T \end{smallmatrix}\right]$ with $\mathbf{z}_n$ and z_T corresponding to $\mathbf{q}_{in}$ and $\dot{Q}_{ext}$, respectively. (4.45) is called *the factorization of extended concentration data*.

If $\mathbf{C}_{in}$, $\mathfrak{C}_{p,in}$, $\mathbf{T}_{in}$, $\mathbf{\Phi}_{in}$, $\mathbf{c}(k)$, $T^\dagger(k)$, $\dot{Q}_{co}(k)$, $\dot{Q}_{sti}(k)$, $\mathbf{q}_{in}(k)$, $q_{out}(k)$, and $\mathbf{c}_p(k)$ are known or measured for K observations, a $K \times S_T$ extended RV-concentration matrix, $\mathbf{D}_T$, can be constructed from (4.45) that relates directly to $\mathbf{X}$ and $\mathbf{N}_T$:

$$\mathbf{D}_T \equiv \left[\mathbf{D} \quad \mathbf{d}_T\right] = \mathfrak{H}\,\mathbf{C}_T - \mathbf{Z}\,\mathbf{C}_{T,in}^{\mathrm{T}} - V_0\,\mathbf{1}_K\,\mathbf{c}_{T,0}^{\mathrm{T}} = \mathbf{X}\,\mathbf{N}_T, \tag{4.47}$$

where $\mathbf{d}_T$ is the K-dimensional calorimetric vector. (4.47) is called the *factorization of extended concentration data in reaction-variant form*. Note that in an isothermally-controlled reaction system, the reactor temperature stays constant, i.e., $\mathfrak{T}^\dagger = T_0 \mathfrak{c}_{p,0}\,\rho_0\,\mathbf{1}_K$, while the calorimetric measurements $\mathbf{d}_T$ may still vary with respect

to time. Thus, only the factorization of extended concentration data in reaction-variant form is useful in such a case.

Note that all the results derived for the factorization of concentration data (4.4) and (4.11) are readily applicable to (4.45) and (4.47) for which $\mathbf{C} := \mathbf{C}_T$, $\mathbf{D} := \mathbf{D}_T$, $\mathbf{X}_e := \mathbf{X}_{e,T}$, $\mathbf{N}_e := \mathbf{N}_{e,T}$, $\mathbf{C}_x := \mathbf{C}_{x,T}$, $\mathbf{N} := \mathbf{N}_T$, $p := p_T$, and $S := S_T$.

Rank increase for $S_m < S$

For $S_m < S$ measured species,

$$\mathbf{D}_{T,m} \equiv \begin{bmatrix} \mathbf{D}_m & \mathbf{d}_T \end{bmatrix}, \quad \mathbf{N}_{T,m} \equiv \begin{bmatrix} \mathbf{N}_m & -\Delta\mathbf{h}_R \end{bmatrix} \tag{4.48}$$

$$\mathbf{C}_{T,m} \equiv \begin{bmatrix} \mathbf{C}_m & \mathfrak{T}^\dagger \end{bmatrix}, \quad \mathbf{N}_{e,T,m} \equiv \begin{bmatrix} \mathbf{N}_m & -\Delta\mathbf{h}_R \\ \mathbf{C}_{in,m}^{\mathrm{T}} & \mathbf{T}_{in}^\dagger \\ \mathbf{c}_{0,m}^{\mathrm{T}} & T_0^\dagger \end{bmatrix}, \tag{4.49}$$

where $\mathbf{D}_{T,m}$, $\mathbf{N}_{T,m}$, $\mathbf{C}_{T,m}$, and $\mathbf{N}_{e,T,m}$ are matrices of dimension $K \times (S_m + 1)$, $(R + 1) \times (S_m + 1)$, $K \times (S_m + 1)$, and $R + p_T + 1 \times (S_m + 1)$, respectively.

Proposition 4.22 *Let Assumptions A1–4 and A6–7 in Appendix D be verified. Furthermore, assume that concentration measurements for $S_m < S$ species and temperature/calorimetric measurements from reaction system* (2.40) *are available, and the R_o observed independent reactions have a significant heat effect. Then, the following properties hold:*

(a) If $(-\Delta\mathbf{h}_R) \notin \mathcal{S}_c(\mathbf{N}_m)$, then

$$\operatorname{rank}(\mathbf{D}_{T,m}) = \operatorname{rank}(\mathbf{D}_m) + 1 = R_o + 1, \tag{4.50}$$

where $\mathbf{D}_{T,m}$ is defined as in (4.48).

(b) If $(-\Delta\mathbf{h}_R) \notin \mathcal{S}_c(\mathbf{N}_m)$ and $\operatorname{rank}(\mathbf{X}_{e,T}) = R + p_T + 1$, then

$$\operatorname{rank}(\mathbf{C}_{T,m}) = \operatorname{rank}(\mathbf{C}_m) + 1 \leq \min(R_o + p + 1, S_m) + 1, \tag{4.51}$$

where $\mathbf{C}_{T,m}$ is defined as in (4.49).

(See Appendix H.3.2 for proof)

Proposition 4.22 says that appending of calorimetric measurements to the RV-concentration matrix leads to a rank increase of 1, if $(-\Delta\mathbf{h}_R) \notin \mathcal{S}_c(\mathbf{N}_m)$. This situation can be encountered when $S_m < S$. If, in addition, $\mathbf{X}_{e,T}$ is full rank, then rank increase of 1 is also observed for concentration data. Thus, in such cases, calorimetric/temperature measurements contain (valuable) additional information.

4.5 Special cases

In this section, factorization of concentration data are considered for the following special cases: (i) continuous stirred-tank reaction systems (CSTR), (ii) batch or semi-batch reaction systems, and (iii) reactions in quasi-equilibrium conditions. Furthermore, non-reacting data exhibiting closure are studied. All special cases are based on

the assumptions of constant density and temperature but could easily be extended to the cases of varying density and temperature.

4.5.1 CSTR

For CSTRs, $V(t) = V_0$. Thus, from (4.5), $\mathfrak{h}(\mathbf{z}(k)) = V_0/\lambda(k)$, and the factorization of concentration data conserve the structures of (4.4) and (4.11). Thus,

$$\mathbf{C} = \frac{1}{V_0}\,\mathbf{\Lambda}\,\begin{bmatrix}\mathbf{X} & \mathbf{Z} & V_0\,\mathbf{1}_K\end{bmatrix}\begin{bmatrix}\mathbf{N}\\ \mathbf{C}_{in}^{\mathrm{T}}\\ \mathbf{c}_0^{\mathrm{T}}\end{bmatrix}, \quad \mathbf{D} = V_0\mathbf{\Lambda}^{-1}\,\mathbf{C} - \mathbf{Z}\,\mathbf{C}_{in}^{\mathrm{T}} - V_0\mathbf{1}_K\mathbf{c}_0^{\mathrm{T}}, \tag{4.52}$$

4.5.2 Batch and semibatch reaction systems

For reaction systems in semibatch mode ($\mathbf{q}_{in} \neq \mathbf{0}_p$, $q_{out} = 0$), $\mathbf{\Lambda} = \mathbf{I}_K$ and, for reaction systems in batch mode ($p = 0$, $q_{out} = 0$), $\mathbf{\Lambda} = \mathbf{I}_K$ and $\mathbf{V} = V_0\mathbf{I}_K$. Thus, the factorization of concentration data (4.4) and (4.11) for semibatch and batch reaction systems simplify to:

$$\mathbf{C} = \mathbf{V}^{-1}\begin{bmatrix}\mathbf{X} & \mathbf{Z} & V_0\,\mathbf{1}_K\end{bmatrix}\begin{bmatrix}\mathbf{N}\\ \mathbf{C}_{in}^{\mathrm{T}}\\ \mathbf{c}_0^{\mathrm{T}}\end{bmatrix}, \quad \mathbf{D} = \mathbf{V}\,\mathbf{C} - \mathbf{Z}\,\mathbf{C}_{in}^{\mathrm{T}} - V_0\mathbf{1}_K\mathbf{c}_0^{\mathrm{T}}, \tag{4.53}$$

$$\mathbf{C} = \begin{bmatrix}\frac{1}{V_0}\mathbf{X} & \mathbf{1}_K\end{bmatrix}\begin{bmatrix}\mathbf{N}\\ \mathbf{c}_0^{\mathrm{T}}\end{bmatrix}, \quad \mathbf{D} = V_0\,(\mathbf{C} - \mathbf{1}_K\mathbf{c}_0^{\mathrm{T}}). \tag{4.54}$$

Rank analysis

First, the rank of $\mathbf{C}$ from batch reaction systems is studied. Then, a non-exhaustive list of practical situations is examined in which column-mean centering of $\mathbf{C}$ causes a rank drop. This is equivalent to checking whether the column space of $\mathbf{C}$ contains the $\mathbf{1}_K$-vector.

Proposition 4.23 *Let Assumptions A1–7 in Appendix D be verified. Then, the rank of* $\mathbf{C}$ *of batch reaction system is determined by:*

$$\operatorname{rank}(\mathbf{C}) = R + 1. \tag{4.55}$$

(See Appendix H.3.5 for proof)

Proposition 4.23 says that the rank of $\mathbf{C}$ of batch reaction systems is determined exactly by the number of independent reactions plus 1 from the initial conditions.

For $\mathbf{N}_e$ being full rank, it is shown next that, in the special cases of single or multiple runs of constant-density batch and semibatch reaction systems, $\mathbf{1}_K \in \mathcal{S}_c(\mathbf{C})$, since $\mathbf{1}_K \in \mathcal{S}_c(\mathbf{X}_e)$.

Proposition 4.24 *Let Assumptions A1–7 in Appendix D be verified. For single or multiple runs of constant-density batch and semibatch reaction systems, if* $\operatorname{rank}(\mathbf{C}) = R+p+1$*, then* $\mathbf{1}_K \in \mathcal{S}_c(\mathbf{X}_e)$*. A similar result holds for constant-density batch reaction systems. (See Appendix H.3.5 for proof)*

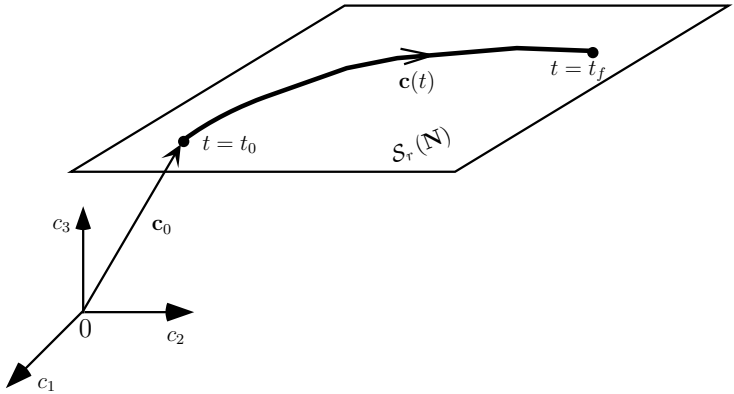

Figure 4.13. Geometric interpretation of the factorization of concentration data from a constant-density batch reaction system. The concentrations $\mathbf{c}$ evolve in an R-dimensional hyperplane spanned by the rows of $\mathbf{N}$ ($\mathcal{S}_r(\mathbf{N})$).

For a single constant-density batch run, it is seen from (4.54) that the $\mathbf{1}_K$-vector is always contained in the column space of $\mathbf{X}_e$. Thus, a rank drop of one occurs when column-mean centering $\mathbf{C}$. For multiple constant-density batch runs, the rank drops by column-mean centering only if $\mathbf{X}_e$ and $\mathbf{N}_e$ are of full rank. Note that this also includes the case of $\mathbf{C}$ being of rank S when $\text{rank}(\mathbf{N}_e) = S = R + p + 1$. For constant-density semibatch reaction systems, rank drop by one occurs both for single and multiple runs if $\text{rank}(\mathbf{C}) = R + p + 1$.

For constant-density batch reaction system, Figure 4.13 illustrates that the concentrations $\mathbf{c}$ evolve in an R-dimensional hyperplane spanned by the rows of $\mathbf{N}$ ($\mathcal{S}_r(\mathbf{N})$). Note that the hyperplane does not include the initial concentrations or the origin, the reason for the rank drop by one upon column mean centering of $\mathbf{C}$. This can also be observed for constant-density semibatch reaction systems

Note that the rank of concentration data from non-reacting mixtures with closure (see Section 4.5.4 below) also drops by one when column-mean centered, since such data is equivalent to data from constant-density batch reaction systems.

A rank drop can provide valuable structural information: Assuming $\text{rank}(\mathbf{C}) = R + p + 1$, a rank drop by one upon mean centering $\mathbf{C}$ indicates that (i) there is no outlet stream and (ii) the density remains constant.

Furthermore, for constant-density batch reaction systems,

$$\bar{\mathbf{C}}_c = \mathbf{J}_c\,(\mathbf{1}_K\,\mathbf{c}_0^{\mathrm{T}} + \mathbf{D}) = \mathbf{D}/V_0 + \mathbf{1}_K\mathbf{c}_0^{\mathrm{T}} = \bar{\mathbf{D}}_c \tag{4.56}$$

$$\bar{\mathbf{H}}_c = \bar{\mathbf{D}}_c\,\mathbf{E} = \bar{\mathbf{C}}_c\,\mathbf{E} = \bar{\mathbf{A}}_c. \tag{4.57}$$

Similarly, since $\dot{\mathbf{X}}_e = [\dot{\mathbf{X}},\, \mathbf{0}_K]$, it follows that

$$\dot{\mathbf{A}} = \dot{\mathbf{H}}; \qquad \dot{\mathbf{C}} = \dot{\mathbf{D}} \tag{4.58}$$

It follows from (4.57) and (4.58) that

$$\operatorname{rank}\left(\bar{\mathbf{C}}_c\right) = \operatorname{rank}\left(\bar{\mathbf{D}}_c\right) = \operatorname{rank}\left(\dot{\mathbf{C}}\right) = \operatorname{rank}\left(\dot{\mathbf{D}}\right). \tag{4.59}$$

Target factor analysis

For batch reaction systems, the two assumptions of Theorem 4.18 are implicitly verified. $\operatorname{rank}\left(\mathbf{c}_0^{\mathrm{T}}\boldsymbol{\Gamma}\right) = 1$ as can be seen from Proposition 2.7c. For the proof of $\operatorname{rank}\left(\mathbf{X}_e\right) = \operatorname{rank}\left([\mathbf{X},\, \mathbf{1}_K]\right) = R + 1$, see Proposition 4.5. Also, only one column of $\boldsymbol{\Gamma}$ is required.

4.5.3 Systems with reactions in quasi-equilibrium

The factorization of concentration data from systems with reactions in quasi-equilibrium conserve the structures of (4.4) and (4.11) for which $\mathbf{x}$, $\mathbf{z}$, and λ in (4.7) are simply reparameterized in terms of s:

$$\frac{d\mathbf{x}}{ds} = \mathbf{r}(\mathbf{c})\, V/(\lambda\, \dot{s}), \quad \frac{d\mathbf{z}}{ds} = \mathbf{q}_{in}/(\lambda\, \dot{s}), \quad \frac{d\lambda}{ds} = -\frac{q_{out}}{V\, \dot{s}}\, \lambda, \tag{4.60}$$

where the initial conditions for $\mathbf{x}$, $\mathbf{z}$, and λ remain the same as in (4.7).

4.5.4 Non-reacting data with closure

Non-reacting mixtures with closure (i.e., with one or several constraints acting on the concentrations) can be cast in the form of the factorization of concentration data for reacting mixtures. Consider the situation where a weighted sum of the concentrations remains constant:

$$\mathbf{C}\,\boldsymbol{\beta} = \alpha \mathbf{1}_K, \tag{4.61}$$

where α is a scalar and $\boldsymbol{\beta}$ a vector with elements $\beta_1, \beta_2, \ldots, \beta_S$. Without loss of generality, it is assumed that $\beta_1 = 1$. Then, the concentration matrix (4.61) can be written as

$$\begin{aligned} \mathbf{C} &= \mathbf{C}^\star \left[-\boldsymbol{\beta}^\star \quad \mathbf{I}_{S-1}\right] + \mathbf{1}_K \left[\alpha \quad \mathbf{0}_{S-1}^{\mathrm{T}}\right] \\ &= \left[\begin{array}{c|c} \mathbf{C}^\star & \mathbf{1}_K \end{array}\right] \left[\begin{array}{cc} -\boldsymbol{\beta}^\star & \mathbf{I}_{S-1} \\ \hline \alpha & \mathbf{0}_{S-1}^{\mathrm{T}} \end{array}\right] = \mathbf{X}_e\, \mathbf{N}_e, \end{aligned} \tag{4.62}$$

where $\boldsymbol{\beta}^\star$ is $\boldsymbol{\beta}$ without the first element, and $\mathbf{C}^\star$ a $K \times S-1$ submatrix of $\mathbf{C}$ without the first column.

Non-reacting concentration data with closure (4.62) can be considered as a special case of data from constant-density batch reaction systems (4.54) with

$$\mathbf{X} := V_0 \mathbf{C}^\star, \quad \mathbf{N} := \left[-\boldsymbol{\beta}^\star \quad \mathbf{I}_{S-1}\right], \quad \mathbf{c}_0^{\mathrm{T}} := \left[\alpha \quad \mathbf{0}_{S-1}^{\mathrm{T}}\right], \tag{4.63}$$

$R := (S-1)$, and unity volume ($V_0 = 1$).

The RV-concentration matrix becomes: $\mathbf{D} = \mathbf{X}\,\mathbf{N} = V_0 \mathbf{C}^\star\, [-\boldsymbol{\beta}^\star,\, \mathbf{I}_{S-1}]$.

Thus, for the closure problem, an equivalent reaction pathway is:

$$\beta_i\, \mathcal{X}_i \rightarrow \mathcal{X}_1, \quad \forall i = 2, \ldots, S.$$

4.6 Summary

Based on the transformation to normal form, a factorization of concentration data was derived that explicits the reaction and flow variants/invariants (the factorization of concentration data). Data pre-treatment to reaction-variant form was proposed, which subtracts the reaction-invariant part from the measured concentrations (the factorization of concentration data in reaction-variant form). This way, information related to the reaction dynamics can be separated from stoichiometric information. It will be shown in the next chapter that the factorization of concentration data play an important role in the analysis of spectral measurements.

Based on the factorization of concentration data, it was possible, without knowledge of reaction kinetics, (i) to reconstruct on-line the concentrations of all species from the concentrations of a few measured species, thereby providing new results on the asymptotic behavior of the estimates, and (ii) to identify experimentally the stoichiometry of a complex reaction system using TFA.

The results obtained for constant-density, isothermal reaction systems (basic dynamic model) were extended to varying-density and non-isothermal situations, to reaction systems expressed in weight fractions, and to process runs with different stoichiometries. Furthermore, the factorization of concentration data were extended to include temperature or calorimetric measurements, as it was shown that these measurements can provide valuable additional information when the concentrations of not all species are measured.

Special cases such as CSTRs, semibatch and batch reaction systems, systems with reactions in quasi-equilibrium conditions, and non-reacting data with closure were also considered.

5

Reaction and flow variants/invariants in the factorization of spectral data

In the previous chapter, factorization of concentration data from reaction systems was proposed. In current practice, however, analytical concentration measurements are not generally available during the course of a reaction. If concentrations are measured, they are mostly used to provide postproduction quality control assessment and demonstrate waste discharge compliance (American Chemical Society *et al.*, 1996).

In rare cases, the concentrations of a limited number of species are analyzed for a few samples taken *during* the course of a reaction. There are several reasons for this: (i) A species concentration is measured mostly through wet-chemical analyses. Owing to high toxicity/reactivity and/or extreme reaction conditions, sampling and sample preparation are often difficult or even impossible; (ii) the development of analytical methods (including validation) is often time consuming and costly; (iii) regarding process monitoring, control, or optimization, another disadvantage of direct concentration measurements is that the time delay between the sampling and the analysis result is usually large (e.g., about 10min for on-line gas chromatographs); (iv) since usually only one process analyzer is available, the long chemical analysis time limits the number of samples that can be taken during a reaction.

Owing to new measurement technologies, spectral data (a one-dimensional signal per observation) are now available both in the laboratory and production (see Appendix F.1.1 for a non-exhaustive list of spectral instrument types). Numerous spectral instruments enable non-destructive indirect concentration measurement of most of the species *in-situ*/on-line during the course of reaction. Measurements are available at short sampling times (often less than a second) at low costs and delay free.

However, similarly to traditional wet-chemical analysis methods, a model must be estimated that provides concentration estimates. Numerous types of spectral data are linear, i.e., the mixture spectrum is a linear combination of the pure-component spectra weighted by the concentrations (linear model; see Appendix F.1.1). Thus, *factorization of spectral data* from reacting mixtures (*reacting spectral data*) can be derived based on the factorization of concentration data. Consequently, *linear* statistical methods are also applicable to spectral data. The two methods that are studied in this chapter are: multivariate calibration and curve-resolution (factor analysis; FA) for the *quantitative*

estimation of concentrations from reacting spectral data.

Factorization of spectral data will be studied for chemical reaction systems under the usual assumptions of constant density and temperature. Furthermore, it is assumed that all species both react and absorb (have a non-zero spectrum) and the corresponding pure-component spectra are linearly independent. Since the factorization of concentration data of the extensions in Subsection 4.4 have the same structure as that of the basic factorization of concentration data, the approaches derived herein can readily be extended to the cases of varying density and/or temperature (under the assumption that the pure-component spectra are temperature independent). Furthermore, the following extensions are studied: (i) presence of non-absorbing reacting species, (ii) presence of non-reacting absorbing species, (iii) linearly-dependent pure-component spectra, (iv) spectral data from process runs with different stoichiometries, and (v) the factorization of spectral data extended by temperature or calorimetric measurements.

The chapter is organized as follows: Section 5.1 introduces the factorization of spectral data (the factorization of spectral data and the factorization in reaction-variant form) for measurements from reacting mixtures. Section 5.2 investigates the ranks of the spectral matrices and methods to reduce their nullity. Section 5.3 studies calibration and factor analysis. Extensions to the basic factorization are provided in Section 5.4.

5.1 Factorization of spectral data

The factorization of general (linear) spectral measurements is briefly reviewed next. Then, the factorization of spectral data from reacting systems are derived.

5.1.1 Factorization of general spectral data

It is assumed that (i) spectral data are available from instruments that measure absorbances and depend *linearly* on molar concentrations (Assumption A8 in Appendix D) and (ii) all species both react and absorb (Assumption A12). The following definition will be useful herein.

Definition 5.1 (Number of absorbing species) *Let S be the number of absorbing species that cause a linearly-independent instrumental response.*

Let $\mathbf{a}(k)$ denote the spectral (absorbance) vector of an L-channel instrument at the observation instant k. For unit pathlength and Beer's law being valid,

$$\mathbf{a}^{\mathrm{T}}(k) = \mathbf{c}^{\mathrm{T}}(k)\,\mathbf{E}, \tag{5.1}$$

where $\mathbf{E}$ is the $S \times L$ pure-component spectra matrix. For K observations, (5.1) can be written in matrix form as:

$$\mathbf{A} = \mathbf{C}\,\mathbf{E} \tag{5.2}$$

with $\mathbf{A}$ being the $K \times L$ spectral (absorbance) matrix.

(5.2) is the factorization of spectral absorbances. However, this factorization is equally valid for other types of spectral data provided they obey the linear relationship (*factorization of general spectral data*; see Appendix F for a list of such instruments). Furthermore, depending on the type of instrument and data pre-treatment used (see Appendix F.1.2), $\mathbf{E}$ can represent (i) the pure-component spectra as a function of wavelengths, wavenumbers, or number of integrated peaks, or (ii) the first or second derivative of the pure-component spectra.

5.1.2 Factorization of spectral data from reacting systems

Factorization of spectral data from reacting systems are presented under the assumption that all species both react and absorb.

Substituting (4.4) into (5.2) leads to *the factorization of spectral data*:

$$\mathbf{A} = \mathbf{C}\,\mathbf{E} = \mathbf{X}_e\,\mathbf{N}_e\,\mathbf{E} = \mathcal{H}^{-1}\,\mathbf{X}\,\mathbf{N}\,\mathbf{E} + \mathcal{H}^{-1}\,\mathbf{Z}\,\mathbf{A}_{in} + V_0\,\mathcal{H}^{-1}\,\mathbf{1}_K\,\mathbf{a}_0^{\mathrm{T}} \tag{5.3}$$

with $\mathbf{A}_{in} \equiv \mathbf{C}_{in}^{\mathrm{T}}\mathbf{E}$ denoting the spectra of the inlet streams (*inlet spectra*) and $\mathbf{a}_0^{\mathrm{T}} = \mathbf{c}_0^{\mathrm{T}}\,\mathbf{E}$ the initial spectrum. The $(p+1) \times L$ *experimental spectral matrix* $\mathbf{A}_x$ is defined using $\mathbf{C}_x$:

$$\mathbf{A}_x = \begin{bmatrix} \mathbf{A}_{in} \\ \mathbf{a}_0^{\mathrm{T}} \end{bmatrix} = \mathbf{C}_x\mathbf{E}. \tag{5.4}$$

Similar to the definition of $\mathbf{D}$ (see 4.11), let the $K \times L$ *spectral matrix in reaction-variant form (RV-spectral matrix)* $\mathbf{H}$ be defined as:

$$\mathbf{H} \equiv \mathcal{H}\,\mathbf{A} - \mathbf{Z}\,\mathbf{A}_{in} - V_0\,\mathbf{1}_K\,\mathbf{a}_0^{\mathrm{T}} = \mathcal{H}\,\mathbf{A} - \mathbf{Z}_x\mathbf{A}_x = \mathbf{D}\,\mathbf{E} = \mathbf{X}\,\mathbf{N}\,\mathbf{E}. \tag{5.5}$$

(5.5) is termed the *factorization of spectral data in reaction-variant form (factorization of RV-spectral data)*.

V_0, $\mathbf{q}_{in}(k)$, and $q_{out}(k)$ for all k need to be known for the computation of $\mathbf{Z}$ and $\mathcal{H}$ in (5.5). Furthermore, the initial and inlet spectra $\mathbf{a}_0$ and $\mathbf{A}_{in}$ or, equivalently, $\mathbf{A}_x$ need to be known. Note that the computation of $\mathbf{H}$ does not involve the concentration matrix $\mathbf{C}$, which is a clear advantage over the alternate data pre-treatment in Appendix E.2. Thus, (4.11) represents a *generic* data pre-treatment for spectral measurements.

Remark 5.2 (Initial concentrations)
It is assumed that the reaction is conducted isothermally. In practical applications, however, there is often a heating phase before the first measurement $\mathbf{a}_0$ at time t_0. Thus, especially for fast reactions, the concentrations of the species initially placed in the reactor, $\mathbf{c}_0^\star$, can be significantly different from $\mathbf{c}_0$ at t_0. Thus, in such a case, $\mathbf{c}_0$ must be measured or estimated.

The contribution of the inlet and initial spectra are subtracted so that the spectral data contain only the influence of the reaction variants. However, if $\mathbf{Z}\,\mathbf{A}_{in}$ cannot be computed, then one may want to choose another data pre-treatment, as given in the next remark.

Remark 5.3 (Subtracting only the initial concentrations and spectrum)
The modified RV-concentration $\mathbf{D}^\star$ ($K \times S$) and RV-spectral matrices $\mathbf{H}^\star$ ($K \times L$) are defined as:

$$\begin{aligned} \mathbf{D}^\star &\equiv \mathcal{H}\mathbf{C} - V_0\,\mathbf{1}_K\,\mathbf{c}_0^{\mathrm{T}} = \left[\begin{array}{c|c} \mathbf{X} & \mathbf{Z} \end{array}\right] \left[\begin{array}{c} \mathbf{N} \\ \hline \mathbf{C}_{in}^{\mathrm{T}} \end{array}\right] = \mathbf{X}_e^\star\,\mathbf{N}_e^\star \\ \mathbf{H}^\star &\equiv \mathcal{H}\mathbf{A} - V_0\,\mathbf{1}_K\,\mathbf{a}_0^{\mathrm{T}} = \mathbf{D}^\star\,\mathbf{E}, \end{aligned} \tag{5.6}$$

where $\mathbf{X}_e^\star$ and $\mathbf{N}_e^\star$ are matrices of dimension $K \times (R+p)$ and $(R+p) \times S$, respectively. For the computation of $\mathbf{D}^\star$, the knowledge of $\mathbf{C}$, V_0, $\mathbf{q}_{in}(k)$, and $q_{out}(k)$ for all k is required, and for $\mathbf{H}^\star$, that of $\mathbf{A}$, V_0, $\mathbf{q}_{in}(k)$, and $q_{out}(k)$ for all k.

5.2 Rank analysis of spectral matrices

5.2.1 Rank of selected matrices

The ranks of $\mathbf{A}$ and $\mathbf{H}$ are investigated next. According to the definition in Property B.3, the matrices $\mathbf{A}$ and $\mathbf{H}$ are rank deficient if their ranks are less than L. However, since the spectral matrices $\mathbf{A}$ and $\mathbf{H}$ are linked to $\mathbf{C}$ and $\mathbf{D}$ through $\mathbf{E}$, respectively, rank deficiency of $\mathbf{A}$ and $\mathbf{H}$ is defined as follows (and not directly from Property B.3).

Definition 5.4 (Rank deficiency of A and H and full rank of A) *The $K \times L$ matrices $\mathbf{A}$ and $\mathbf{H}$ are said to be rank deficient, if their ranks are less than the number of absorbing species S, i.e.,* $\mathrm{rank}\,(\mathbf{A}) < S$ *and* $\mathrm{rank}\,(\mathbf{H}) < S$*, respectively. The rank of $\mathbf{A}$ is said to be full, if* $\mathrm{rank}\,(\mathbf{A}) = S$.

In many applications to spectral data, $\mathbf{A}$ is rank deficient, and data pre-treatment such as column-mean centering affects the rank. The rank of spectral data matrices from constant-density batch reaction systems has been investigated by Amrhein *et al.* (1996). Rank deficiency of a particular constant-density batch system with fast equilibrium reactions was also observed by Nørgaard and Ridder (1994). Some indications about the rank of spectral data matrices from constant-density reaction systems with inlet streams have been provided by Garland *et al.* (1997). Here, the ranks of $\mathbf{A}$ and $\mathbf{H}$ from varying-density reaction systems with inlet and outlets are studied.

Proposition 5.5 *Let Assumptions A8–11 in Appendix D be verified. Then,*

$$\begin{aligned} \mathrm{rank}\,(\mathbf{A}) &= \mathrm{rank}\,(\mathbf{C}), \qquad \mathrm{rank}\,(\mathbf{H}) = \mathrm{rank}\,(\mathbf{D}) = R \\ \mathrm{rank}\,(\bar{\mathbf{A}}_c) &= \mathrm{rank}\,(\bar{\mathbf{C}}_c), \qquad \mathrm{rank}\,(\bar{\mathbf{H}}_c) = \mathrm{rank}\,(\bar{\mathbf{D}}_c) \end{aligned} \tag{5.7}$$

(See Appendix H.4.1 for proof)

As far as the ranks of the spectral matrices $\mathbf{A}$ and $\mathbf{H}$ (or $\bar{\mathbf{A}}_c$ and $\bar{\mathbf{H}}_c$) are concerned, if the pure-component spectra of the S absorbing species are independent ($\mathrm{rank}\,(\mathbf{E}) = S$), then the ranks are determined by the ranks of $\mathbf{C}$ and $\mathbf{D}$ (or $\bar{\mathbf{C}}_c$ and $\bar{\mathbf{D}}_c$; see Section 4.2.1), and it suffices to study the ranks of $\mathbf{C}$ and $\mathbf{D}$ (or $\bar{\mathbf{C}}_c$ and $\bar{\mathbf{D}}_c$) to comment on the ranks of $\mathbf{A}$ and $\mathbf{H}$ (or $\bar{\mathbf{A}}_c$ and $\bar{\mathbf{H}}_c$), respectively.

5.2.2 Nullity-reducing operations

In calibration and factor analysis of spectral data (see Subsections 5.3.1 and 5.3.2 below), maximizing the rank or, equivalently, minimizing its nullity will become important.

The experimental rank-increasing operation proposed in Subsection 4.2.2.2 can also be applied to spectral data. In addition, an alternative rank-increasing operation is proposed that is essentially non-experimental. It appends pure-component spectra to $\mathbf{A}$ in an attempt to increase its rank. For both methods, $\operatorname{rank}(\mathbf{E}) = S$ is assumed. Alternatively, Subsection 5.2.2.3 proposes channel selection for nullity reduction.

5.2.2.1 Experimental rank-increasing operations

In Subsection 4.2.2.2, an experimental rank-increasing operation for the concentration matrix $\mathbf{C}$ was proposed. This method is directly applicable to spectral data.

It is important to note that since the rank of $\mathbf{C}$ can be computed from the rank of $\mathbf{A}$, the *values* of the elements of the experimental concentration matrix $\mathbf{C}_x$ are not required to be explicitly known for the rank-increasing operation. Thus, the following two situations can be handled:
(i) Although the concentrations $\mathbf{c}_0^\star$ of the species initially placed in the reactor are usually known, the initial spectrum $\mathbf{a}_0$ may be taken at a later time instant and, thus, does not correspond to $\mathbf{c}_0^\star$ (see Remark 5.2). In such a case, the values of the initial concentrations $\mathbf{c}_0$ are unknown and so the corresponding entries in $\mathbf{C}_x$.

(ii) The mixtures for the experimental rank-increasing operation are usually prepared by putting together species of known amounts, and experimental design guarantees linear independence of these mixtures ($\operatorname{rank}(\mathbf{C}_x) = p + 1$). An alternative possibility of generating such mixtures is to take samples from the running reaction and pass them through a separation process (e.g., a distillation column) in an attempt to break the linear dependencies in $\mathbf{C}$ resulting from the underlying reactions. For this, the separation does not need to be complete and, thus, the concentration values of the corresponding mixture may be unknown.

5.2.2.2 Non-experimental rank-increasing method

An alternative way of rank increasing is proposed. Assuming the number of absorbing species S to be known, the key idea is as follows: When absorbing species are known to be present in the mixture, their pure-component spectra (e.g., available from a database) can be appended to $\mathbf{A}$ in an attempt to increase the rank of the augmented spectral matrix. The procedure is the following:

Let S_p be the number of species known *a priori* to be present in the mixture for which the pure-component spectra are available. Without loss of generality, it can be assumed that the last S_p rows of $\mathbf{E}$ are the pure-component spectra of the S_p species:

$$\mathbf{E}_p = \mathbf{F}_p \mathbf{E}, \quad \mathbf{F}_p = \left[\begin{array}{c|c} \mathbf{0}_{S_p \times (S - S_p)} & \mathbf{I}_{S_p} \end{array} \right]. \tag{5.8}$$

The spectral matrix can be augmented to an $(K+S_p)\times L$ matrix, $\mathbf{A}^{aug}$, by appending $\mathbf{E}_p$ to $\mathbf{A}$:

$$\mathbf{A}^{aug} \equiv \left[\frac{\mathbf{A}}{\mathbf{E}_p}\right] \tag{5.9}$$

By substituting (5.3) into (5.9),

$$\begin{aligned}
\mathbf{A}^{aug} &= \mathbf{C}^{aug}\,\mathbf{E} = \mathbf{X}_e^{aug}\,\mathbf{N}_e^{aug}\,\mathbf{E}, \\
\mathbf{C}^{aug} &\equiv \left[\frac{\mathbf{C}}{\mathbf{F}_p}\right], \quad \mathbf{X}_e^{aug} \equiv \left[\begin{array}{c|c} \mathbf{X}_e & \mathbf{0}_{K\times S_p} \\ \hline \mathbf{0}_{S_p\times(R+p+1)} & \mathbf{I}_{S_p} \end{array}\right], \quad \mathbf{N}_e^{aug} \equiv \left[\frac{\mathbf{N}_e}{\mathbf{F}_p}\right],
\end{aligned}$$

where $\mathbf{C}^{aug}$, $\mathbf{X}_e^{aug}$, and $\mathbf{N}_e^{aug}$ are matrices of dimension $(K+S_p)\times S$, $(K+S_p)\times(R+p+1+S_p)$, and $(R+p+1+S_p)\times S$, respectively.

Corollary 5.6 *Let Assumptions A1–12 in Appendix D be verified. Proposition 4.11 holds with the following redefinitions:* $\mathbf{C}_- := \mathbf{A}$, $\mathbf{C}_+ := \mathbf{A}^{aug}$, $\mathbf{X}_{e,-} := \mathbf{X}_e$, $\mathbf{X}_{e,+} := \mathbf{X}_e^{aug}$, $\mathbf{N}_{e,-} := \mathbf{N}_e$, $\mathbf{N}_{e,+} := \mathbf{N}_e^{aug}$, $R+p_-+1 := R+p+1$, $R+p_++1 := R+p+1+S_p$, *and* $(p_- = S-R-1) := (p_- = S-R-1-S_p)$.

The condition to guarantee full rank of $\mathbf{X}_e^{aug}$ is given in the corollary below.

Corollary 5.7 (Full rank of $\mathbf{X}_e^{aug}$) *Let Assumptions A1–7 in Appendix D be verified. If* $\operatorname{rank}(\mathbf{X}_e) = (R+p+1)$, *then* $\operatorname{rank}(\mathbf{X}_e^{aug}) = (R+p+1+S_p)$ *is trivially satisfied.(See Appendix H.4.1 for proof)*

Corollary 5.7 shows that, as far as the rank of $\mathbf{X}_e^{aug}$ is concerned, appending a pure-component spectrum is equivalent to appending a batch run. Compared to the experimental rank-increasing operations, fewer inlets are required to reach $\operatorname{rank}(\mathbf{A}^{aug}) = S$ since the condition becomes $p \geq S-R-1-S_p$.

Instead of appending pure-component spectra to $\mathbf{A}$, it is also possible to append mixture spectra with known concentrations (of species known to be present in the unknown mixture) to $\mathbf{A}$. In such a case, $\mathbf{F}_p$ needs to be defined appropriately.

The main advantage of the proposed procedure is that it requires less experimental effort for rank increase of the spectral matrix. When S_p pure-component spectra are available only for *some* of the S_x species with $S_p < S_x$, a *hybrid method* for rank increase can be envisaged: (i) the rank of the spectral matrix is partially increased by appending $\mathbf{E}_p$ to $\mathbf{A}$ such that $\operatorname{rank}(\mathbf{A}^{aug}) > \operatorname{rank}(\mathbf{A})$. (ii) An experimental rank-increasing procedure uses the remaining $(S_x - S_p)$ species for independent variation in an attempt to increase the rank of $\mathbf{A}$ to its maximal attainable rank $\min(R+S_x, S)$.

5.2.2.3 Channel selection for reducing the number of observed absorbing species

First, some notations of Chapter 4 are redefined. Let S be the total number of absorbing species, S_r the total number of reacting species, and S_m and R_o the numbers

of absorbing species and independent reactions observed in the selected spectral data, respectively.

Channel selection methods can be used to select only spectral regions where a few of the S_r reacting species absorb ($S < S_r$). It is assumed that before channel selection, all species both absorb and react (Assumption A12 in Appendix D). After channel selection this assumption is relaxed to the case where some of the reacting species does not absorb. However, it is assumed that all the absorbing species do react.

Since channel selection is equivalent to measuring the concentrations of fewer (reacting) species (see Subsection 4.2.2.2), the number of the independent reactions R_o observed in the selected spectral data may be less than the total number of independent reactions R. The following result with respect to channel selection is a corollary to Proposition 4.12.

Corollary 5.8 *Let Assumptions A1–11 in Appendix D be verified, and the subscripts $-$ and $+$ denote a quantity before and after channel selection, respectively. If (i) all species both absorb and react before channel selection ($S_- = S$), (ii)* $\operatorname{rank}(\mathbf{A}_-) = R + p + 1$, *and (iii)* $\operatorname{rank}(\mathbf{E}_+) = S_+$, *then channel selection ($S_+ = S_m < S_-$) leads to the nullity reduction Δd of the concentration matrices $\mathbf{C}_-$ and $\mathbf{C}_+$ corresponding to the spectral matrices $\mathbf{A}_- = \mathbf{C}_-\mathbf{E}_-$ and $\mathbf{A}_+ = \mathbf{C}_+\mathbf{E}_+$, respectively:*

$$\max(S_m - S + R - R_o, R + p + 1 - S) \leq \Delta d \leq S_m - S + R - R_o + p + 1. \quad (5.10)$$

Channel selection can reduce but never increase the nullity of the concentration matrices corresponding to the spectral matrices $\mathbf{A}_-$ and $\mathbf{A}_+$. To attain rank S_+, the channels must be so chosen that $\operatorname{rank}(\mathbf{A}_+) = S_+$. The same results also hold for $\mathbf{H}_-$ and $\mathbf{H}_+$ with the redefinition $(p + 1) := 0$.

5.3 Some implications

5.3.1 Multivariate calibration

Multivariate forward calibration allows predicting the concentrations of the *species of interest* from *new spectra* (see Appendix C.2 for a brief review of calibration in general and Appendix F.2 for the reasons why multivariate forward calibration was chosen). For the calibration of the species of interest, the spectra and the corresponding concentrations must be available for a few observations (e.g., from non-reacting mixtures of known concentrations or wet-chemical analyses of reacting mixtures).

Calibration methods require that a new spectrum lies in the space spanned by the calibration spectral data (*space-inclusion condition*) and the pure-component spectra of the absorbing species are linearly independent. The space-inclusion condition represents an *a posteriori* condition for the applicability of a calibration model, since it requires the generation of a spectrum from a new experimental run. However, it would be desirable to know *a priori* whether the space-inclusion condition will be fulfilled, i.e., *before* the generation of a spectrum from a new experimental run. In practical applications, this goal is usually achieved by guaranteeing the following conditions for the calibration data: (i) the absorbing species present in calibration set should include

those in the new set, and (ii) the rank of the calibration spectral data equals the number of absorbing species. For *non-reacting (static) calibration data*, Condition ii can often be met, assuming that all the absorbing species are available for linearly-independent composition variation. Numerous mixtures with concentrations selected according to an experimental plan are usually prepared (Box *et al.*, 1978). However, also possibly highly-reactive intermediates must be available — a requirement that is often difficult and sometimes impossible to meet.

This problem is circumvented by taking calibration samples during the course of the reaction (*reacting calibration data*). This way, the intermediates are produced by the reactions, and the need to vary them externally disappears. This advantage over non-reacting calibration data may justify the additional costs of sampling, sample preparation, and wet-chemical analyses. For the verification of Condition ii in case of reacting calibration data, it is proposed to use nullity-reducing operations for the calibration data (rank-increasing operations and/or channel selection; see Subsection 5.2.2) in an attempt to reduce to zero the nullity of the calibration concentration matrix.

However, there are many practical situations where the nullity-reducing operations cannot be applied, e.g., when the required species are not all available for independent variation, or when the spectral data exhibit highly-overlapped bands. For such cases, Condition ii cannot be guaranteed, and an alternative method for the verification of the space-inclusion condition must be found. For this purpose, a method based on the pre-treatment of both the calibration and the new data to reaction-variant form is proposed. This way, the resulting calibration model only predicts that part of the concentrations corresponding to the new spectrum which is usually unknown due to unknown kinetics, namely the reaction-variant part. To obtain the concentrations, the typically known reaction-invariant part (such as the initial conditions and the effects of the inlet and outlet streams) must be added. Both the nullity-reducing operations and the data pre-treatment to reaction-variant form enable free specification of those terms of the new runs related to the reaction invariants.

This subsection is organized as follows: Subsection 5.3.1.1 presents calibration for general rank-deficient concentration and spectral data in the cases of calibration data with and without interferents. For reacting mixtures, Subsection 5.3.1.2 proposes conditions to check the space-inclusion condition *before* the generation of a new spectrum. Subsection 5.3.1.3 focuses on methods that enable the choice of the experimental concentration matrix for the new set independently of the calibration set. Subsection 5.3.1.4 summarizes the space-inclusion results, and Subsection 5.3.1.5 illustrates them by a simulated example.

5.3.1.1 Prediction in the case of rank-deficient spectral data

For simplicity of notation, the results that follow are stated for the noise-free case.

Let S denote the number of the union of absorbing species present in both the calibration *and* the new data, S_k the number of those absorbing species for which calibration concentrations are available by a reference method (called *known species* or *species of interest*), and S_u the number of the remaining absorbing species (called *interferents* or *unknown species*), with $S = S_k + S_u$.

The prediction of concentrations from a new spectrum $\mathbf{a}_n$ will be addressed using a calibration model that is based on calibration data with either (i) no interferent or (ii) S_u interferents. The subscript n indicates a quantity related to a new spectrum.

Concentrations of all species available for calibration

For the case that the concentrations of all (absorbing) species are measured (i.e., $S_k = S$; no interferent in the calibration data), the forward calibration model with the spectral data $\mathbf{A}$ being the inputs and the concentration data $\mathbf{C}$ the outputs is:

$$\mathbf{C} = \mathbf{A}\,\mathbf{B}, \tag{5.11}$$

where $\mathbf{B}$ is an $L \times S$ regressor matrix.

Proposition 5.9 *Let* $\operatorname{rank}(\mathbf{E}) = S$. *Then, the concentrations of the* S *species are predicted correctly from a new spectrum,* $\mathbf{a}_n$, *using forward calibration:*

$$\hat{\mathbf{c}}_n^{\mathrm{T}} = \mathbf{a}_n^{\mathrm{T}}\hat{\mathbf{B}}, \quad \textit{where} \quad \hat{\mathbf{B}} = \mathbf{A}^{+}\,\mathbf{C}, \tag{5.12}$$

iff $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$. *(See Appendix H.4.2 for proof)*

Proposition 5.9 says that when (rank-deficient) spectral and concentration data follow (5.11), the concentrations $\hat{\mathbf{c}}_n$ are predicted correctly from a new spectrum $\mathbf{a}_n$ under the following two assumptions: (i) $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$, i.e., the new spectrum $\mathbf{a}_n$ lies in the row space of the calibration spectral data $\mathbf{A}$ (*space-inclusion condition*), and (ii) $\operatorname{rank}(\mathbf{E}) = S$.

The space-inclusion condition $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$ can be checked by calculating $\varepsilon_p(\mathbf{a}_n, \mathbf{A})$, the projection error of $\mathbf{a}_n$ on the row space of $\mathbf{A}$ using (C.15).

Proposition 5.9 is valid independently of the regression method used. The difference between the various regression methods such as principal component regression (PCR; Jackson, 1991), partial least squares (PLS; Wold, 1966; Wold and Lyttkens, 1970), a modification of PLS (SIMPLS; de Jong, 1993; Phatak and de Jong, 1997), and continuum regression (CR; Stone and Brooks, 1990; Phatak *et al.*, 1992) with respect to calibration comes from the way in which noise in $\mathbf{A}$ and/or $\mathbf{C}$ is handled. Since noise handling is not an issue in this work, it is simply referred to Appendix C.2.2 for a short review that illustrates how the forward calibration problem is solved in the framework of bilinear models using biased regression methods such as PCR and SIMPLS.

Concentrations of some species unavailable for calibration

Let the concentrations of (absorbing) $S_k \leq S$ species be available for calibration. Thus, the concentrations of $S_u = (S - S_k)$ species (interferents) are not for calibration. Then, (5.11) can be partitioned into the known and unknown parts of the concentration matrix:

$$\begin{aligned} \mathbf{C} &= \left[\begin{array}{c|c} \mathbf{C}_k & \mathbf{C}_u \end{array}\right] = \mathbf{A}\left[\begin{array}{c|c} \mathbf{B}_k & \mathbf{B}_u \end{array}\right], \\ \mathbf{C}_k &= \mathbf{C}\,\mathbf{J}_k = \mathbf{A}\,\mathbf{B}_k, \quad \text{where} \quad \mathbf{J}_k \equiv \left[\begin{array}{c} \mathbf{I}_{S_k} \\ \hline \mathbf{0}_{S_u \times S_k} \end{array}\right], \end{aligned} \tag{5.13}$$

$$\mathbf{C}_u \;=\; \mathbf{C}\,\mathbf{J}_u = \mathbf{A}\,\mathbf{B}_u, \quad \text{where} \quad \mathbf{J}_u \equiv \begin{bmatrix} \mathbf{0}_{S_k \times S_u} \\ \hline \mathbf{I}_{S_u} \end{bmatrix}, \tag{5.14}$$

with subscripts k and u denoting the parts of a quantity corresponding to the S_k and S_u species, respectively. Thus, the forward formulation enables the separation of the concentrations corresponding to the known and unknown species into two decoupled submodels, and the following corollary to Proposition 5.9 can be formulated.

Corollary 5.10 *Let* $\operatorname{rank}(\mathbf{E}) = S$. *If* $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$, *then the concentrations of the* S_k *known species are predicted correctly using forward calibration:*

$$\hat{\mathbf{c}}_{k,n}^{\mathrm{T}} = \mathbf{a}_n^{\mathrm{T}} \hat{\mathbf{B}}_k, \quad \textit{where} \quad \hat{\mathbf{B}}_k = \mathbf{A}^{+}\,\mathbf{C}_k. \tag{5.15}$$

Contrary to the case of no interferent, $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$ is only a *sufficient* condition for correct prediction of $\mathbf{c}_{k,n}$ from $\mathbf{a}_n$. Thus, for $\mathbf{a}_n \notin \mathcal{S}_r(\mathbf{A})$, there are cases where $\mathbf{c}_{k,n}$ is correctly predicted. However, $\mathbf{c}_{k,u}$ will then be in error. The necessary conditions are provided in Appendix F.2.

Although the remainder of this subsection assumes no interferent to be present in the calibration data, all the calibration-related results also apply to the case with interferents being present. However, for reasons similar to those in Corollary 5.10, the necessity of the conditions only holds for the case of no interferents.

5.3.1.2 Space-inclusion condition for reacting calibration data

Proposition 5.9 was stated for the general case of rank-deficient data. In this subsection, it is assumed that the rank deficiency is due to underlying reactions in the calibration and new data (see Proposition 5.5), i.e., the calibration concentrations are measured from samples taken during the course of reaction run(s), and the calibration model is used to predict concentrations based on spectra from a new reaction run.

To verify *a priori* the space-inclusion condition $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$, experimental conditions in terms of the experimental matrix and the stoichiometric matrix of both the calibration and the new data will be specified.

It is assumed that (i) the absorbing species in the new set are a subset of those in the calibration set, (ii) $\operatorname{rank}(\mathbf{E}) = S$, and (iii) either $\operatorname{rank}(\mathbf{A}) = S$ or $\operatorname{rank}(\mathbf{A}) = R + p + 1 < S$.

From (4.8),

$$\mathbf{C}_x = \begin{bmatrix} \mathbf{C}_{in}^{\mathrm{T}} \\ \mathbf{c}_0^{\mathrm{T}} \end{bmatrix}, \quad \mathbf{N}_e = \begin{bmatrix} \mathbf{N} \\ \mathbf{C}_x \end{bmatrix}, \qquad \mathbf{C}_{x,n} = \begin{bmatrix} \mathbf{C}_{in,n}^{\mathrm{T}} \\ \mathbf{c}_{0,n}^{\mathrm{T}} \end{bmatrix}, \quad \mathbf{N}_{e,n} = \begin{bmatrix} \mathbf{N}_n \\ \mathbf{C}_{x,n} \end{bmatrix}, \tag{5.16}$$

where $\mathbf{C}_x$ $(p+1 \times S)$ and $\mathbf{C}_{x,n}$ $(p_n+1 \times S)$ are the experimental matrices for the calibration and the new sets, p and p_n the corresponding number of (generalized) inlets, $\mathbf{N}$ $(R \times S)$ and $\mathbf{N}_n$ $(R_n \times S)$ the corresponding stoichiometric matrices, and $\mathbf{N}_e$ $(R+p+1 \times S)$ and $\mathbf{N}_{e,n}$ $(R_n+p_n+1 \times S)$ the corresponding extended stoichiometric matrices. Note that the number of independent reactions and the stoichiometries are allowed to differ between the calibration and the new data.

Proposition 5.11 *If* $\operatorname{rank}(\mathbf{A}) = S$*, then* $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$*. (See Appendix H.4.2 for proof)*

Proposition 5.11 shows that the space-inclusion condition $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$ is trivially satisfied when the calibration spectral matrix $\mathbf{A}$ is full rank.

When the calibration spectral data $\mathbf{A}$ is possibly rank deficient, i.e., $\operatorname{rank}(\mathbf{A}) = (R + p + 1) \leq S$, an additional condition must be specified for the verification of $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$, as shown next.

Proposition 5.12 *Let Assumptions A1–12 in Appendix D be verified,* $\operatorname{rank}(\mathbf{A}) = (R + p + 1)$*, and* $\mathbf{R}$ *an* $(R_n + p_n + 1) \times (R + p + 1)$ *matrix of full rank. If*

$$\mathbf{N}_{e,n} = \mathbf{R}\,\mathbf{N}_e, \tag{5.17}$$

then $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$*. (See Appendix H.4.2 for proof)*

Proposition 5.12 shows that, in the case of rank-deficient calibration spectral data, if the row spaces of the extended stoichiometric matrices, in which the spectral measurements of the calibration and the new sets live, are the same, then one can be sure that the space-inclusion condition is satisfied. The conditions $\mathbf{N}_{e,n} = \mathbf{R}\,\mathbf{N}_e$ can be checked by knowing $\mathbf{N}_{e,n}$ and $\mathbf{N}_e$ and computing the Frobenius norm of $\varepsilon_p(\mathbf{N}_{e,n}, \mathbf{N}_e)$ (see C.16): If $\varepsilon_p = 0$, then $\mathbf{N}_{e,n} = \mathbf{R}\,\mathbf{N}_e$ is satisfied. Note that

$$\mathbf{N}_{e,n} = \left[\begin{array}{c} \mathbf{N}_n \\ \hline \mathbf{C}_{x,n} \end{array}\right] = \mathbf{R}\,\mathbf{N}_e = \left[\begin{array}{c} \mathbf{R}_N \\ \hline \mathbf{R}_C \end{array}\right] \mathbf{N}_e, \text{ which implies } \mathbf{N}_n = \mathbf{R}_N\mathbf{N}_e \text{ and } \mathbf{C}_{x,n} = \mathbf{R}_C\mathbf{N}_e, \tag{5.18}$$

where $\mathbf{R}_N$ and $\mathbf{R}_C$ are matrices of dimension $R_n \times (R+p+1)$ and $(p_n+1) \times (R+p+1)$, respectively. Thus, both the stoichiometries and the initial and inlet concentrations of the new set must lie in the row space of $\mathbf{N}_e$.

Note that Proposition 5.12 is independent of the initial volume $V_{0,n}$. Also, since no assumption on $\mathbf{X}_{e,n}$ was made, the kinetics $\mathbf{r}_n$ and the inlet flowrates $\mathbf{q}_{in,n}(t)$ can be arbitrary and even dependent, i.e., $\operatorname{rank}(\mathbf{X}_{e,n}) < (R_n + p_n + 1)$.

To check the condition $\mathbf{N}_{e,n} = \mathbf{R}\,\mathbf{N}_e$ in Proposition 5.12, both the experimental and the stoichiometric matrices of both the calibration and the new sets must be known. In many practical situations, however, the stoichiometries are unknown, but the reactions are the same for both the calibration and the new sets (i.e., $\mathbf{N}_n = \mathbf{N}$). For such a situation (and its generalized form $\mathbf{N}_n = \mathbf{R}^\dagger\,\mathbf{N}$), the following corollary can be stated.

Corollary 5.13 *Let Assumptions A1–12 in Appendix D be verified,* $\operatorname{rank}(\mathbf{A}) = (R + p + 1)$*,* $\mathbf{N}_n = \mathbf{R}^\dagger\,\mathbf{N}$*, and* $\mathbf{R}^\dagger$ *and* $\mathbf{R}^\ddagger$ *matrices of full rank of dimension* $R_n \times R$ *and* $(p_n + 1) \times (p + 1)$*, respectively. If*

$$\mathbf{C}_{x,n} = \mathbf{R}^\ddagger\,\mathbf{C}_x, \tag{5.19}$$

then $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$*. (See Appendix H.4.2 for proof)*

Since the initial and inlet concentrations of the new set must lie in the row space of $\mathbf{C}_x$, which is smaller than that of $\mathbf{N}_e$, Corollary 5.13 provides a more conservative sufficient condition than Proposition 5.12. However, the condition $\mathbf{C}_{x,n} = \mathbf{R}^\ddagger\,\mathbf{C}_x$ is easier to check than the condition $\mathbf{N}_{e,n} = \mathbf{R}\,\mathbf{N}_e$, since it does not require the knowledge of $\mathbf{N}$.

5.3.1.3 Methods for meeting the space-inclusion condition for reacting calibration data

For meeting the space-inclusion condition $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$, Proposition 5.12 and Corollary 5.13 showed that the experimental concentration matrix of the new set $\mathbf{C}_{x,n}$ must be so chosen that the condition $\mathbf{C}_{x,n} = \mathbf{R}_C \mathbf{N}_e$ and the condition $\mathbf{C}_{x,n} = \mathbf{R}^{\ddagger} \mathbf{C}_x$ are satisfied, respectively. However, it would be desirable to choose $\mathbf{C}_{x,n}$ independently of the calibration set. The next two subsections propose methods to that effect.

Nullity-reducing operations

Proposition 5.11 showed that the space-inclusion condition is trivially satisfied if $\operatorname{rank}(\mathbf{A}) = S$ can be guaranteed, thereby enabling the specification of $\mathbf{C}_{x,n}$ independently of the calibration data.

Nullity-reducing operations and maximal attainable rank The idea is to eliminate the null space of the concentration matrix corresponding to calibration spectral matrix $\mathbf{A} = \mathbf{C}\,\mathbf{E}$ by using the nullity-reducing operations proposed in Subsection 5.2.2 (experimental and non-experimental rank-increasing operations; channel selection for reducing the number of observed absorbing species).

Let S_x be redefined as the total number of species that have nonzero elements in $\mathbf{C}_x$ with $\mathbf{C}_x$ fulfilling (5.19), i.e., the species that can be varied linearly independently in the calibration set include those in the new set. In practical applications, often only $S_x < S$ species are available for independent variation of the initial and inlet concentrations (see also Subsection 5.2.2.1). Thus, the rank of $\mathbf{A}$ can be at most $(R + S_x)$. Even in such a case, guaranteeing $\operatorname{rank}(\mathbf{C}_x) = S_x$ gives a free choice of $\mathbf{C}_{x,n}$ for the S_x species, as shown in the next corollary that follows from Corollary 5.13.

Corollary 5.14 *Let Assumptions A1–12 in Appendix D be verified,* $\mathbf{N}_n = \mathbf{R}^{\dagger}\mathbf{N}$ *and* $\operatorname{rank}(\mathbf{A}) = R + p + 1$. *If* $\operatorname{rank}(\mathbf{C}_x) = S_x (= p + 1)$, *then* $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$.

Reacting vs. non-reacting calibration data In many practical calibrations, non-reacting mixtures of well-defined composition are prepared for which the spectra are measured (non-reacting calibration data). The resulting calibration model is then often used to predict the concentrations from spectra measured in reacting mixtures.

Non-reacting calibration data are a special case of reacting calibration data with $\mathbf{N}_e = \mathbf{C}_x$. According to Proposition 5.12, $\mathbf{N}_{e,n} = \mathbf{R}\,\mathbf{N}_e$ guarantees space inclusion. Thus, not only the initial and inlet concentrations $\mathbf{C}_{x,n}$, but also the $R_n \times S$ stoichiometric matrix $\mathbf{N}_n$ of the new set must lie in the row space of $\mathbf{C}_x$. In other words, possibly highly-reactive absorbing intermediates that are produced by the R_n independent reactions in the new set must be available for linearly-independent composition variation in the calibration set. Unfortunately, the latter assumption is often difficult, and sometimes even impossible, to verify.

In reacting calibration data, these intermediates are implicitly included, thereby increasing the rank of the spectral calibration data by R. Thus, R (possibly) intermediates are not required for linearly-independent variation during the calibration phase.

Data pre-treatment to reaction-variant form

The disadvantage of nullity-reducing operations is that they require the availability of $(S-R)$ or (S_m-R_o) species for linearly-independent variation of the *reaction-invariant part*.

This paragraph proposes a method that is based on the pre-treatment of the calibration and the new data to reaction-variant form. It does not require the availability of $S_x = (S-R)$ or $S_x = (S_m - R_o)$ species. The resulting calibration model only predicts that part of the concentrations that is usually unknown due to unknown kinetics, namely the R-dimensional *reaction-variant part*. To obtain an estimate of concentrations, the (usually known) reaction-invariant part (such as initial conditions and effects of the inlet and outlet streams) of the new data must be added. Propositions 5.9 and 5.12 will be stated in terms of the data pair $\{\mathbf{H}, \mathbf{D}\}$. Since the proofs are similar to those of Propositions 5.9 and 5.12, they are omitted here.

Theorem 5.15
Let Assumptions A1–12 in Appendix D be verified and the following quantities be known/measured: (i) for calibration, $\mathbf{A}$, $\mathbf{A}_{in}$, $\mathbf{C}$, $\mathbf{C}_{in}$, V_0, $\mathbf{q}_{in}(k)$, *and* $q_{out}(k)$ *for all* k, *and (ii) for prediction,* $\mathbf{a}_n$, $\mathbf{a}_{0,n}$, $\mathbf{A}_{in,n}$, $V_{0,n}$, $\mathbf{q}_{in,n}(k)$, *and* $q_{out,n}(k)$ *for all* k. *Then, the RV-concentrations*$\mathbf{d}_n$ *for the* S *species are predicted correctly using forward calibration:*

$$\hat{\mathbf{d}}_n^{\mathrm{T}} = \mathbf{h}_n^{\mathrm{T}} \hat{\mathbf{B}}_d, \quad \textit{where} \quad \hat{\mathbf{B}}_d = \mathbf{H}^{+}\,\mathbf{D}, \tag{5.20}$$

iff $\mathbf{h}_n \in \mathcal{S}_r(\mathbf{H})$, *with* $\hat{\mathbf{B}}_d$ *being the estimate of an* $L \times S$ *regressor matrix, and* $\mathbf{D}$ *and* $\mathbf{H}$ *defined as in* (4.11) *and* (5.5), *respectively.*

If, furthermore, $\mathbf{c}_{0,n}$ *and* $\mathbf{C}_{in,n}$ *are known/measured, then the concentrations of the* S *species can be predicted by:*

$$\hat{\mathbf{c}}_n = \frac{\lambda_n}{V_n}\left(\hat{\mathbf{d}}_n + \mathbf{C}_{in,n}\mathbf{z}_n + V_{0,n}\,\mathbf{c}_{0,n}\right). \tag{5.21}$$

Theorem 5.15 is based on the calibration and new data, both pre-treated to reaction-variant form. If the space-inclusion condition $\mathbf{h}_n \in \mathcal{S}_r(\mathbf{H})$ is verified, the resulting calibration model (5.20) predicts the R-dimensional reaction-variant part of the concentrations. To obtain an estimate of the concentrations $\mathbf{c}_n$, the reaction-invariant part ($\mathbf{c}_{0,n}$ and $\mathbf{C}_{in,n}$) of the new data must be added.

In the presence of interferents, (5.20) and (5.21) become

$$\begin{aligned} \hat{\mathbf{d}}_{n,k}^{\mathrm{T}} &= \mathbf{h}_n^{\mathrm{T}} \hat{\mathbf{B}}_{d,k}, \qquad \hat{\mathbf{B}}_{d,k} = \mathbf{H}^{+}\,\mathbf{D}_k, \\ \hat{\mathbf{c}}_{n,k} &= \frac{\lambda_n}{V_n}\left(\hat{\mathbf{d}}_{n,k} + \mathbf{C}_{in,n,k}\mathbf{z}_n + V_{0,n}\,\mathbf{c}_{0,n,k}\right). \end{aligned} \tag{5.22}$$

Proposition 5.16 *Let Assumptions A1–12 in Appendix D be verified.* $\mathbf{h}_n \in \mathcal{S}_r(\mathbf{H})$ *if* $\mathbf{N}_n = \mathbf{R}^{\dagger}\,\mathbf{N}$.

Proposition 5.16 tells that the space-inclusion condition $\mathbf{h}_n \in \mathcal{S}_r(\mathbf{H})$ is satisfied by guaranteeing that the rows of $\mathbf{N}_n$ lie in the row space of $\mathbf{N}$. Since the latter condition is independent of the reaction invariants, the initial and inlet concentrations of the new set can be freely chosen. Note that $\mathbf{N}_n = \mathbf{R}^{\dagger}\,\mathbf{N}$ is trivially satisfied when the reaction system is the same for both the calibration and the new runs.

Remark 5.17 (Subtracting only the initial concentrations and spectrum) Consider the situation where the inlet concentrations, the inlet spectra, and/or the inlet flowrates of either the calibration or the new set are not available. Then, $\mathbf{A}$ and $\mathbf{C}$ can be pre-treated to $\mathbf{H}^\star$ and $\mathbf{D}^\star$ using (5.6), respectively. Similarly to (5.20), calibration can be performed for the (pre-treated) calibration pair $\{\mathbf{H}^\star, \mathbf{D}^\star\}$. The sufficient condition for meeting the space-inclusion condition $\mathbf{h}_n^\star \in \mathcal{S}_r(\mathbf{H}^\star)$ becomes $\varepsilon_p(\begin{bmatrix} \mathbf{N}_n \\ \mathbf{C}_{in,n} \end{bmatrix}, \begin{bmatrix} \mathbf{N} \\ \mathbf{C}_{in} \end{bmatrix}) = 0$, i.e., in contrast to calibration using data in reaction-variant form, also the inlet concentrations $\mathbf{C}_{in,n}$ of the new set is dependent on the calibration set ($\varepsilon_p(\mathbf{C}_{in,n}, \begin{bmatrix} \mathbf{N} \\ \mathbf{C}_{in} \end{bmatrix}) = 0$).

5.3.1.4 Discussion of the space-inclusion results

Subsection 5.3.1.1 showed for general rank-deficient spectral data that if the space-inclusion condition $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$ is verified, then the concentrations of the species of interest can be predicted from new spectra, even in the presence of interferents.

If rank deficiency in the spectral data is caused by reactions,

Subsection 5.3.1.2 proposed sufficient conditions to check *a priori* the space-inclusion condition $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$, i.e., *before* the generation of $\mathbf{a}_n$. In Subsection 5.3.1.3, methods were proposed to help choose the experimental concentration matrix and/or the stoichiometries of the new set independently of the calibration set. For all results, it is assumed that (AA1) $\mathbf{A}$ and $\mathbf{a}_n$ are available, and (AA2) the calibration experimental concentration matrix $\mathbf{C}_x$ is so chosen that the R independent reactions are *present and active* in the calibration set.

Table 5.1 gives an overview of the sufficient conditions required for meeting the space-inclusion depending on the corresponding calibration pair ($\{\mathbf{A}, \mathbf{C}_k\}$ or $\{\mathbf{H}, \mathbf{D}_k\}$), the knowledge required to enable meeting the sufficient conditions, and implications of the sufficient conditions with respect to the possibility of choosing the stoichiometries and/or the experimental concentration matrices (initial and inlet concentrations) of the new set independently of the calibration set. Table 5.2 indicates the knowledge required and proposes guidelines to meet the various expressions of the sufficient conditions in Table 5.1.

Five different sets of sufficient conditions are proposed that guarantee the space-inclusion condition $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$, and one sufficient condition that guarantees the space-inclusion condition $\mathbf{h}_n \in \mathcal{S}_r(\mathbf{H})$. For the first two sufficient conditions regarding $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$ (5.23 and 5.24), the stoichiometries $\mathbf{N}_n$ and the experimental concentration matrix $\mathbf{C}_{x,n}$ of the new set are *fixed* by the choice of those of the calibration set. The difference between both is that if the reaction system is the same for both the calibration and the new set, then, in contrast to (5.23), the knowledge of the (possibly unknown) stoichiometries is not required for (5.24), since the choice of $\mathbf{C}_{x,n}$ does not depend on $\mathbf{N}$. However, (5.24) is more conservative than (5.23), as it is illustrated by Examples 5.18 and 5.19 below.

For (5.25)–(5.28), the experimental concentration matrix $\mathbf{C}_{x,n}$ can be chosen *independently* of the calibration set. In addition, the stoichiometries can be chosen freely for (5.26) and (5.27), and the experimental concentration matrix of the calibration set for (5.28). The full-rank conditions in (5.26) and (5.27) can be met by using the

Table 5.1. Overview of the sufficient conditions required for meeting the space-inclusion depending on the corresponding calibration pair ($\{\mathbf{A}, \mathbf{C}_k\}$ or $\{\mathbf{H}, \mathbf{D}_k\}$), the knowledge required to enable meeting the sufficient conditions, and implications of the sufficient conditions with respect to the possibility of choosing the stoichiometries and/or the experimental concentration matrices (initial and inlet concentrations) of the new set independently of the calibration set.

Calibration pair	Reference	Sufficient conditions for meeting the space-inclusion conditions	Knowledge required for sufficient conditions	Free choice of	
$\{\mathbf{A}, \mathbf{C}_k\}$	Prop. 5.12:	$\operatorname{rank}(\mathbf{A}) = (R+p+1) \wedge \varepsilon_p(\mathbf{N}_{e,n}, \mathbf{N}_e) = 0$ [a]	$\mathbf{N}$, $\mathbf{N}_n$; $\mathbf{C}_x$, $\mathbf{C}_{x,n}$; $S_x \geq p+1$, $p \leq S-R-1$	–	(5.23)
	Cor. 5.13:	$\operatorname{rank}(\mathbf{A}) = (R+p+1) \wedge \varepsilon_p(\mathbf{N}_n, \mathbf{N}) = 0$ $\wedge \varepsilon_p(\mathbf{C}_{x,n}, \mathbf{C}_x) = 0$ [a]	$\mathbf{N}$, $\mathbf{N}_n$; $\mathbf{C}_x$, $\mathbf{C}_{x,n}$; $S_x \geq p+1$, $p \leq S-R-1$	–	(5.24)
	Cor. 5.14:	$\operatorname{rank}(\mathbf{A}) = (R+p+1) \wedge \varepsilon_p(\mathbf{N}_n, \mathbf{N}) = 0$ $\wedge \operatorname{rank}(\mathbf{C}_x) = S_x$ [a]	$\mathbf{N}$, $\mathbf{N}_n$; $\mathbf{C}_x$; $p = S_x - 1$	$\mathbf{C}_{x,n}$	(5.25)
	Prop. 5.11:	$\operatorname{rank}(\mathbf{A}) = S$	S; $S_x \geq S-R$, $(p \geq S_x - 1)$	$\mathbf{C}_{x,n}$, $\mathbf{N}_n$, $\mathbf{N}$	(5.26)
	Channel selection:	$\operatorname{rank}(\mathbf{A}_+) = S_m$	S_m; $S_{x,m} \geq S_m - R_o$, $(p \geq S_{x,m} - 1)$	$\mathbf{C}_{x,n}$, $\mathbf{N}_n$, $\mathbf{N}$	(5.27)
$\{\mathbf{H}, \mathbf{D}_k\}$	Prop. 5.16:	$\varepsilon_p(\mathbf{N}_n, \mathbf{N}) = 0$ [a]	$\mathbf{N}$, $\mathbf{N}_n$ [b]	$\mathbf{C}_{x,n}$, $\mathbf{C}_x$	(5.28)

[a] It is assumed that $\operatorname{rank}(\mathbf{E}) = S$.

[b] For calibration/prediction of concentrations, the following quantities must be known in addition: $\mathbf{A}_x$, $\mathbf{A}_{x,n}$, $\mathbf{C}_{x,k}$, $\mathbf{C}_{x,n,k}$, $\mathcal{H}$, $\mathbf{h}_n$.

Table 5.2. Knowledge required and guidelines proposed to meet the various expressions of the sufficient conditions in Table 5.1.

Expression	Knowledge required	Guidelines
$\text{rank}\,(\mathbf{A}) = R{+}p{+}1$	$\mathbf{N}$; $\mathbf{C}_x$; $S_x \geq p{+}1$, $p \leq S - R - 1$	Experimental or non-experimental rank-increasing operations (Subsections 5.2.2.1 or 5.2.2.2)
$\text{rank}\,(\mathbf{A}) = S$	S; $S_x \geq S - R$, $(p \geq S_x - 1)$	
$\text{rank}\,(\mathbf{A}_+) = S_m$	S_m; $S_{x,m} \geq S_m - R_o$, $(p \geq S_{x,m} - 1)$	Channel selection (Subsection 5.2.2.3)
$\varepsilon_p(\mathbf{N}_{e,n}, \mathbf{N}_e) = 0$	$\mathbf{N}$, $\mathbf{N}_n$; $\mathbf{C}_x$, $\mathbf{C}_{x,n}$	Proper choice of $\mathbf{C}_{x,n}$ or $\mathbf{N}_n$ depending on $\mathbf{N}$ and/or $\mathbf{C}_x$
$\varepsilon_p(\mathbf{N}_n, \mathbf{N}) = 0$	$\mathbf{N}$, $\mathbf{N}_n$	
$\varepsilon_p(\mathbf{C}_{x,n}, \mathbf{C}_x) = 0$	$\mathbf{C}_x$, $\mathbf{C}_{x,n}$	
$\text{rank}\,(\mathbf{C}_x) = S_x$	$\mathbf{C}_x$; $p = S_x - 1$	Linearly-independent variation of all S_x species

rank-increasing operations proposed in Subsections 5.2.2.1–5.2.2.3. The disadvantage with these conditions is that at least $(S - R)$ or $(S_m - R_o)$ species must be available for linearly-independent variation, respectively[1]. However, often less than $(S - R)$ or $(S_m - R_o)$ are available for linearly-independent variation. For such a case, (5.25) can be applied, which suggests to vary the available S_x species independently such that full rank of $\mathbf{C}_x$ is guaranteed. In contrast to (5.26) and (5.27), however, $\mathbf{N}$ and $\mathbf{N}_n$ cannot be freely specified for (5.25), and no other species can be present in the initial and inlet concentrations of the new set than the S_x species in the calibration set.

Example 5.18 (Conservatism of the sufficient conditions in (5.24)**)**
Consider the batch reaction system $\mathcal{X}_1 \rightarrow \mathcal{X}_2 \rightarrow \mathcal{X}_3$ with all three species absorbing ($S = 3$ $R = 2$, $p = 0$), and assume $\text{rank}\,(\mathbf{E}) = S$. According to Propositions 5.5 and 4.23, $\mathbf{A}$ and $\mathbf{N}_e$ are full rank $(R + p + 1) = 3 = S$ when $c_{A,0} \neq 0$. Thus, according to (5.26), the initial concentrations $\mathbf{C}_{x,n}$ and the stoichiometries $\mathbf{N}_n$ of the new set can be arbitrary. This is also found in (5.23), since the condition $\varepsilon_p(\mathbf{N}_{e,n}, \mathbf{N}_e) = 0$ is trivially satisfied.
According to (5.24), the following two conditions are required for meeting the space-inclusion condition $\mathbf{a}_n \in \mathcal{S}_r\,(\mathbf{A})$: (i) for meeting the condition $\varepsilon_p(\mathbf{N}_n, \mathbf{N}) = 0$, the stoichiometries of the new set must be a linear combination of those of the calibration set, and (ii) the condition $\varepsilon_p(\mathbf{C}_{x,n}, \mathbf{C}_x) = 0$, which imposes $\mathbf{c}_{0,n} = \alpha\,\mathbf{c}_0$, where α is a constant. This illustrates the conservatism of the sufficient conditions in (5.24).

Example 5.19 (Rank-increasing operations)
Consider the following reaction system with one inlet stream ($p = 1$): $\mathcal{X}_1 \rightarrow \mathcal{X}_2 + \mathcal{X}_3$, $\mathcal{X}_2 \rightarrow \mathcal{X}_4 + \mathcal{X}_5$. It is assumed that all five species absorb ($S = 5$), the reactions are the same for both the calibration and the new sets ($\varepsilon_p(\mathbf{N}_n, \mathbf{N}) = 0$), and only Species

[1]Note that the channels must be so chosen that the S_k species are a subset of the S_m selected species.

$\mathcal{X}_1$ and $\mathcal{X}_2$ can be used in the experimental concentration matrices of the calibration and the new sets, i.e., $S_x = 2$. Let $\mathbf{c}_0$ and $\mathbf{c}_{in}$ be so chosen that rank $(\mathbf{C}_x) = 2 = S_x$. Although rank $(\mathbf{A}) = R+p+1 = 4 < S$, $\mathbf{C}_{x,n}$ can be freely specified, since the sufficient conditions in (5.25) are met.

The disadvantage of (5.25)–(5.27) is that certain conditions for the S_x species in calibration set and/or new set must be fulfilled (in addition to Assumption AA2). This disadvantage is not encountered in (5.28), since the reaction-invariant part is subtracted from the spectral and concentration measurements (data pre-treatment to reaction-variant form). In contrast to (5.26) and (5.27), however, (5.28) involves the knowledge of $\mathbf{N}$ and $\mathbf{N}_n$. Data pre-treatment to reaction-variant form requires the knowledge of the reaction invariants including the inlet spectra $\mathbf{A}_{in}$ of both the calibration and new sets. In practical applications, this might be a drawback, since the inlet spectra measured externally can be different from those inside the reacting mixtures due to changing chemical-physical interactions (Burns and Ciurzak, 1992).

In Table 5.1, channel selection is used for attaining full rank of the spectral matrix, i.e., a special case of (5.25). In addition, channel selection can be applied to (5.23)–(5.25) and (5.28), where the projection errors of the extended stoichiometries, the stoichiometries, and the experimental concentration matrices must be computed for the S_m species.

5.3.1.5 Simulation study

Data generation (Example 2.13c)

Certain aspects of the space-inclusion results are illustrated by means of an isothermal, varying-density semibatch reaction system (2.54), involving $R = 2$ independent reactions and conducted in a solvent $\mathcal{X}_6$. Two types of reaction rates are studied for the first reaction:

(1) Elementary reaction:

$$r_1^{(1)}(c_1) = \kappa_1 \, c_1^2$$

(2) Inhibited reaction:

$$r_1^{(2)}(c_1, c_5) = \frac{\kappa_{11} \, c_1^2}{1 + \kappa_{12} \, c_5}$$

The reaction is conducted in a non-absorbing solvent $\mathcal{X}_6$. The vectors $\mathbf{c} \equiv [c_1, \, c_2, \, c_3, \, c_4, \, c_5]^{\mathrm{T}}$, $\mathbf{c}_t \equiv [\mathbf{c}^{\mathrm{T}}, \, c_6]^{\mathrm{T}}$, and $\mathbf{c}_k = [c_1, \, c_2]^{\mathrm{T}}$ represent the concentrations of the $S = 5$ absorbing species, all $S_t = 6$ species, and the $S_k = 2$ known (measured) species, respectively. See Appendix A.2 for the numerical values of the parameters.

Four different runs are generated. Table 5.3 lists the numerical values of the number of observations K, the initial volume V_0, the final batch time t_f, the initial concentrations, the inlet concentrations $\mathbf{c}_0$, and the kinetic expression used for the first reaction. Note that the inlet streams do not contain solvent.

Table 5.3. Number of observations K, initial volume V_0 [l], final batch time t_f [h], initial concentrations $\mathbf{c}_{t,0}$ [M], and inlet concentrations $\mathbf{c}_{in}$ [M] for the Runs R1–R4.

				$\mathbf{c}_{t,0}$					$\mathbf{c}_{in}$		
Run	K	V_0	t_f	$\mathcal{X}_1$	$\mathcal{X}_2$	$\mathcal{X}_3$	$\mathcal{X}_4$	$\mathcal{X}_6$	$\mathcal{X}_1$	$\mathcal{X}_4$	Kinetics r_1
R1	85	10	7	13.73	0	0	0.86	3.22	8.55	8.55	$r_1^{(1)}$
R2	73	15	6	8.41	0	0	1.68	8.41	16.15	0	$r_1^{(2)}$
R3	61	15	5	13.49	2.41	0	0.96	1.20	8.55	8.55	$r_1^{(1)}$
R4	73	15	6	8.94	1.12	8.94	0.71	3.58	16.15	0	$r_1^{(2)}$

The concentration profiles are presented in Figure 5.1. The density, volume, and feed profiles for the 4 runs R1–R4 are illustrated in Figure 5.2. Figure 5.3 depicts the (highly-overlapping) pure-component spectra ($L = 301$) of all absorbing species, generated from real NIR pure-component spectra. These are used to construct $\mathbf{A}$.

Three cases are considered:

(C1) Calibration set: Run R1 ($p = 1$ generalized inlet). New set: Run R2.

(C2) Calibration set: Run R1 ($p = 1$ generalized inlet). New set: Run R3.

(C3) Calibration set: Runs R1 and R3 (the corresponding concentration and spectral matrices are appended with respect to the common column order; *column-wise appending*; $p = 2$ generalized inlets, since the inlet concentrations of Runs R1 and R3 are the same). New set: Run R4.

Key differences between Runs R1 and R2 (Case C1) are the inlet concentrations, the feed profiles, and the kinetics of the first reaction. As far as the feed is concerned, the number of species involved is also different. However, the set of species initially present in the reactor are the same, though the concentrations are different. Comparing Runs R1 and R3 (Case C2), it can be noted that the kinetics, the inlet concentrations and the feed profiles are the same. However, the initial concentrations for R3 involves Species $\mathcal{X}_2$, which is not initially present in the calibration run R1 but is produced by the first reaction (in both the calibration and new runs). Finally, the initial concentrations of R4 (Case C3) involves Species $\mathcal{X}_3$, which is not present initially in calibrations runs R1 and R3 but is produced by the second reaction (in both the calibration and new runs).

Nullity reduction and data pre-treatment to reaction-variant form

Calibration models are built using the pairs $\{\mathbf{A},\mathbf{C}_k\}$ and $\{\mathbf{H},\mathbf{D}_k\}$ according to (5.15) and (5.22).

A hypothetical Case C4 that uses nullity reduction is introduced. Assume that Species $\mathcal{X}_2$ does not absorb. This is equivalent to saying that channel selection has been performed according to Subsection 5.2.2.3 with $S_- = 5$ and $S_+ = S_m = 4$. Case C4 is similar to Case C2 but for $\mathcal{X}_2$ not absorbing, i.e., the calibration set is based on Run R1, and the new set on Run R3. A calibration model is built, in which the prediction

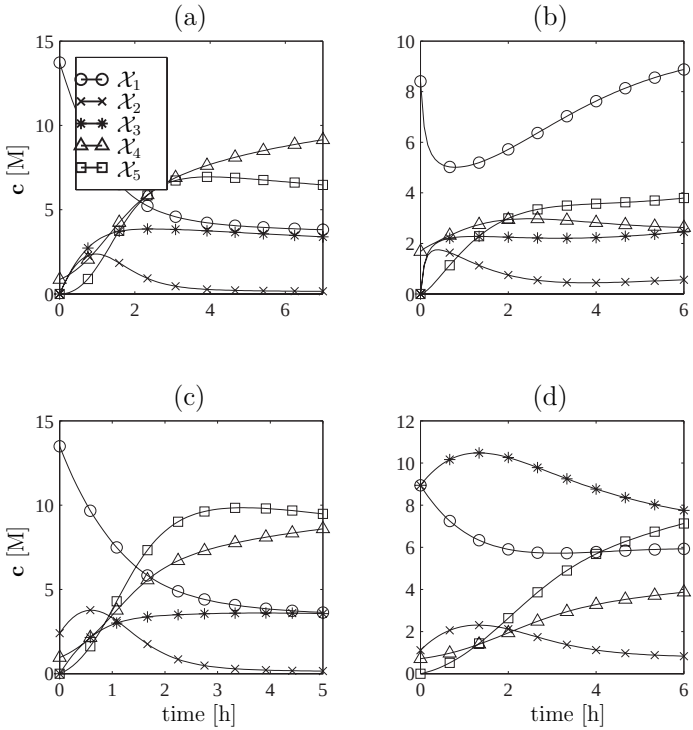

Figure 5.1. Example 2.13c: The concentrations **c** of the absorbing species for the 4 runs. (a) Run R1, (b) Run R2, (c) Run R3, (d) Run R4.

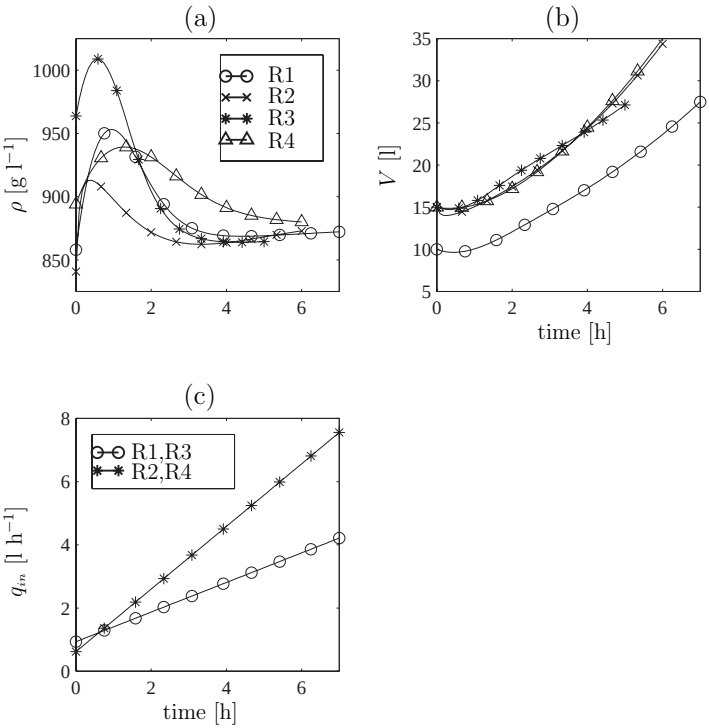

Figure 5.2. Example 2.13c: (a) Density ρ, (b) volume V, and (c) feed profiles q_{in} for Runs R1–R4.

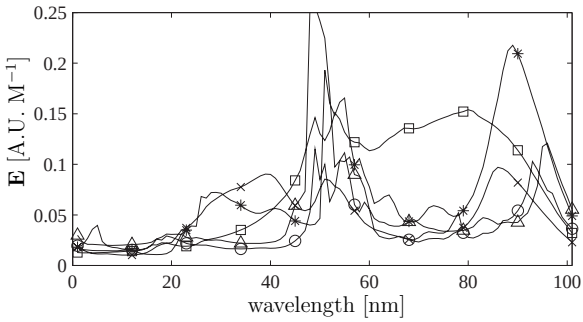

Figure 5.3. NIR pure-component spectra **E** for Example 2.13c. See Figure 5.1 for legend.

of only Species $\mathcal{X}_1$ is considered ($S_k = 1$) using the calibration and prediction pair $\{\mathbf{A},\mathbf{C}_k\}$ with $\mathbf{C}_k = \mathbf{c}_1$.

Statistical considerations

The spectral data of all calibration and new sets are corrupted by noise generated by multiplying the signal by a number from a uniform distribution in the range $[-0.01, 0.01]$ (multiplicative or heteroscedastic noise). Such noise was experimentally determined for spectra in the MIR range using ReactIR 1000 (ReactIR, 1994). Also, $\mathbf{C}_k$ is corrupted by Gaussian noise generated with a standard deviation chosen as 2.5% of the maximum of the column-means of $\mathbf{C}_k$.

The internal (fit) and external prediction performances of the calibration models based on the calibration and the new set, respectively, are compared as a function of the number of principal components (PCs) retained in the calibration model (see Appendix C.4.2).

The normalized root mean squared error (NRMS; see C.39) is calculated for both the calibration set (NRMS_C) and the new set (NRMS_N). For each of the noisy calibration sets, a calibration model is estimated and cross validated (CV) to determine the number of significant PCs (see Appendix C.4.2.2). Here, a leave-one-block-out CV with a block size of 4 spectra and the average of 3 random shuffles in time direction is used.

Since the sample covariances of principal components strongly depend on the particular noise added (Jackson, 1991), extreme values for NRMS can occur. To avoid this, the NRMS values for 200 calibration sets corrupted by different noise sequences are considered (Monte–Carlo simulation). The averages of NRMS_C and NRMS_CV for all noisy calibration sets are calculated.

For each calibration model, the prediction performance of 200 noisy new sets is studied. As far as the new sets are concerned, the averages of both NRMS_N and $\hat{\mathbf{C}}_{k,n}$ for each combination of a noisy calibration set with all noisy new sets are calculated, thereby performing 200×200 simulation runs in total.

Discussion of the results

The following results are presented in Table 5.4 for Cases C1–C4: (i) the ranks of the noise-free spectral matrices $\mathbf{A}$ and $\mathbf{H}$ of the calibration sets, (ii) the projection errors for $\mathbf{N}_e$ and $\mathbf{N}$, and (iii) the projection error for $\mathbf{C}_x$. From CV, the optimal numbers of PCs, A^*, to retain in the calibration models (PCR and SIMPLS) correspond to the rank predicted in Table 5.4. Furthermore, Table 5.4 contains the NRMS_N of SIMPLS for the calibration pair given in curly brackets and A^*. The ranks of the noise-free matrices $\mathbf{A}$ and $\mathbf{H}$ for Cases C1–C4 are $(R + p + 1)$ and R, respectively.

For Case C1, the sufficient conditions (5.24) for space inclusion are fulfilled (see Table 5.4) and, thus, correct external prediction is expected. Figure 5.4 shows the NRMS for external prediction, cross validation, and internal prediction for the calibration pair $\{\mathbf{A}, \mathbf{C}_k\}$ as a function of (i) the number of principal components (PC) retained in the calibration model and (ii) the regression methods PCR and SIMPLS. Using $A^* = 4$, the concentrations of Species $\mathcal{X}_1$ and $\mathcal{X}_2$ are correctly predicted (internally, cross validation, externally) from the calibration sets (see Figure 5.4). Note that, since the prediction

Table 5.4. Numerical results for Cases C1–C4 with $S = 5$, $R = 2$, $S_k = 2$, $p = 1$, and $\text{rank}(\mathbf{E}) = S$: (i) the ranks of the noise-free matrices $\mathbf{A}$ and $\mathbf{H}$ of the calibration sets, (ii) the projection errors for $\mathbf{N}_e$ and $\mathbf{N}$, (iii) the projection error for $\mathbf{C}_x$, and (iv) NRMS_N of SIMPLS for the calibration pair given in curly brackets and A^* given by the ranks in (i). The cases where the space-inclusion condition $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$ or $\mathbf{h}_n \in \mathcal{S}_r(\mathbf{H})$ is not met are marked bold.

	Case C1 R1/R2	Case C2 R1/R3	Case C3 (R1+R3)/R4	Case C4 R1/modified R3
$\text{rank}(\mathbf{A}) = R + p + 1$	4	4	5	4
$\text{rank}(\mathbf{H}) = R$	2	2	2	2
$\varepsilon_p(\mathbf{N}_{e,n}, \mathbf{N}_e)$[a]	0	**1.6**	0	0
$\varepsilon_p(\mathbf{N}_n, \mathbf{N})$[a]	0	0	0	0
$\varepsilon_p(\mathbf{C}_{x,n}, \mathbf{C}_x)$[a]	0	**2.4**	8.94	0
$\text{NRMS}_\text{N}\{\mathbf{A}, \mathbf{C}_k\}$	0.024	**0.342**	0.038	0.026
$\text{NRMS}_\text{N}\{\mathbf{H}, \mathbf{D}_k\}$	0.024	0.022	0.027	0.023

[a] For Case C4, the projection errors of the extended stoichiometries, the stoichiometries, and the experimental concentration matrices are computed for the S_m species.

performance of PCR and SIMPLS for A^* is nearly the same (also for Cases C2–C4), only the concentrations predicted from the SIMPLS model are shown in Figures 5.5 and 5.6, which compare true and predicted concentrations.

The other scenarios where the sufficient conditions for space inclusion are fulfilled are: (i) In Cases C1–C4 for the calibration pair $\{\mathbf{H},\mathbf{D}_k\}$, since the same reaction network is considered for both the calibration and the new sets, the condition $\varepsilon_p(\mathbf{N}_n, \mathbf{N}) = 0$ is trivially fulfilled. (ii) In Case C3 for the calibration pair $\{\mathbf{A},\mathbf{C}_k\}$, since the rank of the spectral matrix is increased to $S = 5$ by appending the spectral data of Run R3 to those of Run R1, the space-inclusion condition $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$ is trivially fulfilled according to (5.26). (iii) In Case C4 for the calibration pair $\{\mathbf{A},\mathbf{C}_k\}$, since the rank of the spectral matrix is full ($\text{rank}(\mathbf{A}_+) = 4 = S_m$) after "channel selection" ($R_o = R = 2$), the space-inclusion condition $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$ is trivially fulfilled according to (5.27). As expected from the sufficient conditions, Figures 5.5 and 5.6 show that the concentrations of Scenarios i–iii are well predicted.

For Case C3 using the calibration pair $\{\mathbf{A},\mathbf{C}_k\}$, Table 5.4 shows that $\varepsilon_p(\mathbf{C}_{x,n}, \mathbf{C}_x) \neq 0$, while $\varepsilon_p(\mathbf{N}_{e,n}, \mathbf{N}_e) = 0$. This is due to the fact that for Case C3, Species $\mathcal{X}_3$ is present in $\mathbf{C}_{x,n}$ but not in $\mathbf{C}_x$. However, it cannot be concluded from (5.24) that the concentrations will be predicted correctly, since (5.24) represents only *sufficient* conditions. This illustrates the conservatism of the sufficiency of condition (5.24).

For Case C2 using the calibration pair $\{\mathbf{A},\mathbf{C}_k\}$, $\varepsilon_p(\mathbf{N}_{e,n}, \mathbf{N}_e) \neq 0$ and $\varepsilon_p(\mathbf{C}_{x,n}, \mathbf{C}_x) \neq 0$ (see Table 5.4). This is due to Species $\mathcal{X}_2$ being initially present in the new set, but absent initially in the calibration set. Note that Species $\mathcal{X}_2$ is present also in the calibration set since it is produced by the first reaction. However, the corresponding

concentrations depend linearly on the concentrations of the remaining species due to the reactions. Thus, the calibration data do not span the extra dimension necessary to predict the concentrations of the new Run R3. By the violation of the sufficient conditions in (5.23) and (5.24), there usually is a high possibility that the space-inclusion condition $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$ is not fulfilled, and the concentrations are not predicted correctly. Figure 5.5b shows that the concentrations are not predicted correctly. Note that, in contrast to Case C2, $\varepsilon_p(\mathbf{C}_{x,m,n}, \mathbf{C}_{x,m}) = 0$ for Case C4, since the effect of Species $\mathcal{X}_2$ is eliminated.

In general, better prediction performance is obtained by using the pair $\{\mathbf{H}, \mathbf{D}_k\}$ compared to $\{\mathbf{A}, \mathbf{C}_k\}$ (see NRMS_n in Table 5.4). This is only true when the terms related to the reaction invariants (e.g., the initial and inlet concentrations and the inlet spectra of the calibration and the new sets, the inlet flowrates and the volume) are known perfectly. For such cases, the noise is restricted to a lower-dimensional space (here, 2-dimensional), which often results in better performance.

An interesting unexplained case occurs when 6 PCs are retained in Case C2 with calibration pair $\{\mathbf{A}, \mathbf{C}_k\}$: the concentration of Species $\mathcal{X}_1$ is correctly predicted using SIMPLS, though Species $\mathcal{X}_2$ is not predicted correctly.

5.3.2 Multivariate curve-resolution techniques (FA)

As mentioned earlier, concentration measurements of some species of interest are often not available (due to difficulties/costs in sampling, sample preparation, and development of analytical techniques) and, thus, calibration of spectral measurements for the purpose of concentration prediction is not possible.

For such situations, factor analysis (FA) or *multivariate curve-resolution techniques* have been used successfully for *non-reacting mixtures* to resolve the concentrations and the corresponding pure-component spectra from spectral data. It involves two steps: (i) PCA (see Section C.1), (ii) rotation of scores and loading matrices into physically-meaningful quantities using prior knowledge about the reaction system and the type of instrumental response.

FA methods are used to resolve the rotational ambiguities and, in some cases, also the intensity ambiguities (see also Appendix C.3). For non-reacting species, Manne (1995) provides conditions for the unique resolution of the rotational ambiguity. For the *quantitative* estimation of concentrations and pure-component spectra, an external standard must be provided.

Most of the FA techniques rely on the assumption that the rank of the spectral data is determined by the number of absorbing species S and, thus, the column and row spaces of the spectral data correspond to those of the (unknown) concentration and pure-spectra matrices, respectively.

For reacting mixtures, however, this assumption is generally not satisfied due to rank-deficient spectral matrices, as was shown in Section 5.2.1. If a sufficient number of species is available for independent variation of the initial and inlet concentrations, Amrhein *et al.* (1996) proposed to use nullity-reducing operations that guarantee rank S for the reacting spectral data (see Subsection 5.2.2).

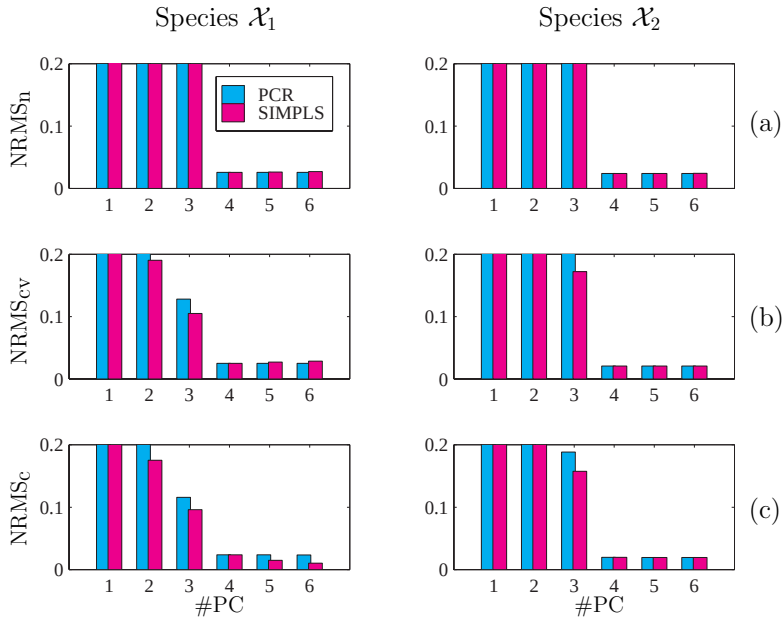

Figure 5.4. Example 2.13c: External prediction, cross validation and internal prediction using the calibration pair $\{\mathbf{A}, \mathbf{C}_k\}$ in Case C1: (a) internal prediction, (b) cross validation, and (c) external prediction for Species $\mathcal{X}_1$ and $\mathcal{X}_2$ as a function of (i) the number of principal components (PC) retained in the calibration model, and (ii) the regression methods PCR and SIMPLS. Bars are cut off at 0.2.

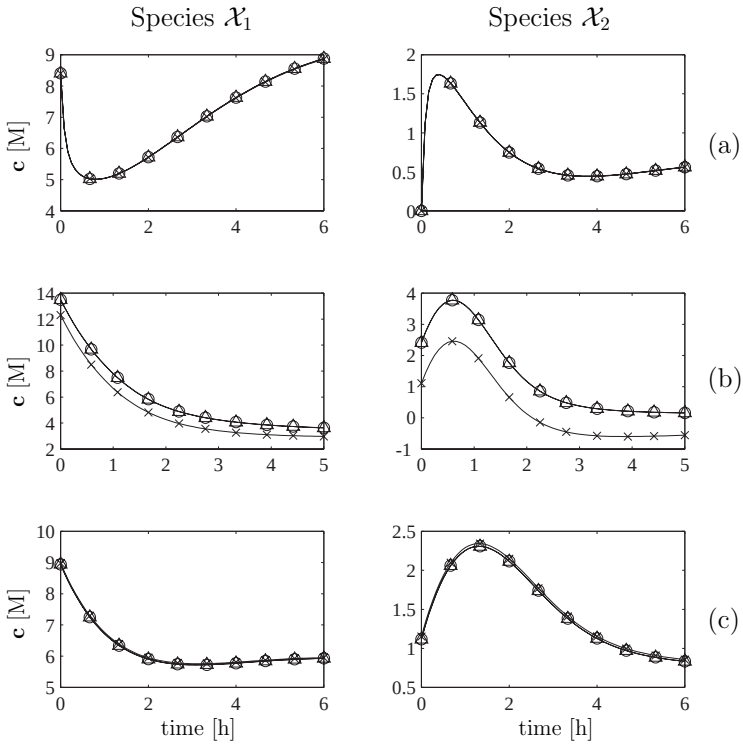

Figure 5.5. Example 2.13c: Predicted concentrations of Species $\mathcal{X}_1$ and $\mathcal{X}_2$ using SIMPLS for: (a) Case C1, (b) Case C2, and (c) Case C3. ($\circ$) true, ($\times$) predicted from $\{\mathbf{A}, \mathbf{C}_k\}$, ($\triangle$) predicted from $\{\mathbf{H}, \mathbf{D}_k\}$. The optimal number of principal components A^* retained in the calibration models is given in Table 5.4.

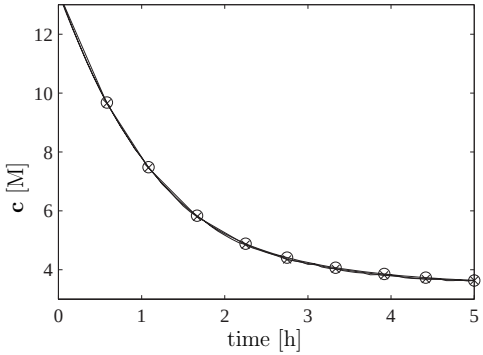

Figure 5.6. Example 2.13c: Predicted concentrations of Species $\mathcal{X}_1$ using SIMPLS for Case C4. (∘) true, (×) predicted from $\{\mathbf{A}, \mathbf{c}_1\}$. The optimal number of principal components A^* retained in the calibration model is 4.

In this subsection, FA methods will be proposed for the *quantitative* estimation of concentrations for *all* the species at *all* the observations from (rank-deficient) reacting spectral data (resolving both the rotational *and* the intensity ambiguities). In contrast to standard FA methods, the proposed FA methods are usually not able to estimate the pure-component spectra. It will be assumed that the pure-component spectra are not available, nullity-reducing operations may not guarantee rank S for the reacting spectral data, but the reaction-invariant part is known. Thus, the (unknown) reaction-variant part of the spectral data (i.e., also of the concentration data) remains to be estimated. It will be shown that, since no concentration measurement is available for calibration, some implicit or explicit information regarding the kinetic structure need to be available for the construction of a calibration set. Similarly to calibration, a space-inclusion condition must be verified.

Subsection 5.3.2.1 formulates the problem of quantitative concentration estimation from reacting spectral data using factor analysis. In Subsection 5.3.2.2, methods to determine the number of independent reactions and absorbing species from reacting spectral data are discussed. Subsections 5.3.2.3 and 5.3.2.4 propose various rotational methods based on the knowledge of kinetic structure (explicit information) and the presence of irreversible or reversible reactions (implicit information), respectively. Simulated examples illustrate the proposed methods.

5.3.2.1 Factor analysis as a calibration problem

The goal is to estimate the concentration matrix $\mathbf{C}$, i.e., the concentrations of *all* S species at *all* K observations. It will be shown that concentration estimation of *some* of the S species is possible as well. Let Assumptions A1–12 in Appendix D be verified and the spectral matrix and the quantities necessary to reconstruct the reaction invariants be known/measured.

Since the reaction-invariant part of $\mathbf{C}$ is known, only the reaction-variant part remains to be estimated. However, in contrast to the calibration methods proposed in Subsection 5.3.1, no calibration concentration measurement is available. Instead, prior knowledge regarding the kinetic structure or the type of reactions will be used to construct the calibration concentrations. The following prior-knowledge candidates are considered:

(K1) Non-negativity of concentrations,
(K2) Stoichiometric matrix $\mathbf{N}$,
(K3) Structure of the R reaction rates $\mathbf{r}(\mathbf{c}, \boldsymbol{\theta})$ with unknown parameters $\boldsymbol{\theta}$,
(K4) Type of the reactions (reversible vs. irreversible),
(K5) Identity of the species/reaction that cause local rank changes (see Appendix G for the definition of local rank).

Knowledge K3 considers *explicit* information, while Knowledge K4 uses *implicit* information regarding the kinetic structure.

With the knowledge/measurement of $\mathbf{A}$, $\mathbf{A}_x$, V_0, $\mathbf{q}_{in}(k)$ and $q_{out}(k)$ for all k, the spectral data can be brought to reaction-variant form $\mathbf{H}$ (see 5.5):

$$\mathbf{H} = \mathbf{X}\,\boldsymbol{\Theta}, \qquad \boldsymbol{\Theta} \equiv \mathbf{N}\,\mathbf{E}, \tag{5.29}$$

where $\boldsymbol{\Theta}$ is the $R \times L$ *reaction spectral matrix*. $\mathbf{A}$ and $\mathbf{H}$ are assumed to be noise free.

Let $\mathbf{C}_r$ be a $K_r \times S$ matrix defined as $\mathbf{C}_r \equiv \left[\begin{smallmatrix} \mathbf{C}_x \\ \acute{\mathbf{C}} \end{smallmatrix}\right]$, where $\acute{\mathbf{C}}$ is a $\acute{K} \times S$ submatrix of $\mathbf{C}$ that contains concentrations during the ongoing reactions. Furthermore, let $\mathbf{A}_r$ $(K_r \times L)$ and $\acute{\mathbf{A}}$ $(\acute{K} \times L)$ be the spectral matrices corresponding to $\mathbf{C}_r$ and $\acute{\mathbf{C}}$, respectively. Let $\mathbf{D}_r$ be the corresponding $K_r \times S$ submatrix of $\mathbf{D}$, $\mathbf{X}_r$ be the corresponding $K_r \times R$ submatrix of $\mathbf{X}$, and $\mathbf{H}_r$ the corresponding $K_r \times L$ submatrix of $\mathbf{H}$.

The pair $\{\mathbf{A}_r, \mathbf{C}_r\}$ forms a *constructed* calibration set, while each row of $\mathbf{A}$, $\mathbf{a}_n$ $(n = 1, \ldots, K)$, can be considered as a new spectrum. And similarly for the pair $\{\mathbf{H}_r, \mathbf{D}_r\}$ and the rows of $\mathbf{H}$, $\mathbf{h}_n$ $(n = 1, \ldots, K)$.

The results obtained in Subsection 5.3.1.1 apply directly to the problem of estimating $\mathbf{C}$ from a calibration model that is built from the calibration pair $\{\mathbf{A}_r, \mathbf{C}_r\}$ with the following redefinitions: $\mathbf{A} := \mathbf{A}_r$, $\mathbf{C} := \mathbf{C}_r$. Note that the case of resolving *some* of the species concentrations (partial resolution) is treated in Subsection 5.3.1.1 in the framework of interferents. Thus, this subject is not treated any further here.

Next, the space-inclusion results obtained in Subsection 5.3.1.2 will be adapted to the FA problem.

Proposition 5.20 *Let Assumptions A1–12 in Appendix D be verified. If* $\operatorname{rank}(\mathbf{A}_r) = S$, *then* $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A}_r)$.

Proposition 5.21 *Let Assumptions A1–12 in Appendix D be verified and* $\mathbf{R}$ *an* $(R + p + 1) \times K_r$ *matrix. If*

$$\mathbf{N}_e = \mathbf{R}\,\mathbf{C}_r, \tag{5.30}$$

then $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A}_r)$.*(See Appendix H.4.3 for proof)*

Both Propositions 5.20 and 5.21 provide (*a priori*) conditions for the resolution of $\mathbf{C}$ on the basis of $\mathbf{A}$ and $\{\mathbf{A}_r, \mathbf{C}_r\}$.

If $\text{rank}\,(\mathbf{C}_x) = S$ (assuming all the species to be available for linearly-independent variation), then $\text{rank}\,(\mathbf{A}_r) = S$ and, thus, according to Proposition 5.20, the space-inclusion condition is trivially met. In this case, no concentration $\acute{\mathbf{C}}$ must be known/-estimated during the course of the reaction.

Proposition 5.21 says that concentration $\acute{\mathbf{C}}$ must be known/estimated during the course of the reaction such that the rows of $\mathbf{C}_r$ span the entire row space of $\mathbf{N}_e$. Note that $\mathbf{X}_e$ can be rank deficient.

Next, the results obtained in Subsection 5.3.1.3 will be adapted to the FA problem. Theorem 5.15 applies with the following redefinitions: $\mathbf{H} := \mathbf{H}_r$, $\mathbf{D} := \mathbf{D}_r$. To guarantee the space-inclusion condition $\mathbf{h}_n \in \mathcal{S}_r\,(\mathbf{H}_r)$, the condition $\mathbf{N}_n = \mathbf{R}^\dagger\,\mathbf{N}$ must be met (Proposition 5.16). This condition is trivially met if $\text{rank}\,(\mathbf{D}_r) = \text{rank}\,(\mathbf{X}_r) = R$ is guaranteed.

However, for meeting $\text{rank}\,(\mathbf{X}_r) = R$, R must be known. In the following subsection, methods will be discussed that attempt to determine R from spectral data. The estimation of $\acute{\mathbf{C}}$ and $\mathbf{X}_r$ from Knowledge K1–5 is the subject of Subsections 5.3.2.3 and 5.3.2.4.

5.3.2.2 Determination of R and S

Next, methods to determine R and S from spectral data are discussed.

Determination of R

From Proposition 5.5, $\text{rank}\,(\mathbf{A}) = \text{rank}\,(\mathbf{C})$, and $\text{rank}\,(\mathbf{H}) = \text{rank}\,(\mathbf{D})$. Thus, for the determination of R, it suffices to study the ranks of $\mathbf{C}$ and $\mathbf{D}$ when all the species both absorb and react.

Corollary 5.22 *Let Assumptions A1–12 in Appendix D be verified. Then, the number of independent reactions R can be determined from $K \times S$ spectral matrix $\mathbf{A}$:*

(a) If $\mathbf{A}_{in}$, $\mathbf{A}$, V_0, $\mathbf{q}_{in}(k)$, and $q_{out}(k)$ of reaction systems described by (2.38) are known/measured for all k, then R can be determined from $\mathbf{H}$ in (5.5) as

$$R = \text{rank}\,(\mathbf{H})\,. \tag{5.31}$$

(b) If $\text{rank}\,(\mathbf{C}_x\mathbf{\Gamma}) = p + 1$ and $\mathbf{X}_e$ is full rank, then R can be determined from $\mathbf{A}$ in (5.3) as

$$R = \text{rank}\,(\mathbf{A}) - p - 1. \tag{5.32}$$

Corollary 5.22a shows that, if data pre-treatment to reaction-variant form (5.5) is possible, the number of independent reactions R can be determined directly from $\mathbf{H}$.

If data pre-treatment to reaction-variant form is not possible, the number of independent reactions R must be determined from $\mathbf{A}$. However, similarly to $\mathbf{C}$ (see Theorem 4.18), this is only possible by verifying certain reaction-invariant relationships.

On the determination of S

The number of absorbing species S in a reacting mixture is often unknown *a priori* and, thus, the (column dimension of the) stoichiometric matrix $\mathbf{N}$ is also unknown. In such a case, it is desirable to determine S from spectral data in a way similar to that proposed for R in the previous subsection. However, it will be shown that this goal is impossible to attain, even with the assumption $\text{rank}\,(\mathbf{E}) = S$.

The previous subsection presented methods for the determination of R from the spectral matrix $\mathbf{A}$ or RV-spectral matrix $\mathbf{H}$. For the determination of the number of absorbing species S from $\mathbf{A}$, the rank of $\mathbf{A}$ or, equivalently, $\mathbf{C}$ must be increased to S. Thus, data pre-treatment to reaction-variant form as used for $\mathbf{H}$ is not useful, and the focus will be on the analysis of $\mathbf{A}$ or $\mathbf{C}$.

To guarantee $\text{rank}\,(\mathbf{C}) = S$, according to Corollary 4.4, the two conditions $\text{rank}\,(\mathbf{X}_e) = (R + p + 1)$ and $\text{rank}\,(\mathbf{N}_e) = S$ must be guaranteed. Full rank of $\mathbf{X}_e$ can be experimentally guaranteed according to Proposition 4.5. To guarantee $\text{rank}\,(\mathbf{N}_e) = S$, either S must be known to guarantee $\text{rank}\,(\mathbf{C}_x) = S$ (which implies $\text{rank}\,(\mathbf{N}_e) = S$) or, when $\text{rank}\,(\mathbf{C}_x) < S$, $\mathbf{N}$ (i.e., also its column dimension S) must be known for the computation of $\text{rank}\,(\mathbf{N}_e)$. Thus, the determination of S from $\mathbf{A}$ is a fundamental problem that cannot be solved!

5.3.2.3 Rotation — explicit use of kinetic structures

This subsection discusses the estimation of $\acute{\mathbf{C}}$ (or $\mathbf{X}_r$) from knowledge of the kinetic structure and spectral measurements. It is assumed that the stoichiometric matrix (Knowledge K2) and the structure of the reaction rates $\mathbf{r}(\mathbf{c}, \boldsymbol{\theta})$ (Knowledge K3) are known. Thus, the concentrations are completely described by (4.3) and (4.7) except for the unknown reaction rate parameters $\boldsymbol{\theta}$.

For such a scenario, various methods have been proposed to estimate the unknown reaction rate parameters directly from the spectral data by solving a constrained nonlinear regression problem (Sylvestre and Maggio, 1974; Koch, 1977; Sheiner, 1984, 1985, 1986; Gampp *et al.*, 1988; Maeder and Zuberbühler, 1990; Perkampus and Kaufmann, 1991; Crouch, 1993; Lamberti *et al.*, 1993; Mottola, 1993; Bugnon *et al.*, 1994; Dyson *et al.*, 1997; Van Stokkum, 1997). However, no formal proof for correct resolution in the case of rank-deficient spectral data has been provided. Here, a rigorous proof is provided.

Two approaches are proposed: (M1) estimation of $\acute{\mathbf{C}}$ from measured (rank-deficient) spectral data $\mathbf{A}$, and (M2) estimation of $\mathbf{X}_r$ from measured/reconstructed spectral data in reaction-variant form $\mathbf{H}$. Let Assumptions A1–12 in Appendix D be verified.

For Approach M1, the following constrained nonlinear regression problem is solved:

$$\begin{aligned}
\min_{\boldsymbol{\theta}} J_A, \qquad & J_A \equiv \|\mathbf{A} - \mathbf{A}(\boldsymbol{\theta})\| = \| \left[\mathbf{I}_K - \mathbf{C}(\boldsymbol{\theta})\, \mathbf{C}^{+}(\boldsymbol{\theta})\right]\, \mathbf{A}\| \\
\text{s.t. } & \dot{\mathbf{x}}(\boldsymbol{\theta}) = \mathbf{r}(\mathbf{c}, \boldsymbol{\theta})\, V/\lambda, \qquad \mathbf{x}(0) = \mathbf{0}_R, \\
& \dot{\mathbf{z}} = \mathbf{q}_{in}/\lambda, \qquad \mathbf{z}(0) = \mathbf{0}_p, \\
& \dot{\lambda} = -\frac{q_{out}}{V}\,\lambda, \qquad \lambda(0) = 1, \\
& \mathbf{c}(\boldsymbol{\theta}) = \left(\mathbf{N}^{\mathrm{T}}\, \mathbf{x}(\boldsymbol{\theta}) + \mathbf{C}_{in}\, \mathbf{z} + V_0 \mathbf{c}_0\right) / \mathrm{h}(\mathbf{z}) \\
& \mathbf{C}(\boldsymbol{\theta}) = \boldsymbol{\mathcal{H}}^{-1}\, \left(\mathbf{X}(\boldsymbol{\theta})\, \mathbf{N} + \mathbf{Z}\, \mathbf{C}_{in}^{\mathrm{T}} + V_0 \mathbf{1}_K \mathbf{c}_0^{\mathrm{T}}\right),
\end{aligned} \tag{5.33}$$

where J_A is the cost to minimize, $\|\cdot\|$ an appropriate matrix norm (e.g., the Frobenius norm), and $\mathbf{A}(\boldsymbol{\theta})$, $\mathbf{x}(\boldsymbol{\theta})$, $\mathbf{X}(\boldsymbol{\theta})$, $\mathbf{c}(\boldsymbol{\theta})$, and $\mathbf{C}(\boldsymbol{\theta})$ the $K \times L$ spectral matrix, the reaction variants, the $K \times R$ matrix of reaction variants, the concentrations, and the $K \times S$ concentration matrix calculated from the dynamic model, respectively.

The second expression for J_A was obtained by assuming $\mathrm{rank}\,(\mathbf{A}) = A$. Thus, $\mathbf{C} = \mathbf{T}\,\mathbf{P}^{\mathrm{T}}$ and $\mathbf{A} = \mathbf{T}\,\mathbf{P}^{\mathrm{T}}\,\mathbf{E}$ with $\mathbf{T}$, $\mathbf{P}$ being full-rank matrices of dimension $K \times A$ and $S \times A$, respectively. Therefore, $\mathbf{A}(\boldsymbol{\theta}) = \mathbf{C}(\boldsymbol{\theta})\,\mathbf{E} = \mathbf{C}(\boldsymbol{\theta})\,\mathbf{C}^{+}(\boldsymbol{\theta})\,\mathbf{A} = \mathbf{T}\,(\mathbf{P}^{\mathrm{T}}\mathbf{P})\,(\mathbf{T}^{+}\mathbf{T})\,\mathbf{P}^{\mathrm{T}}\mathbf{E} = \mathbf{A}$. This way, (5.33) minimizes the projection error of $\mathbf{A}$ on the column space of $\mathbf{C}$ by optimizing the reaction rate parameters $\boldsymbol{\theta}$.

The solution to the optimization problem is $\hat{\boldsymbol{\theta}}$ with the corresponding estimate $\hat{\mathbf{C}} = \hat{\mathbf{C}}(\hat{\boldsymbol{\theta}})$. Thus, $\hat{\boldsymbol{C}}$ is implicitly estimated.

For Approach M2, the following cost function is minimized with respect to $\boldsymbol{\theta}$:

$$J_H \equiv \|\mathbf{H} - \mathbf{H}(\boldsymbol{\theta})\| = \| \left[\mathbf{I}_K - \mathbf{X}(\boldsymbol{\theta})\, \mathbf{X}^{+}(\boldsymbol{\theta})\right]\, \mathbf{H}\| \tag{5.34}$$

subject to the same constraints as in (5.33), where $\mathbf{H}(\boldsymbol{\theta})$ is the $K \times L$ RV-spectral matrix estimated from the dynamic model. Note that (5.34) minimizes the projection error of $\mathbf{H}$ on the column space of $\mathbf{X}$ by optimizing the reaction rate parameters $\boldsymbol{\theta}$. The solution to this optimization problem is $\hat{\boldsymbol{\theta}}$ with the corresponding estimate $\hat{\mathbf{X}} = \hat{\mathbf{X}}(\hat{\boldsymbol{\theta}})$. Thus, an estimate of $\mathbf{X}_r$ is obtained as $\hat{\mathbf{X}}_r = \hat{\mathbf{X}}$ with $K_r = K$.

Note that no theoretical advantage of one of the proposed approaches over the other can be stated.

Example 2.2f (Parameter estimation of a reversible reaction; cont'd from page 18**)**
The esterification reaction described in Example 2.2a ($R = 2$, $S = 4$, i.e., the catalyst $\mathcal{X}_5$ does not absorb) is conducted isothermally in semibatch mode with ethanol $\mathcal{X}_1$ being fed with the volumetric flowrate q_{in} shown in Figure 5.7a. The mixture initially placed in the reactor has concentrations $\mathbf{c}_0^{\star} = [2.90, 14.20, 0, 0]^{\mathrm{T}}$ M and volume $V_0 = 0.5$ l. The first spectrum $\mathbf{a}_0$ is taken at time t_0 after a 10-minute heating-up phase with corresponding concentrations $\mathbf{c}_0 = [2.65, 13.95, 0.25, 0.25]^{\mathrm{T}}$ M. Figure 5.9 depicts the pure-component spectra of all absorbing species ($L = 101$) which are used to construct $\mathbf{A}$ ($K = 100$). Gaussian noise with standard deviation 0.041 is added to $\mathbf{A}$. The evolution of the noisy $\tilde{\mathbf{A}}$ is illustrated in Figure 5.7b.
For the estimation of $\mathbf{C}$, it is assumed that (i) the reaction kinetics follow the power law, i.e., the reaction rate is given by $r = \kappa_1\, c_1\, c_2 - \kappa_2\, c_3\, c_4$, (ii) the density is constant, and

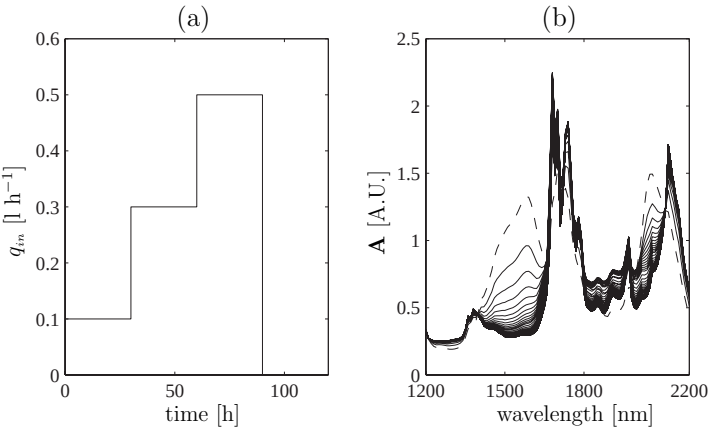

Figure 5.7. Example 2.2f: (a) The inlet volumetric flowrate q_{in}, and (b) the time-evolving spectra (first spectrum indicated by dashed line; every second spectrum shown).

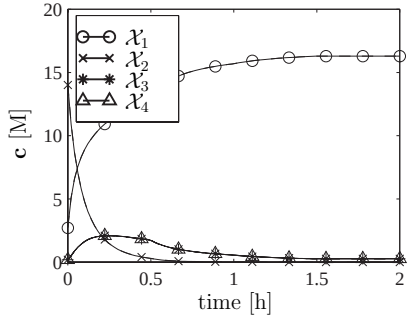

Figure 5.8. Example 2.2f: The true (–) and the estimated (– –) concentrations.

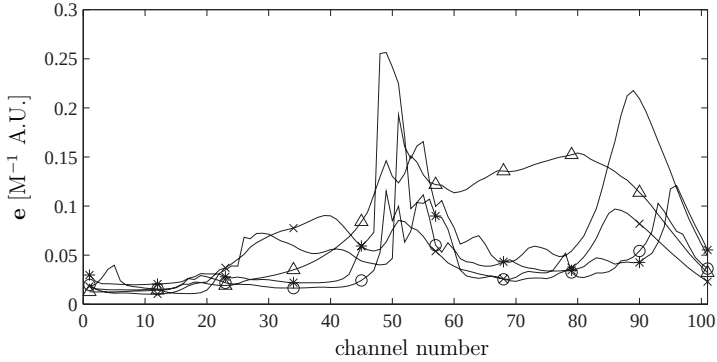

Figure 5.9. Pure-component spectra **E** for Example 2.2f. For legend, see Figure 5.8.

(iii) $\mathbf{c}_0^\star$ is known. In addition to the parameters κ_1 and κ_2, the initial concentration of water ($\mathcal{X}_4$) at time t_0 will be estimated ($\boldsymbol{\theta} := [\kappa_1, \kappa_2, c_{0,4}]^{\mathrm{T}}$). The initial concentrations of the remaining species are computed from (4.29) and (4.30), where the measured species correspond to the estimated water concentration ($\mathbf{z} = 0$, since the flowrate is assumed to be zero during the heating-up phase).
$\boldsymbol{\theta}$ is estimated by solving the nonlinear regression problem (5.33) with $\|\cdot\|$ being the Frobenius norm. Table 5.5 lists the simulated (true) values $\boldsymbol{\theta}_t$, the initial guesses $\boldsymbol{\theta}_0$, and the optimal estimates $\hat{\boldsymbol{\theta}}$. The difference between the true values and the optimal estimates can be attributed to noise. Figure 5.8 shows that the concentrations are correctly estimated (no visible difference).

5.3.2.4 Rotation — use of final conditions

This subsection discusses the estimation of $\acute{\mathbf{C}}$ (or $\mathbf{X}_r$) from certain conditions that result from the knowledge of the reaction type (reversible or irreversible).

Problem formulation

If explicit expressions for the reaction rates are not available, then so-called "model-free" FA methods must be used for the resolution of $\mathbf{C}$. Here, rotation for (i) a set of irreversible reactions with the presence of limiting species, and (ii) a set of reversible reactions is studied. It will be shown that for such reaction types, the rotational problem can be solved without the knowledge of explicit expressions for the kinetic structure. Saurina *et al.* (1998) and Izquierdo-Ridorsa *et al.* (1997) considered special cases of a set of irreversible first-order and reversible reactions, respectively. Furthermore, they used an iterative approach based on alternating regression (AR; see Appendix C.3.2.4).

Here, techniques will be proposed that estimate $\mathbf{C}$ for a *general* set of irreversible or reversible reactions. For irreversible reactions, the solution will be obtained *non-iteratively*. The proposed techniques rely on the estimation of the concentration or the reaction variants at those time instants for which (i) all the R independent irreversible reactions have disappeared due to the presence of limiting reactants, or (ii) all the R independent reversible reactions have reached equilibrium. The corresponding conditions are termed *final conditions*. Several experiments will be conducted for which the final concentrations (or reaction variants) are appended to form $\acute{\mathbf{C}}$ (or $\mathbf{X}_r$).

For reaction types i or ii, the reaction rates $r_i(t)$ converge asymptotically to 0 for $t \to \infty$ for all R independent reactions. In practice, due to noise, it is assumed that

Table 5.5. The true values $\boldsymbol{\theta}_t$, the initial guesses $\boldsymbol{\theta}_0$ for the nonlinear regression problem (5.33), and the estimates $\hat{\boldsymbol{\theta}}$ for Example 2.2f.

Parameter	$\boldsymbol{\theta}_t$	$\boldsymbol{\theta}_0$	$\hat{\boldsymbol{\theta}}$
κ_1 [M^{-1} h^{-1}]	0.0300	0.0800	0.0305
κ_2 [M^{-1} h^{-1}]	0.0100	0.0010	0.0129
$c_{4,0}$ [M]	0.1924	0	0.1306

the reaction rate $r_b(t)$ of Reaction b is negligible for $t \geq \acute{t}_b$, when $r_b(t)$ is less than or equal to a given limit $\underline{r}_b$:

$$r_b(t) \leq \underline{r}_b, \quad \forall t \in [\acute{t}_b, t_K], \tag{5.35}$$

where t_K corresponds to the time instant of the final observation K. The aim of Subsections 5.3.2.4 and 5.3.2.4 is to provide conditions and algorithms for the estimation of $\acute{\mathbf{C}}$ (or $\mathbf{X}_r$).

Irreversible reaction systems

Let Assumptions A1–12 in Appendix D be verified. Furthermore, it is assumed that (i) $\operatorname{rank}(\mathbf{X}_e) = R + p + 1$, and (ii) a set of R independent irreversible reactions is present (Knowledge K4). From the latter assumption follows that, in the absence of feeding after a certain time instant, if the consumption of a reactant $\mathcal{X}_j$ is greater than its production rate, the concentration c_j and the consumption rates of the corresponding reaction rates r_i become asymptotically zero:

$$\lim_{t\to\infty} c_j(t) = 0, \quad \lim_{c_j\to 0} r_i(c_1, \ldots, c_j, \ldots, c_S) = 0 \quad \text{for } i = \{s \,|\, j \in I^{rs}\}. \tag{5.36}$$

B experiments are conducted. Let R_b denote the subset of the R reactions that are present in the bth experiment and, similarly, S_b the subset of the S species. Let $\acute{t}_b$ denote the time instant for which all R_b reactions of the bth experiment have (asymptotically) terminated, and $\acute{\mathbf{c}}_b$ the corresponding (final) concentrations. The initial and inlet concentrations of the m experiments are appended to give $\mathbf{C}_x$.

Let $A_x \equiv \operatorname{rank}(\mathbf{C}_x)$ and $A \equiv \operatorname{rank}(\mathbf{A}_r)$. From $\operatorname{rank}(\mathbf{E}) = S$, $\operatorname{rank}(\mathbf{X}_e) = R+p+1$, and $\mathbf{N}_e = \mathbf{R}\,\mathbf{C}_r$, it follows that $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{N}_e) = A$. Thus, for the calibration pair $\{\mathbf{A}_r, \mathbf{C}_r\}$, $(A - A_x)$ final conditions are required that satisfy $\mathbf{N}_e = \mathbf{R}\,\mathbf{C}_r$. Alternatively, for the calibration pair $\{\mathbf{H}_r, \mathbf{D}_r\}$, R final conditions are required.

Next, a guideline for the choice of the calibration pair $\{\mathbf{A}_r, \mathbf{C}_r\}$ or $\{\mathbf{H}_r, \mathbf{D}_r\}$ is provided:

(a) If $A_x > A - R$, then calibration based on the pair $\{\mathbf{A}_r, \mathbf{C}_r\}$ is preferred, since only $(A - A_x) < R$ final conditions are required ($\acute{\mathbf{C}}$ with $\acute{K} = A - A_x$).

(b) If $A_x < A - R$, then calibration based on the pair $\{\mathbf{H}_r, \mathbf{D}_r\}$ is preferred, since only $R < (A - A_x)$ final conditions are required ($\mathbf{X}_r$ with $K_r = R$).

(c) If $A_x = A - R$, then calibration based on either $\{\mathbf{A}_r, \mathbf{C}_r\}$ or $\{\mathbf{H}_r, \mathbf{D}_r\}$ can be selected, since both calibration sets require R final conditions.

For the estimation of the final condition $\acute{\mathbf{c}}_b$ of the bth experiment, the following procedure is proposed:

(S1) Choose the experimental conditions in such a way that the R_b reactions (asymptotically) terminate at distinct time instants in the time interval $[t_{0,b}, t_{K,b}]$. This can be checked *a posteriori* by the following algorithm:

  (a) Determine R_b using (5.31): $R_b = \operatorname{rank}(\mathbf{H}_b)$, where $\mathbf{H}_b$ is the RV-spectral matrix for the bth experiment.

(b) Determine the termination instants of individual reactions by applying EFA backward in time on pre-treated submatrices of $\mathbf{H}_b$. The approximate-rank is estimated using the methods described in Appendix C.4.1.2.

(c) If R_b local approximate-rank reductions are observed, then the assumption of termination at distinct time instants is verified.

(P2) Termination of the reactions at distinct time instants guarantees that R_b reactants run out at distinct time instants. Thus, the R_b-dimensional concentration vector $\acute{\mathbf{c}}_{m,b}$ will be zero at the final instant $\acute{t}_b$: $\acute{\mathbf{c}}_{m,b} = \mathbf{0}_{R_b}$. Let assume, without loss of the generality, that the first R_b species are the limiting species. Then, the concentrations of the remaining $(S_b - R_b)$ species, $\acute{\mathbf{c}}_{u,b}$, are estimated from (4.29):

$$\hat{\acute{\mathbf{c}}}_{u,b} = \frac{\acute{\lambda}_b}{\acute{V}_b}\left[\mathbf{N}_u^{\mathrm{T}}\hat{\mathbf{x}}_b + \mathbf{C}_{in,u}^b\acute{\mathbf{z}}_b + V_0^b\mathbf{c}_{u,0}^b\right], \quad \hat{\mathbf{x}}_b = \mathbf{N}_m^{+\mathrm{T}}\left(-\mathbf{C}_{in,m}^b\acute{\mathbf{z}}_b - V_0^b\mathbf{c}_{m,0}^b\right), \tag{5.37}$$

where the subscripts b, m, and u denote quantities related to the bth experiment, the R_b reactants, and $(S_b - R_b)$ remaining species, respectively. Thus, $\acute{\mathbf{c}}_b = \begin{bmatrix}\acute{\mathbf{c}}_{m,b}\\ \acute{\mathbf{c}}_{u,b}\end{bmatrix}$.

The following remarks are in order:

- If a reactant is known to be responsible for the termination of two or more reactions in the B experiments, then its concentrations can be predicted for all K observations without explicit expressions for the kinetic structure (partial resolution). However, for the concentration estimation of the remaining species, explicit expressions for the kinetic structure are required.

- The choice of $\mathbf{C}_x$ is important to guarantee the space-inclusion condition in Proposition 5.21, since it is present in $\mathbf{C}_r$ and, furthermore, it indirectly influences $\acute{\mathbf{C}}$ and $\mathbf{X}_r$.

- The identity of the limiting reactant responsible for the termination of a *particular* reaction of an experiment may be unknown. However, the identity of the limiting reactants responsible for the termination of *all* the reactions of an experiment must be determined from chemical insight (Knowledge K5), since local (approximate-)rank reductions can have several reasons (for a non-exhaustive list, see Appendix G.1).

Example 2.4g (Limiting reactant $\mathcal{X}_2$ in irreversible reactions; cont'd from page 12**)**
Consider the irreversible parallel reaction system (2.8) ($R = 2$). Assume power-law kinetics, i.e., $r_1 = \kappa_1 c_1\, c_2$ and $r_2 = \kappa_2 c_1^2$. Furthermore, it is assumed that all four species absorb ($S = 4$).
Two batch runs are conducted ($B = 2$). The first batch run is conducted with $\mathbf{c}_{0,1} = [3,\ 0,\ 0,\ 0]^{\mathrm{T}}$ M and $V_{0,1} = 0.5$ l, where $\mathbf{c}_{0,1}$ was chosen such that only the second reaction ($2\,\mathcal{X}_1 \rightarrow \mathcal{X}_4$) is activated. The second run corresponds to that presented in Example 2.4f in Appendix G.1.2. Thus, $\mathbf{C}_x = \left[\begin{smallmatrix}3 & 0 & 0 & 0\\ 3 & 0.8 & 0 & 0\end{smallmatrix}\right]$. Figure 5.10 presents the concentrations and RV-concentrationsof the first batch run for $K = 200$ observations.

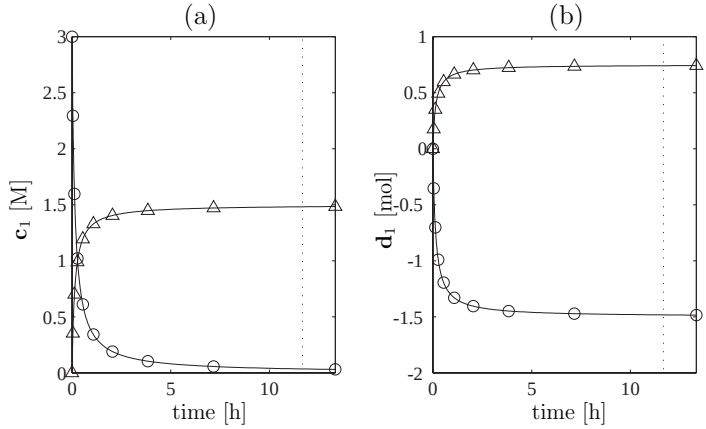

Figure 5.10. (a) The concentrations $\mathbf{c}_1$, and (b) the RV-concentrations$\mathbf{d}_1$ of the first batch run in Example 2.4g. See Figure 4.12 on page 86 for legend of the species.

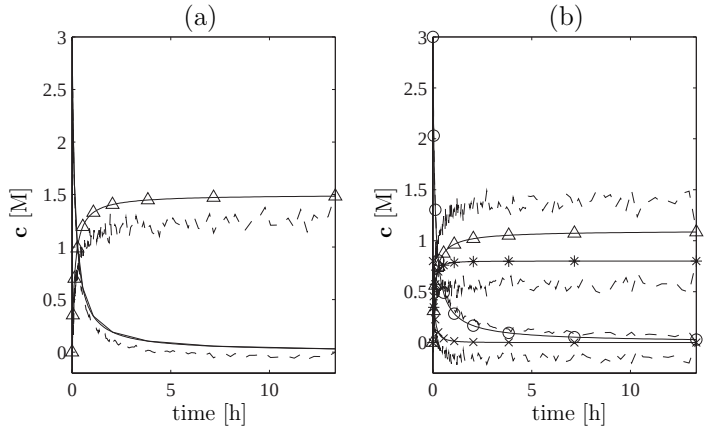

Figure 5.11. The concentrations of: (a) the first batch run, and (b) the second batch run in Example 2.4g. (–) true, (– –) estimated using $\{\mathbf{A}_r, \mathbf{C}_r\}$. See Figure 4.12 on page 86 for legend of the species.

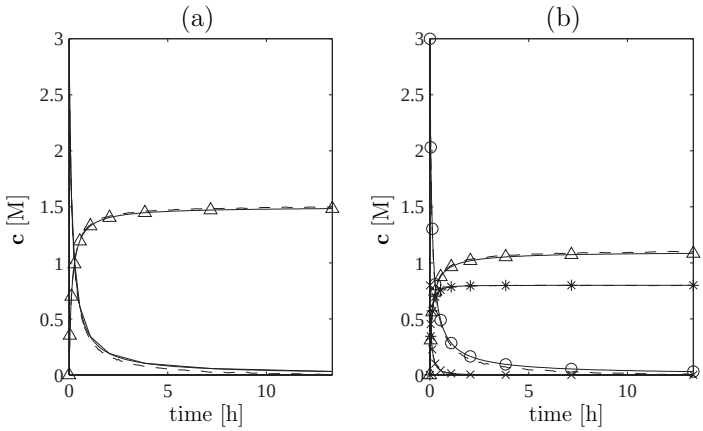

Figure 5.12. The concentrations of: (a) the first batch run, and (b) the second batch run in Example 2.4g. (–) true, (– –) estimated using $\{\mathbf{H}_r, \mathbf{D}_r\}$. See Figure 4.12 on page 86 for legend of the species.

Uniform noise in the range [-0.0039, 0.0039] is added to $\mathbf{A}_1$. The matrix $\tilde{\mathbf{A}}_2$ is noisy as well (see Example 2.4f. $\tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{A}_1 \\ \tilde{\mathbf{A}}_2 \end{bmatrix}$.

From $R = 2$, $A_x = \operatorname{rank}(\mathbf{C}_x) = 2$, and $A = \operatorname{rank}(\mathbf{A}_r) = 4$, it follows that $A_x = A - R = 2$. Thus, calibration based on either $\{\mathbf{A}_r, \mathbf{C}_r\}$ or $\{\mathbf{H}_r, \mathbf{D}_r\}$ can be selected, since both calibration sets require $R = 2$ final conditions.

From EFA backward in time applied to the column-mean-centered $\tilde{\mathbf{H}}_1$ (not shown), it can be concluded that the second reaction (the only reaction in this run) terminates after about 11.5 h ($R_1 = 1$). Furthermore, it was already shown in Example 2.4f that the two reactions in the second batch run terminate at distinct time instants ($R_2 = 2$). Thus, the concentrations of Species $\mathcal{X}_1$ and $\mathcal{X}_2$ at the final observations of both batch runs are known to be zero. Therefore, $\mathbf{X}_r$ can be estimated with

$$\mathbf{D}_{r,m} = \begin{bmatrix} -V_0^1 c_{1,0}^1 & 0 \\ -V_0^2 c_{1,0}^2 & -V_0^2 c_{2,0}^2 \end{bmatrix} \text{ and } \mathbf{N}_m = \begin{bmatrix} -1 & -1 \\ -2 & 0 \end{bmatrix}$$
$$\text{as } \hat{\mathbf{X}}_r = \mathbf{D}_{r,m}\,\mathbf{N}_m^+ = \begin{bmatrix} 0 & 0.5\,V_0^1 c_{1,0}^1 \\ 0.5\,V_0^2 c_{2,0}^2 & 0.5\,V_0^2\,(c_{1,0}^2 - c_{2,0}^2) \end{bmatrix}. \tag{5.38}$$

$\mathbf{C}_r$ is estimated as $\hat{\mathbf{C}}_r = \begin{bmatrix} \mathbf{C}_{x,m} & \mathbf{C}_{x,u} \\ \mathbf{C}_m & \hat{\mathbf{C}}_u \end{bmatrix}$, where $\mathbf{C}_m = \mathbf{0}_{2\times 2}$ and $\mathbf{C}_u$ estimated from (5.37). $\hat{\mathbf{C}}$ predicted from the calibration models based on $\{\mathbf{A}_r, \mathbf{C}_r\}$ and $\{\mathbf{H}_r, \mathbf{D}_r\}$ are shown in Figures 5.11 and 5.12, respectively. In contrast to the prediction from the calibration model based on $\{\mathbf{A}_r, \mathbf{C}_r\}$, the concentrations are correctly predicted from the calibration model based on $\{\mathbf{H}_r, \mathbf{D}_r\}$. The reason for this difference was already stated in Subsection 5.3.1.5: Since the terms related to the reaction-invariant part are known perfectly, the noise is restricted to a lower-dimensional space (here, 2-dimensional), which results in much better performance with $\{\mathbf{H}_r, \mathbf{D}_r\}$.

Figure 5.12 shows that the estimated concentrations for Species $\mathcal{X}_1$ exhibit a small offset from the true value. This offset is due to the incorrect assumption in $\mathbf{C}_r$ that

$\mathcal{X}_1$ is not completely consumed at the end of the batch runs and, thus, the assumption that the concentration is perfectly 0 is not verified.
In summary, for the resolution of the concentrations of all four species, $B = 2$ experimental runs were required. Furthermore, by appropriate choice of the experimental concentration matrix $\mathbf{C}_x$, it was possible to activate only the second reaction in the first batch run. This way, $\text{rank}\,(\mathbf{C}_r) = S$ and $\text{rank}\,(\mathbf{X}_r) = R$ are satisfied and, thus, also the space-inclusion conditions.

Reversible reaction systems

Let Assumptions A1–12 in Appendix D be verified. Furthermore, it is assumed that a set of R independent reversible reactions is present (Knowledge K4). Then, in the absence of feeding after a certain time instant, reversible Reaction b reaches equilibrium, if the production and consumption rates (summarized in r_b) corresponding to this reaction become (nearly) equal or, equivalently, r_b becomes negligible. An approach based on the calibration pair $\{\mathbf{H}_r, \mathbf{D}_r\}$ will be proposed next.

Let B be the number of experiments conducted. Let $\mathbf{H}_r$ $(B \times L)$ and $\mathbf{X}_r$ $(B \times R)$ denote the matrices containing the L-dimensional RV-spectral vectors, $\acute{\mathbf{h}}_b$, and the R-dimensional reaction-variant vector, $\acute{\mathbf{x}}_b$, for each experiment b after all reactions have (asymptotically) reached equilibrium at $k \geq \acute{k}_b$. Let $\acute{\mathbf{c}}_b$ denote the S-dimensional concentration vector of the bth experiment at equilibrium, and $\mathbf{H}_b$ the RV-spectral matrix of the bth experiment.

Let Assumptions A1–12 in Appendix D be verified and $\mathbf{N}$ be known. Furthermore, it is assumed that the structure of the mass action laws (2.60) of all reactions are known but with unknown equilibrium constants $\boldsymbol{\mathcal{K}}$.

A procedure that is based on the two-step FA-approach is proposed for the estimation of $\mathbf{X}_r$. In the first step of FA (see also Appendix C.3), the abstract matrices $\mathbf{X}_a$ $(K \times R)$ and $\boldsymbol{\Theta}_a$ $(R \times L)$ of $\mathbf{X}$ and $\boldsymbol{\Theta}$, respectively, are computed from $\mathbf{H}$ using PCA:

$$\mathbf{H} = \mathbf{X}_a \, \boldsymbol{\Theta}_a. \tag{5.39}$$

In the second step of FA, an $R \times R$ rotation matrix $\mathbf{R}$ is estimated so as to estimate the physically-meaningful quantities $\mathbf{X}$ and $\boldsymbol{\Theta} = \mathbf{N}\,\mathbf{E}$:

$$\hat{\mathbf{X}} = \mathbf{X}_a \mathbf{R}, \qquad \hat{\boldsymbol{\Theta}} = \mathbf{R}^{-1} \boldsymbol{\Theta}_a, \tag{5.40}$$

where $\hat{\mathbf{X}}$ and $\hat{\boldsymbol{\Theta}}$ are the estimates of $\mathbf{X}$ and $\boldsymbol{\Theta}$, respectively.

The following procedure for the estimation of $\mathbf{X}_r$ is proposed:

(S1) Conduct $B \geq (R + 1)$ experiments where the reaction invariants of each experiment are so chosen that (i) all reactions are in equilibrium at the end of the experiment, and (ii) $\text{rank}\,(\mathbf{H}_r) = R$. The first condition can be checked e.g. by computing the approximate-rank of a column-mean-centered submatrix of $\mathbf{H}_b$ for the R last observations[2].

(S2) From (2.60):

$$\mathsf{k}_i(\acute{\mathbf{c}}_b) = \mathsf{K}_i, \qquad \forall i = 1, \ldots, R, \quad b = 1, \ldots, B. \tag{5.41}$$

[2] All reactions are in equilibrium when the approximate-rank is zero.

Since K_i is constant ($i = 1, \ldots, R$), it is the same for all experiments. Thus, the following set of $R\,(B-1)$ equations can be derived:

$$\mathsf{k}_j(\acute{\mathbf{c}}_1) = \mathsf{k}_j(\acute{\mathbf{c}}_b), \qquad \forall j = 1, \ldots, R, \quad i = 2, \ldots, B. \tag{5.42}$$

(S3) Compose $\mathbf{H}$ by appending all $\mathbf{H}_b$. Compute the abstract space $\mathbf{X}_a$ from (5.39). Select from $\mathbf{X}_a$ those rows, $\hat{\acute{\mathbf{x}}}_{a,b}$, that correspond to $\acute{k}_b$ ($b = 1, \ldots, B$). Thus, from (5.40),

$$\hat{\acute{\mathbf{x}}}_b^{\mathrm{T}} = \acute{\mathbf{x}}_{a,b}^{\mathrm{T}}\,\mathbf{R}, \qquad i = 1, \ldots, B. \tag{5.43}$$

Express $\acute{\mathbf{c}}_b$ in (5.42) as a function of the (unknown) rotation matrix $\mathbf{R}$ using (4.3):

$$\mathsf{k}_j(\acute{\mathbf{x}}_{a,1}^{\mathrm{T}}\,\mathbf{R}) = \mathsf{k}_j(\acute{\mathbf{x}}_{a,b}^{\mathrm{T}}\,\mathbf{R}), \qquad \forall i = 1, \ldots, R, \quad b = 2, \ldots, B. \tag{5.44}$$

(S4) Solve (5.44) for $\mathbf{R}$, leading to an estimate of $\mathbf{X}_r$.

Note that at least $(R+1)$ experiments must be conducted, since all R^2 elements of $\mathbf{R}$ are unknown and $R\,(B-1)$ equations are available:

$$R\,(B-1) \geq R^2 \quad \Rightarrow \quad B \geq R+1. \tag{5.45}$$

Furthermore, in Subsection 5.3.2.4, it was required that the irreversible reactions terminate at distinct time instants. Here, the reversible reactions are *not* required to reach equilibrium at distinct time instants.

(5.42) represents a set of nonlinear algebraic equations that may contain multiple solutions (Press *et al.*, 1994). In such a case, prior knowledge such as non-negativity of the estimated concentrations $\hat{\mathbf{C}}$ (Knowledge K1) may be needed to help select the physically-meaningful solution.

An alternative approach is the following: By performing B_1 experiments in addition to the minimum required number $(R+1)$, $\binom{B}{R+1}$ sets of algebraic equations (5.42) can be generated, with $B = (R+1+B_1)$, and each has a solution set. The intersection of all the solution sets contains the true physical solution. It is hoped that, by this approach, the intersection set becomes smaller with increasing B and, thus, selection of the true physical solution becomes easier.

Example 2.2h (Esterification; cont'd from page 18**)**
It is assumed that all reacting species absorb except for the sulfuric acid and, thus, $S = 4$ ($R = 1$). Two batch runs are conducted with $\mathbf{c}_{0,1} = [3,\, 14,\, 0,\, 0]^{\mathrm{T}}$ M and $V_{0,1} = 0.5$ l and $\mathbf{c}_{0,2} = [14,\, 2,\, 0,\, 0]^{\mathrm{T}}$ M and $V_{0,2} = 0.8$ l, respectively. Figure 5.13 presents the concentrations of both batch runs for $K = 100$ observations (solid curves). Gaussian noise is added to $\mathbf{A}_1$ and $\mathbf{A}_2$ with standard deviations 0.038 and 0.042, respectively.
Following the procedure presented above:

(a) $B = (R+1) = 2$ experiments are conducted. From the approximate-rank of the column-mean centered submatrix of $\tilde{\mathbf{H}}_b$ for the 8 last observations ($b = 1, 2$), it is confirmed that equilibrium is reached at the end of each experiment: $\acute{t}_1 = \acute{t}_2 = 2$ h. The final RV-spectral vectors $\hat{\acute{\mathbf{h}}}_b$ of each experiment is selected from $\tilde{\mathbf{H}}_b$, and appended to $\mathbf{H}_r$. The condition $\mathrm{prank}\,(\mathbf{H}_r) = R$ is verified.

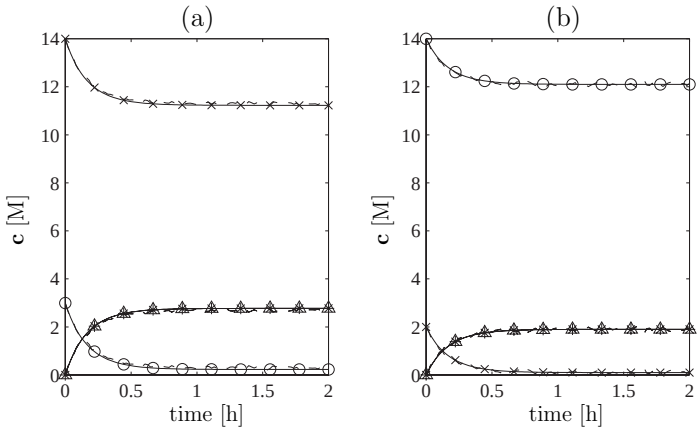

Figure 5.13. The concentrations of: (a) the first batch run, and (b) the second batch run in Example 2.2h. (–) true, (– –) estimated. See Figure 5.8 on page 125 for legend of the species.

(b) The mass-action law for the esterification reaction at equilibrium reads:

$$\mathsf{K} = \frac{\acute{c}_{3,b}\,\acute{c}_{4,b}}{\acute{c}_{1,b}\,\acute{c}_{2,b}}, \qquad b = 1, 2, \tag{5.46}$$

where $\acute{k}_b = 100$ for $b = 1, 2$. Since the (unknown) value of K is constant, the following equation ($R\,(B-1) = 1$) can be derived:

$$\frac{\acute{c}_{3,1}\,\acute{c}_{4,1}}{\acute{c}_{1,1}\,\acute{c}_{2,1}} = \frac{\acute{c}_{3,2}\,\acute{c}_{4,2}}{\acute{c}_{1,2}\,\acute{c}_{2,2}}. \tag{5.47}$$

(c) $\tilde{\mathbf{H}} = \begin{bmatrix} \tilde{\mathbf{H}}_1 \\ \tilde{\mathbf{H}}_2 \end{bmatrix}$. The 200×1 matrix $\mathbf{X}_a$ is obtained from PCA of $\tilde{\mathbf{H}}$ by retaining the first factor ($R = 1$). The $100th$ and $200th$ element (row) of $\mathbf{X}_a$ is selected where the first and second batch are (approximately) at equilibrium, respectively: $\acute{x}_{a,1} = x_a^{100}$, $\acute{x}_{a,2} = x_a^{200}$.
Since $R = 1$, the rotation matrix $\mathbf{R} = \mathcal{R}$ is a scalar. Thus, with $\hat{\acute{x}}_b = \mathcal{R}\,\acute{x,}b$ ($b = 1, 2$),

$$\hat{\mathbf{c}}_b = \mathbf{N}^{\mathrm{T}}\hat{\acute{x}}_b/V_{0,b} + \mathbf{c}_{0,b} = \mathcal{R}\,\mathbf{N}^{\mathrm{T}}\acute{x,}b/V_{0,b} + \mathbf{c}_{0,b}, \quad b = 1, 2. \tag{5.48}$$

Substituting (5.48) in (5.47), an equation for the unknown rotation scalar $\mathcal{R}$ is obtained.

(d) The initial estimate of $\mathcal{R}$ is set to 5. (5.47) is solved for $\mathcal{R}$ under non-negativity constraints for the estimated concentrations. $\hat{\mathcal{R}} = 1.21$ and $\hat{\mathbf{X}}_r = [1.309,\ 1.490]^{\mathrm{T}}$ are obtained, which are in good agreement with the true values $\mathcal{R} = 1.17$ and $\mathbf{X}_r = [1.39,\ 1.52]^{\mathrm{T}}$, respectively.

The concentration profiles of both batch runs are estimated correctly as shown in Figure 5.13 (dotted lines).

5.4 Extensions

In Section 5.1, the factorization of spectral data (5.3) and (5.5) were derived under the assumption that all species both react and absorb. In the following two subsections, this assumption will be relaxed, and the presence of non-absorbing reacting species and non-reacting absorbing species will be considered. Furthermore, the assumption of linearly-independent pure-component spectra will be relaxed, and the factorization of spectral data extended by temperature/calorimetric measurements will be studied.

5.4.1 Non-absorbing reacting species

There are two main reasons why not all reacting species may absorb:

(R1) *By physics:* Some reacting species do not absorb in the spectral regions of interest. This happens, for example, for solid-liquid reactions, where the spectrum of only those reacting species are measured that are in the liquid phase (using e.g. an attenuated-total-reflection probe, ATR). An example for solid-liquid reactions is a biotechnological process involving a *Saccharomyces cerevisiae* culture (Cannizzaro, 1998). This process was followed by an ATR FTIR probe. The non-absorbing reacting species is the (solid) biomass, while ethanol, glucose, and sugars are the species that both react and absorb.

(R2) *By purpose:* In calibration or factor analysis, it is sometimes useful to select spectral regions of interest using the channel selection method proposed in Subsection 5.2.2.3.

Let S_r be the total number of reacting species, S_{ra} the number of reacting and absorbing species, and S_{rn} the number of non-absorbing reacting species with $S_r \equiv S_{ra} + S_{rn}$. Furthermore, let R be the *total* number of independent reactions.

The number of absorbing species S is redefined as $S := S_{ra}$, and the $S \times L$ pure-component matrix as

$$\mathbf{E} := \mathbf{E}_{ra}, \tag{5.49}$$

where the subscript $(\cdot)_{ra}$ denotes quantities related to the S_{ra} species. Then, the $K \times S_r$ matrices $\mathbf{D}_r$ and $\mathbf{C}_r$ corresponding to the S_r reacting species can be partitioned as follows:

$$\mathbf{D}_r = \left[\begin{array}{c|c} \mathbf{D}_{ra} & \mathbf{D}_{rn} \end{array}\right], \qquad \mathbf{C}_r = \left[\begin{array}{c|c} \mathbf{C}_{ra} & \mathbf{C}_{rn} \end{array}\right], \tag{5.50}$$

where the subscript rn denotes quantities related to the S_{rn} species. Such a partitioning implies for $\mathbf{N}_r$ $(R \times S)$ and $\mathbf{N}_{e,r}$ $(R + p + 1 \times S)$:

$$\mathbf{N}_r = \left[\begin{array}{c|c} \mathbf{N}_{ra} & \mathbf{N}_{rn} \end{array}\right], \qquad \mathbf{N}_{e,r} = \left[\begin{array}{c|c} \mathbf{N}_{e,ra} & \mathbf{N}_{e,rn} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{N}_{ra} & \mathbf{N}_{rn} \\ \mathbf{C}_{x,ra} & \mathbf{C}_{x,rn} \end{array}\right]. \tag{5.51}$$

The (redefined) $K \times S$ concentration matrices $\mathbf{D}$ and $\mathbf{C}$ corresponding to the S absorbing species are related to $\mathbf{D}_r$ and $\mathbf{C}_r$ through an $S_r \times S$ matrix $\boldsymbol{\mathcal{P}}_{ra}$.

$$\mathbf{D} := \mathbf{D}_r \boldsymbol{\mathcal{P}}_{ra} = \mathbf{D}_{ra}, \quad \mathbf{C} := \mathbf{C}_r \boldsymbol{\mathcal{P}}_{ra} = \mathbf{C}_{ra}, \qquad \boldsymbol{\mathcal{P}}_{ra} \equiv \left[\begin{array}{c} \mathbf{I}_{S_{ra}} \\ \mathbf{0}_{S_{rn} \times S_{ra}} \end{array}\right]. \tag{5.52}$$

Similarly,

$$\mathbf{N} := \mathbf{N}_r \boldsymbol{\mathcal{P}}_{ra} = \mathbf{N}_{ra}, \quad \mathbf{N}_e := \mathbf{N}_{e,r} \boldsymbol{\mathcal{P}}_{ra} = \mathbf{N}_{e,ra} = \begin{bmatrix} \mathbf{N}_{ra} \\ \mathbf{C}_{x,ra} \end{bmatrix}, \tag{5.53}$$

where $\mathbf{N}$ and $\mathbf{N}_e$ are matrices of dimension $R \times S$ and $(R + p + 1) \times S$, respectively. $\mathbf{X}$ and $\mathbf{X}_e$ are defined as usual.

Since non-absorbing reacting species correspond to the case of unmeasured species in Subsection 4.1.4, all the results in Subsection 4.2 corresponding to the rank of $\mathbf{C}_m$ and $\mathbf{D}_m$ apply directly to the concentration matrices $\mathbf{C}_{ra}$ and $\mathbf{D}_{ra}$ corresponding to $\mathbf{A}$ and $\mathbf{H}$ with the following redefinitions, respectively: $\mathbf{C}_m := \mathbf{C}_{ra}$, $\mathbf{D}_m := \mathbf{D}_{ra}$, and $S_m := S_{ra}$.

5.4.2 Non-reacting absorbing species

In many practical applications, reactions are catalyzed and/or conducted in a mixture of solvents. Catalysts and solvents may absorb in the spectral regions of interest. In such cases, they represent non-reacting absorbing species.

Let S_{na} be the number of non-reacting absorbing species. The number of absorbing species S is redefined as $S := S_r + S_{na}$ ($S_{ra} = S_r$), and the $S \times L$ pure-component matrix as

$$\mathbf{E} := \left[\begin{array}{c} \mathbf{E} \\ \hline \mathbf{E}_{na} \end{array}\right], \tag{5.54}$$

where the subscript na denotes quantities related to the S_{na} species.

Since the S_{na} non-reacting species do not react by definition,

$$\mathbf{N} := \left[\begin{array}{c|c} \mathbf{N} & \mathbf{0}_{R \times S_{na}} \end{array}\right], \tag{5.55}$$

where the stoichiometric matrix $\mathbf{N}$ is augmented by a $R \times S_{na}$ zero-matrix. Thus, the $K \times S$ RV-concentration matrix $\mathbf{D}$ and the concentration matrix $\mathbf{C}$ corresponding to the S absorbing species are redefined as

$$\mathbf{D} := \left[\begin{array}{c|c} \mathbf{D} & \mathbf{0}_{K \times S_{na}} \end{array}\right], \qquad \mathbf{C} := \left[\begin{array}{c|c} \mathbf{C} & \mathbf{C}_{na} \end{array}\right], \tag{5.56}$$

with

$$\mathbf{N}_e := \left[\begin{array}{c|c} \mathbf{N} & \mathbf{0}_{R \times S_{na}} \\ \hline \mathbf{C}_x & \mathbf{C}_{x,na} \end{array}\right], \quad \mathbf{C}_x := \left[\begin{array}{c|c} \mathbf{C}_x & \mathbf{C}_{x,na} \end{array}\right], \tag{5.57}$$

where the redefined $(p + 1) \times S$ experimental concentration matrix $\mathbf{C}_x$ contains the experimental concentrations of both the S_r reacting absorbing and the S_{na} non-reacting absorbing species. $\mathbf{X}$ and $\mathbf{X}_e$ are defined as usual.

5.4.3 Linearly-dependent pure-component spectra

In the spectral regions of interest, the pure-component spectra of some of the species might be similar or identical to a linear combination of the remaining species. Thus, some of the species have linearly-dependent pure-component spectra. This phenomenon is usually known under *spectral overlap* (Tauler *et al.*, 1995).

Let S_{ta} be the total number of absorbing species, and S_d the number of absorbing species for which the pure-component spectra are linearly dependent on those of the remaining S_{id} absorbing species with $S_{ta} = S_{id} + S_d$. S is redefined as the *number of independently absorbing species* $S := S_{id}$, and the $S \times L$ pure-component matrix as

$$\mathbf{E} := \mathbf{E}_{id}, \tag{5.58}$$

where the subscript id denotes quantities related to the S_{id} species. Assuming that the first S_{id} rows of the $S_{ta} \times L$ instrumental matrix, $\mathbf{E}_{ta}$, correspond to the pure-component spectra of the S_{id} species, it follows:

$$\mathbf{E}_{ta} \equiv \left[\frac{\mathbf{E}_{id}}{\mathbf{E}_d}\right] = \boldsymbol{\mathcal{P}}_E \mathbf{E}, \qquad \boldsymbol{\mathcal{P}}_E \equiv \begin{bmatrix} \mathbf{I}_{S_{id}} \\ \boldsymbol{\Gamma}_E \end{bmatrix}, \tag{5.59}$$

where $\boldsymbol{\mathcal{P}}_E$ is an $S_{ta} \times S$ matrix, and the subscripts ta and d denote quantities related to the S_{ta} and S_d species, respectively.

Since $\mathbf{H} = \mathbf{D}_{ta}\mathbf{E}_{ta} = (\mathbf{D}_{ta}\boldsymbol{\mathcal{P}}_E)\mathbf{E}$ and $\mathbf{A} = \mathbf{C}_{ta}\mathbf{E}_{ta} = (\mathbf{C}_{ta}\boldsymbol{\mathcal{P}}_E)\mathbf{E}$,

$$\begin{aligned} &\mathbf{D} := \mathbf{D}_{ta}\boldsymbol{\mathcal{P}}_E, \qquad &&\mathbf{N} := \mathbf{N}_{ta}\boldsymbol{\mathcal{P}}_E, \\ &\mathbf{C} := \mathbf{C}_{ta}\boldsymbol{\mathcal{P}}_E, \qquad &&\mathbf{N}_e := \mathbf{N}_{e,ta}\boldsymbol{\mathcal{P}}_E, \qquad \mathbf{C}_x := \mathbf{C}_{x,ta}\boldsymbol{\mathcal{P}}_E. \end{aligned} \tag{5.60}$$

Comparing the redefinitions of $\mathbf{D}$ (or $\mathbf{C}$) in (5.60) with that in (5.52), it can be seen that $\mathbf{D}_{ta}$ and $\mathbf{D}_r$ (or $\mathbf{C}_{ta}$ and $\mathbf{C}_r$) are multiplied from the right by the matrices $\boldsymbol{\mathcal{P}}_E$ and $\boldsymbol{\mathcal{P}}_{ra}$, respectively. Thus, also in the case of linearly-dependent pure-component spectra, the number of independent reactions observed by $\mathbf{D}$ (or $\mathbf{C}$) may be less than the total number of independent reactions. Therefore, all the results in Subsection 4.2.1 corresponding to the rank of $\mathbf{C}_m$ and $\mathbf{D}_m$ apply directly to the concentration matrices $\mathbf{C}_{ra}$ and $\mathbf{D}_{ra}$ corresponding to $\mathbf{A}$ and $\mathbf{H}$ with the following redefinitions, respectively: $\mathbf{C}_m := \mathbf{C}_{ta}\boldsymbol{\mathcal{P}}_E$, $\mathbf{D}_m := \mathbf{D}_{ta}\boldsymbol{\mathcal{P}}_E$, and $S_m := S_{id}$.

5.4.4 Factorization of extended spectral data

Based on the factorization of extended concentration data derived in Subsection 4.4.5, the factorization of spectral data encompassing temperature/calorimetric measurements are derived. It is assumed that $\mathbf{E}$ is of full rank S and independent of temperature.

Appending the additional data vectors $\boldsymbol{\mathcal{T}}^\dagger$ and $\mathbf{d}_T$ to the spectral and RV-spectral matrices, respectively, leads to:

$$\mathbf{A}_T \equiv \left[\begin{array}{c|c} \mathbf{A} & \boldsymbol{\mathcal{T}}^\dagger \end{array}\right] = \mathbf{X}_{e,T}\,\mathbf{N}_{e,T}\,\mathbf{E}_T, \tag{5.61}$$

$$\mathbf{H}_T \equiv \left[\begin{array}{c|c} \mathbf{H} & \mathbf{d}_T \end{array}\right] = \mathbf{X}\,\mathbf{N}_T\,\mathbf{E}_T, \tag{5.62}$$

where $\mathbf{A}_T$ and $\mathbf{H}_T$ are $K \times (L+1)$ spectral matrices, $\mathbf{X}_{e,T}$ and $\mathbf{N}_{e,T}$ as defined in (4.45), and $\mathbf{N}_T$ as defined in Subsection 2.3.4, and $\mathbf{E}_T$ the $(S+1) \times (L+1)$ pure-component spectra matrix defined as

$$\mathbf{E}_T \equiv \left[\begin{array}{c|c} \mathbf{E} & \mathbf{0}_S \\ \mathbf{0}_L^{\mathrm{T}} & 1 \end{array}\right]. \tag{5.63}$$

(5.63) shows that the $(S+1)$st pure-component spectrum is modeled as a species that has a unity absorbance at a single channel number distinct from the L channels of the remaining S absorbing species.

5.5 Summary

Spectral data from reacting mixtures were studied from the perspective of the factorization of concentration data developed in the previous chapter. Based on this insight, new results for both calibration and FA for the purpose of estimating/predicting concentrations from spectral data were obtained.

The rotation step in FA was formulated as a calibration problem. To account for the lack of calibration concentration measurements in FA, explicit or implicit knowledge about the kinetic structure was used.

The goals of calibration and stoichiometric modeling using TFA can now be compared. In calibration, the objective is to collect calibration data in which new data can *fit*, while in TFA, the objective is to collect data that provide the best *discrimination* power for proposed targets. This leads to nullity-change operations in opposite directions: in calibration, it is aimed at *reducing* the nullity of the spectral calibration data (or of the underlying concentration matrix). In TFA, however, it is aimed at *increasing* the nullity of the concentration data.

As far as data pre-treatment to reaction-variant form is concerned, a calibration model based on pre-treated data *predicts* the reaction variants from a new spectrum, while in TFA, targets can be discriminated directly on the pre-treated data with the goal of *identifying* the reaction-variant space.

The results obtained under the assumption that all species both react and absorb were extended to the situations characterized by the presence of non-absorbing reacting species, non-reacting absorbing species, and linearly-independent pure-component spectra. Furthermore, the factorization of spectral data were extended to include temperature/calorimetric measurements (extended spectral data). This way, the results obtained for extended concentration data are directly applicable to extended spectral data.

6

Final remarks and outlook

6.1 Main theoretical results and their significance

Motivated by the industrial needs to estimate concentrations and identify reaction pathways without the knowledge of kinetics, a nonlinear transformation was proposed to bring the dynamic model of homogeneous reaction systems to normal form (*dynamic model in normal form*). As a result, the concept of reaction and flow variants/invariants enables the segregation of the state evolution (concentrations, volume, temperature) into three parts: (i) the part related to the reactions (reaction variants), (ii) the part related to flows such as inlet and outlet streams and external heat flow (reaction invariants and flow variants), and (iii) the part related to the initial conditions (reaction and flow invariants).

As a direct consequence of this segregation, novel factorizations for measured concentration, temperature, and spectral data were derived. Data pre-treatment to reaction-variant form was proposed to remove the direct contribution of the (usually known) initial conditions and inlet streams (i.e., the reaction-invariant part) from measured concentration and spectral data. This way, the reaction-variant part, which is related to the underlying reactions to be studied, can be extracted (*factorizations of data in reaction-variant form*). For spectral data, the proposed data pre-treatment is generic, since it does not involve the knowledge of the (unknown) concentrations.

Table 6.1 summarizes the practical problems addressed in this dissertation, the tools used to solve them, and the corresponding practical implications. Also, the benefits to the chemist/engineer in the laboratory (academia, vendors, industry) and production (industry) are provided.

The main results of Chapters 3–5 are now analyzed with respect to their practical importance.

Transformation to normal form

The concept of reaction variants/invariants has been around for nearly 25 years. A *linear* transformation enables the separation of reaction variants and invariants. When the latter remain constant, they can be eliminated from the dynamic model, and only R differential equations remain to be integrated. Unfortunately, this straightforward situation holds only for batch reactors, since the presence of inlet and outlet streams affects the reaction invariants. Thus, the system evolves due to the reactions and flows

in a larger space. A major contribution of this work has been to introduce a *nonlinear* transformation that enables to single out the reaction *and* flow variants, i.e., those terms that remain constant and can be dropped from the dynamic model. It is this transformation that enables the analysis of concentration and spectral data (discussed next) through appropriate factorizations.

Factorization of concentration data

The analysis of a reaction system (investigation of reaction network and kinetics) is often based on measured concentration that, typically, depend on both reactions *and* flows. Fortunately, these dependencies can be separated using the transformation to normal form, and the following very important linear relationship results, $\mathbf{C} = \mathbf{X}_e\mathbf{N}_e$. This way, the measured concentration data is factorized into $\mathbf{X}_e$ (which contains *dynamic* information related to the reaction variants, the reaction invariants and flow variants, and the reaction and flow invariants) and $\mathbf{N}_e$ (the extended stoichiometric matrix that contains only structural information, $\mathbf{N}$,$\mathbf{C}_{in}$ and $\mathbf{c}_0$).

When the reaction variants are known, which is often the case, the concentration data can be pre-treated to reaction-variant form for analysis, $\mathbf{D} = \mathbf{X}\,\mathbf{N}$, where $\mathbf{X}$ is the matrix of reaction variants. Thus, it becomes straightforward to identify reaction stoichiometries using TFA.

If pre-treatment to reaction-variant form cannot be obtained (e.g., because the flowrates are unknown), then one must work with the measured concentrations $\mathbf{C}$ directly. The applicability of TFA is a bit more tricky, and only necessary conditions for the acceptance of proposed targets have been available. A side but important result of this work has been the generation of necessary *and* sufficient conditions for the acceptance of targets using some reaction-invariant relationships. Furthermore, the important case of unmeasured species (i.e., when not all the reacting species are measured) has been solved satisfactorily.

Factorization of spectral data

The estimation of concentrations from spectral measurements has become increasingly important in industry. This work has brought the natural factorization of spectral data, $\mathbf{A} = \mathbf{C}\,\mathbf{E}$, to the reaction-variant form $\mathbf{H} = \mathbf{X}\,\mathbf{E}$, which is ideally suited for the analysis of reaction systems.

Most of the published work regarding the estimation of concentrations is limited to spectral data matrices of rank S (full-rank). However, in the case of reacting mixtures, the spectral data matrix is very often rank deficient, and the available results do not apply. This work has recognized this problem and proposed useful techniques for solving it (e.g., experimental rank augmentation, channel selection).

For the estimation of concentrations from spectral measurements, multivariate calibration is the method of choice. It, however, requires that a new spectrum lies in the space spanned by the calibration spectral data (space-inclusion condition). This work proposed *a priori* conditions that must be fulfilled by the calibration data in order to meet the space-inclusion requirement. In reaction-variant form, these conditions reduce very simply e.g. to the requirement that the reaction system is the same for both the calibration and the new runs.

Often in practice, there are too few or even no measurements for calibration purposes. Multivariate curve-resolution techniques can then be used to compute $\mathbf{C}$ from $\mathbf{A}$ (e.g., FA techniques and additional information to resolve the rotational and intensity ambiguities that are typical of FA). However, this approach requires rank $(\mathbf{A}) = S$, which often does not hold for reacting mixtures. For this particular situation, this work proposed to use partial kinetic information to resolve the ambiguities. For instance, given the measured (rank-deficient) spectral matrix $\mathbf{A}$, the stoichiometry and the kinetic model with the *unknown* parameters $\boldsymbol{\theta}$, an optimization approach can solve for $\boldsymbol{\theta}$ and, consequently, also for $\mathbf{C}$. When explicit expressions for the reaction rates are not available, the ambiguities can be solved using implicit knowledge of the type of reactions (e.g., all reactions reversible or irreversible). The proposed techniques rely on the estimation of the reaction variants at those time instants for which all the reversible reactions have reached equilibrium, or all the irreversible reactions have disappeared due to the presence of limiting reactants. A guideline for an experimental plan for several experiments was proposed that must be conducted and for which the final conditions are used to compute the reaction variants.

6.2 Experimental results

Various tools were successfully tested in the laboratories of Lonza Ltd., Visp, Switzerland, and the Chemical Department of the Engineering School of Wallis (ISW) in Sitten, Switzerland. Amrhein *et al.* (1998) illustrated the versatility of the proposed calibration methods by the calibration of water concentration with absorbance data (near-infrared (NIR) and mid-infrared (MIR)) for the esterification reaction described by (2.6) and conducted in the laboratories of Lonza. For the same reaction system, the yield of the ester was optimized on-line based on concentration estimates from NIR measurements (calibration model obtained from FA) and a kinetic model directly fitted on NIR measurements using FA (Perspektive, 1997; Amrhein *et al.*, 1997). For a diketene reaction conducted in semibatch mode in the laboratories of Lonza, the effect of various catalysts and temperature levels were analyzed by fitting kinetic models on MIR spectral measurements (unpublished for proprietary reasons). For a parallel reaction involving styrene and n-butane (2.8) conducted in semibatch mode in the laboratories of ISW, the yield of 1-phenyl-1-butoxyethan was optimized on-line using the same methodology as mentioned for the esterification reaction (yet unpublished). In addition, reaction system (2.8) was feedback linearized.

6.3 Outlook

Although this work presented several solutions to the problem of the efficient use of concentration and spectral data for the study of reaction systems, there remain numerous important problems whose solutions would help chemists and engineers in their daily tasks.

This work proposed a nonlinear transformation for *homogeneous* reaction systems described by *ordinary* differential equations. For inhomogeneous reaction systems, a

species lives in one or several phases, and the mass exchange between the various phases is usually explained by additional algebraic and/or differential equations. It would be a challenging task for future work to extend the concept of reaction and flow variants/invariants to such systems. Also, numerous reaction systems can be described by partial differential equations (e.g., tubular reactors, fixed-bed reactors). For such systems, Waller and Mäkilä (1981) proposed a transformation that separates the reaction variants and invariants. If the transformed boundary conditions of the reaction invariants are constant over the reactor length, the number of partial differential equations can usually be reduced. It would be interesting to be able to extend the results obtained in this dissertation for lumped systems to distributed parameter systems.

Last but not least, most of the results regarding reaction data were developed for the noise-free case, while the simulated examples were usually presented for noisy concentration and spectral data. A rigorous statistical study to investigate the influence of noise on the various techniques proposed constitutes an important future research direction. Especially, the advantages and disadvantages of the data formulation in reaction-variant form for the case of inaccurate knowledge of the reaction invariants would be interesting to study. Also, the discrimination power of concentration and spectral data for various stoichiometric targets and candidate kinetic models using TFA and FA, respectively, would be worth a careful investigation.

Table 6.1. An overview of the practical problems addressed in this dissertation, the tools proposed to solve them, and the corresponding practical implications. The benefits to the chemist/engineer in the laboratory (academia, vendors, industry) and production (industry) are also indicated.

Problems and benefits	Tools studied	Practical implications
Estimation/prediction of concentrations[a, b] • Replace wet-chemical analyses. • Facilitate model building.	Asymptotic observers	• Can estimate on-line, without knowledge of kinetics and initial concentrations, the concentrations of all the species from those of a few species measured routinely.
	Multivariate calibration Factor analysis	• Can predict concentrations on-line from spectral measurements. • Can determine the applicability of calibration models before generating new spectra. • Help design the calibration set using (i) spectral *and* concentration measurements (multivariate calibration) and (ii) spectral measurements *and* prior knowledge about the underlying reactions (factor analysis).
Building first-principles models[a] • Increase process knowledge. • Facilitate scale-up, monitoring, control, and optimization.	Target factor analysis	• Can determine the number of independent reactions. • Can unambiguously discriminate among stoichiometric targets.
	Factor analysis	On the basis of spectral data: • Can discriminate among candidate first-principles models. • Can determine the termination of a reaction or the reaching of an equilibrium.
Monitoring, control and optimization[b] • Improve safety, product quality and productivity.	Model reduction	• Can reduce the number of differential equations.
	State accessibility	• Can determine, without knowledge of kinetics, the maximal dimension of the state domain accessible from the flows.
	Linearizing feedback	• Can linearize the nonlinear reaction model exactly *via* state feedback.

[a] Laboratory.

[b] Production.

Appendices

A

Numerical values for the simulated examples

A.1 Esterification of ethanol and acetic acid — Examples 2.2

Table A.1 lists the numerical values of the molecular weights, the pure-component densities and the reaction rate coefficients of Example 2.2.

Table A.2 lists the numerical values of the initial concentrations, the inlet concentrations $\mathbf{c}_0$, the number of observations K, the initial volume V_0, and the final batch time t_f of the simulated esterification reaction of Example 2.2.

A.2 Consecutive reactions — Example 2.13

The concentration vector of all the species is $\mathbf{c}_t = [c_1,\ c_2,\ c_3,\ c_4,\ c_5,\ c_6]^{\mathrm{T}}$. The total mass density varies with concentration as $\rho(\mathbf{c}_t) = \sqrt{\mathbf{c}_t^{\mathrm{T}}\,\mathbf{M}_w\,\bar{\boldsymbol{\rho}}}$, where $\bar{\boldsymbol{\rho}}$ is the 6-dimensional vector of pure-component mass densities. The numerical values of the parameters are given in Table A.3.

Six different runs are generated. Table A.4 lists the numerical values of the initial concentrations, the inlet concentrations $\mathbf{c}_0$, the number of observations K, the initial

Table A.1. Parameter values for Example 2.2.

parameter	value	unit
$\{M_i\}$	[46.08, 60.06, 88.12, 18.02, 98.08]	g mol^{-1}
$\{\bar{\rho}_i\}$	[774, 1028, 876, 992, 1814]	g l^{-1}
$[k_1,\ k_2]$	[0.03, 0.01]	l mol^{-2} min^{-2}

Table A.2. Initial concentrations $\mathbf{c}_{t,0}$ [M], inlet concentrations $\mathbf{c}_{in}$ [M], number of observations K, initial volume V_0 [l], and final batch time t_f [min] for the runs R1–R3 of Example 2.2f.

				$\mathbf{c}_{t,0}$					$\mathbf{c}_{in}/\mathbf{c}_{\Delta}$	
Run	K	V_0	t_f	$\mathcal{X}_1$	$\mathcal{X}_2$	$\mathcal{X}_3$	$\mathcal{X}_4$	$\mathcal{X}_5$	$\mathcal{X}_1$	$\mathcal{X}_5$
R1	100	0.5	120	2.9	14.2	0	0.23	0	16.15	0
R2	100	0.5	120	5	5	0	0	0.23	0	0
R3	100	0.5	120	2.9	14.2	0	0	0.23	0	0

Table A.3. Parameter values for the simulated consecutive reactions.

parameter	value	unit
$\{M_{w,i}\}$	[50, 80, 20, 50, 15, 40]	g mol^{-1}
$\{\bar{\rho}_i\}$	[867.5, 1357.6, 809.0, 912.8, 712.0, 789.2]	g l^{-1}
$[\kappa_1, \kappa_{11}, \kappa_2]$	[0.028, 0.280, 0.278]	M^{-1} h^{-1}
κ_{12}	1.000	M^{-1}

volume V_0, the final batch time t_f, and the kinetic expression used for the first reaction in Example A.2.

A.3 Consecutive reactions — Simulated example on page 79

The numerical values of the parameters of the simulated example on page 79 are given in Table A.5.

A.4 Parallel reactions — Examples 2.4

The numerical values for the kinetic parameter of the simulated parallel reactions in the Examples 2.4b, d, e, and f are given in Table A.6.

Table A.4. Initial concentrations $\mathbf{c}_{t,0}$ [M], inlet concentrations $\mathbf{c}_{in}$ [M], number of observations K, initial volume V_0 [l], and final batch time t_f [h] for the six runs R1–R6 of the simulated consecutive reactions in Example A.2.

				$\mathbf{c}_{t,0}$					$\mathbf{c}_{in}/\mathbf{c}_{\Delta}$		
Run	K	V_0	t_f	$\mathcal{X}_1$	$\mathcal{X}_2$	$\mathcal{X}_3$	$\mathcal{X}_4$	$\mathcal{X}_6$	$\mathcal{X}_1$	$\mathcal{X}_4$	Kinetics/Inlet
R1	85	10	7	13.73	0	0	0.86	3.22	4	0	$r_1^{(1)}$/impulse
R2	85	10	7	13.73	0	0	0.86	3.22	0	0	$r_1^{(1)}$/–
R3	61	15	5	13.49	2.41	0	0.96	1.20	0	0	$r_1^{(1)}$/–
R4	85	10	7	13.73	0	0	0.86	3.22	8.55	8.55	$r_1^{(1)}$/continuous
R5	73	10	7	8.41	0	0	1.68	8.41	16.15	0	$r_1^{(2)}$/continuous
R6	73	10	7	8.94	1.12	8.94	0.71	3.58	16.15	0	$r_1^{(2)}$/continuous

Table A.5. Numerical values for the simulated reaction system on 79.

Parameter	Value	Unit
V_0	1.5	l
$\{M_{w,i}\}$	$[50, 100, 75, 12.5]$	g mol^{-1}
$\kappa_{1,0}$, $\kappa_{2,0}$	0.028, 0.278	M^{-1} h^{-1}
E_1, E_2	1.5e-3, 2.0e-3	K
$\{\alpha_i\}$	$[0.3, 0.4, 0.05, 0.1]$	g l^{-1} K^{-2}
$\{\beta_i\}$	$[368, 371, 367, 372]$	K
$\{\gamma_i\}$	$[800, 1300, 900, 700]$	g l^{-1}

Table A.6. Kinetic parameter values for the simulated parallel reactions.

Example	κ_1 [M^{-1} h^{-1}]	κ_2 [h^{-1}]	κ_3 [M^{-1} h^{-1}]
2.4b on page 28	–	2.4	–
2.4d on page 87	0.24	0.3	2.4
2.4e on page 197	1.8	2.4	–
2.4f on page 200	0.09	0.02	–

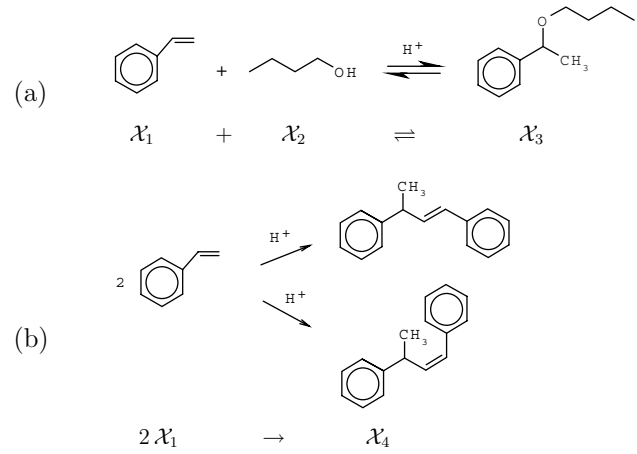

Figure A.1. Parallel reaction system with styrene ($\mathcal{X}_1$). (a) 1-phenyl-1-butoxyethan ($\mathcal{X}_3$) is produced in a hydration reaction from $\mathcal{X}_1$ and n-butanol ($\mathcal{X}_2$). (b) $\mathcal{X}_1$ dimerizes to (trans)-1,3-diphenylbut-1-ene ($\mathcal{X}_4$) and its diastereoisomere, (cis)-1,3-diphenylbut-1-ene.

B

Preliminaries on linear algebra

In this appendix, some preliminaries on linear algebra are provided. They include definitions of the fundamental subspaces of a matrix (null, row, and column spaces; their orthogonal complements), the corresponding dimensions and their relation to the rank, properties of the rank, the numerical construction of the fundamental subspaces by singular value decomposition, and the construction and properties of the pseudo-inverse. The appendix concludes with a brief review of orthogonal projection and its relation to the least-squares problem and the explicit construction of the null space.

B.1 Null, row, and column spaces, and their orthogonal complements

The following definitions hold for the fundamental subspaces of an $m \times n$ matrix $\mathbf{Y}$, involved in the linear transformation $\mathbf{u} = \mathbf{Y}\,\mathbf{b}$:

(D1) **Null space:** The null space or kernel of $\mathbf{Y}$ is the set $\mathcal{N}(\mathbf{Y})$ of all solutions to the homogeneous equation $\mathbf{0}_m = \mathbf{Y}\,\mathbf{b}$. In set notation,

$$\mathcal{N}(\mathbf{Y}) = \{\mathbf{b} : \mathbf{b} \in \mathbb{R}^n \text{ and } \mathbf{Y}\,\mathbf{b} = \mathbf{0}_m.\}$$

(D2) **Row space:** The row space of $\mathbf{Y}$ is the set $\mathcal{S}_r(\mathbf{Y})$ of all linear combinations of the rows of $\mathbf{Y}$. In set notation,

$$\mathcal{S}_r(\mathbf{Y}) = \{\mathbf{b} : \mathbf{u} \in \mathbb{R}^m \text{ and } \mathbf{b} = \mathbf{Y}^{\mathrm{T}}\,\mathbf{u}.\}$$

(D3) **Column space:** The column space or range or image of $\mathbf{Y}$ is the set $\mathcal{S}_c(\mathbf{Y})$ of all linear combinations of the columns of $\mathbf{Y}$. In set notation,

$$\mathcal{S}_c(\mathbf{Y}) = \{\mathbf{u} : \mathbf{b} \in \mathbb{R}^n \text{ and } \mathbf{u} = \mathbf{Y}\,\mathbf{b}.\}$$

It is convenient to note that

$$\text{if } \mathbf{Y} = \mathbf{X}\,\mathbf{Z}, \text{ then } \mathcal{S}_c(\mathbf{Y}) \subset \mathcal{S}_c(\mathbf{X}) \text{ and } \mathcal{S}_r(\mathbf{Y}) \subset \mathcal{S}_r(\mathbf{Z}).$$

The set of all vectors $\boldsymbol{v}$ that are orthogonal to the subspace $\mathcal{W}$ of $\mathbb{R}^n$ is called the *orthogonal complement* of $\mathcal{W}$ and is denoted by $\mathcal{W}^{\perp}$. Then, the following relationships between the orthogonal complements of the fundamental subspaces of $\mathbf{Y}$ hold (see Lay, 1994, p.339).

Theorem B.1
The orthogonal complement of the row space of $\mathbf{Y}$ *is the null space of* $\mathbf{Y}$*, and the orthogonal complement of the column space of* $\mathbf{Y}$ *is the null space of* $\mathbf{Y}^{\mathrm{T}}$*:*

$$\mathcal{S}_r^{\perp}(\mathbf{Y}) = \mathcal{N}(\mathbf{Y}), \qquad \mathcal{S}_c^{\perp}(\mathbf{Y}) = \mathcal{N}(\mathbf{Y}^{\mathrm{T}}).$$

It follows that

$$\mathcal{S}_r(\mathbf{Y}) = \mathcal{S}_c(\mathbf{Y}^{\mathrm{T}}) \text{ and } \mathcal{S}_c(\mathbf{Y}) = \mathcal{S}_r(\mathbf{Y}^{\mathrm{T}}).$$

Theorem B.1 is graphically illustrated in Figure B.1.

B.2 Rank and dimension

The mathematical rank of a matrix $\mathbf{Y}$ (denoted by $\mathrm{rank}(\mathbf{Y})$) is the number of linearly-independent rows and columns of $\mathbf{Y}$.

In this section, the dimensions of the different fundamental subspaces related to the rank and some properties of the rank is discussed.

B.2.1 Dimension

The dimensions of the fundamental subspaces of $\mathbf{Y}$ are related as shown next (see Lay, 1994, p.238).

Theorem B.2
The dimensions of the row and column space of $\mathbf{Y}$ *are equal to the rank* r *of* $\mathbf{Y}$*, i.e.,* $\dim(\mathcal{S}_r(\mathbf{Y})) = \dim(\mathcal{S}_c(\mathbf{Y})) = \mathrm{rank}(\mathbf{Y}) = r$*, which also satisfies:*

$$\mathrm{rank}(\mathbf{Y}) + \dim(\mathcal{N}(\mathbf{Y})) = n, \qquad \mathrm{rank}(\mathbf{Y}^{\mathrm{T}}) + \dim(\mathcal{N}(\mathbf{Y}^{\mathrm{T}})) = m. \tag{B.1}$$

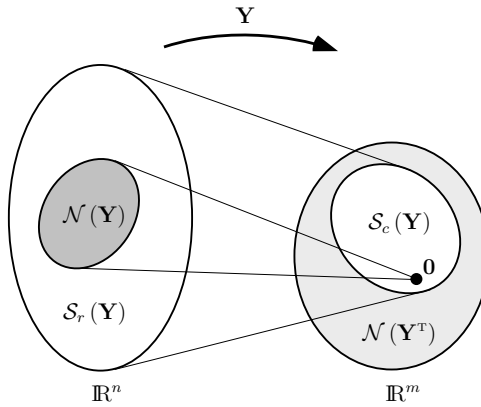

Figure B.1. The fundamental subspaces of an $m \times n$ matrix $\mathbf{Y}$ involved in the linear transformation $\mathbf{u} = \mathbf{Y}\,\mathbf{b}$.

Theorem B.2 uses the following property of the rank:

$$\operatorname{rank}(\mathbf{Y}) = \operatorname{rank}(\mathbf{Y}^{\mathrm{T}}) \tag{B.2}$$

(B.1) tells that the dimension of the null space or *nullity* of $\mathbf{Y}$ is given by the difference between the number of columns of $\mathbf{Y}$ and its rank. Table B.1 lists the dimensions of the different fundamental subspaces of $\mathbf{Y}$.

B.2.2 Properties of the rank

Some properties of the rank are listed below. For the basics of rank, it is referred to Magnus *et al.* (1988).

The rank of a matrix $\mathbf{Y}$ is at most the minimum of its dimension:

Property B.3 *Let $\mathbf{Y}$ be an $m \times n$ matrix. Then,*

$$\operatorname{rank}(\mathbf{Y}) \leq \min(m, n). \tag{B.3}$$

$\mathbf{Y}$ *is said* rank deficient, *if its rank is less than the minimum of its dimension, i.e.,* $\operatorname{rank}(\mathbf{Y}) < \min(m, n)$.

Property B.4 *If $\mathbf{Y}$ can be decomposed into two matrices $\mathbf{T}$ $(m \times r)$ and $\mathbf{P}$ $(n \times r)$,* $\mathbf{Y} = \mathbf{T}\,\mathbf{P}^{\mathrm{T}}$, *then*

$$\operatorname{rank}(\mathbf{Y}) \leq \min\left(\operatorname{rank}(\mathbf{T}), \operatorname{rank}(\mathbf{P})\right). \tag{B.4}$$

Linear dependency of either the rows or the columns of $\mathbf{Y}$ is closely linked to its rank. If $\mathbf{Y}$ contains linearly-dependent rows or columns, it can be decomposed into two matrices of full rank:

Property B.5 *If the last $(m - A)$ rows of $\mathbf{Y}$ are linearly dependent on the first A ones, then $\mathbf{Y}$ can be decomposed into an $m \times A$ matrix $\mathbf{T}$ and an $n \times A$ matrix $\mathbf{P}$:*

$$\mathbf{Y} = \mathbf{T}\,\mathbf{P}^{\mathrm{T}} \qquad \textit{with} \qquad \mathbf{T} = \begin{bmatrix} \mathbf{I}_A \\ \mathbf{T}_1 \end{bmatrix}, \tag{B.5}$$

where $\mathbf{T}_1$ is an $(m - A) \times A$ matrix. A similar property holds for the columns also.

Table B.1. Dimensions of the fundamental subspaces of the $m \times n$ matrix $\mathbf{Y}$. $\operatorname{rank}(\mathbf{Y}) = r$.

Space	$\mathbb{R}^n$	$\mathcal{N}(\mathbf{Y})$	$\mathcal{S}_r(\mathbf{Y})$	$\mathcal{S}_c(\mathbf{Y})$
Dimension	n	$n - r$	r	r
Space	$\mathbb{R}^m$	$\mathcal{N}(\mathbf{Y}^{\mathrm{T}})$	$\mathcal{S}_c(\mathbf{Y}^{\mathrm{T}})$	$\mathcal{S}_r(\mathbf{Y}^{\mathrm{T}})$
Dimension	m	$m - r$	r	r

Column-mean centering of $\mathbf{Y}$ will be become important in factor analysis. If the $\mathbf{1}_m$-vector lies in the column space of $\mathbf{Y}$, then a rank drop of one can be observed. Alternatively, a rank drop of one can be observed when $\mathbf{Y}$ is a wide matrix ($m < n$) of full rank. Otherwise, column-mean centering does not cause a rank drop:

Property B.6 *Let $\bar{\mathbf{Y}}_c = \mathbf{J}_c \mathbf{Y}$ be the column-mean-centered matrix of $\mathbf{Y}$, where $\mathbf{J}_c \equiv \mathbf{I}_m - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m^{\mathrm{T}}$ is the $m \times m$ centering matrix. Then, the bounds on the rank of $\bar{\mathbf{Y}}_c$ are given by*

$$\operatorname{rank}(\mathbf{Y}) - 1 \leq \operatorname{rank}\left(\bar{\mathbf{Y}}_c\right) \leq \min(\operatorname{rank}(\mathbf{Y}), m-1). \tag{B.6}$$

In addition, iff $\mathbf{1}_m \in \mathcal{S}_c(\mathbf{Y})$, then

$$\operatorname{rank}\left(\bar{\mathbf{Y}}_c\right) = \operatorname{rank}(\mathbf{Y}) - 1; \quad \mathbf{1}_m \in \mathcal{S}_c(\mathbf{Y}), \tag{B.7}$$

else $\operatorname{rank}\left(\bar{\mathbf{Y}}_c\right) = \min(\operatorname{rank}(\mathbf{Y}), m-1)$. *A similar property holds for row centering also.*

Proof:
Since $\mathbf{J}_c$ is an idempotent matrix, its rank is determined by the trace:

$$\operatorname{rank}(\mathbf{J}_c) = \operatorname{tr}(\mathbf{J}_c) = m - 1. \tag{B.8}$$

Since $\bar{\mathbf{Y}}_c = \mathbf{J}_c \mathbf{Y}$ is expressed as a product of two matrices, by Property B.4, $\operatorname{rank}\left(\bar{\mathbf{Y}}_c\right) \leq \min(\operatorname{rank}(\mathbf{Y}), m-1)$. Since the null space of $\mathbf{J}_c$ has dimension 1, the dimension of the column space of $\mathbf{Y}$ can be reduced at most by one. Thus, $\operatorname{rank}\left(\bar{\mathbf{Y}}_c\right) \geq \operatorname{rank}(\mathbf{Y}) - 1$. Since the basis for the null space of $\mathbf{J}_c$ is the vector $\mathbf{1}_m$, rank reduction is ensured if this vector $\mathbf{1}_m$ is in the column space of $\mathbf{Y}$. Else, the rank remains intact, only to be reduced if $\operatorname{rank}(\mathbf{Y}) = m$. □

It is shown below that regarding the rank, differentiation in observation direction is similar to column-mean centering:

Property B.7 *Let $\dot{\mathbf{Y}}$ denote the matrix differentiated in observation direction (in this case, calculating the backward differences along the columns of $\mathbf{Y}$). The bounds on the ranks of $\dot{\mathbf{Y}}$ are given by*

$$\operatorname{rank}(\mathbf{Y}) - 1 \leq \operatorname{rank}\left(\dot{\mathbf{Y}}\right) \leq \min(\operatorname{rank}(\mathbf{Y}), m-1). \tag{B.9}$$

In addition, iff $\mathbf{1}_m \in \mathcal{S}_c(\mathbf{Y})$, then

$$\operatorname{rank}\left(\dot{\mathbf{Y}}\right) = \operatorname{rank}(\mathbf{Y}) - 1, \tag{B.10}$$

else $\operatorname{rank}\left(\dot{\mathbf{Y}}\right) = \min(\operatorname{rank}(\mathbf{Y}), m-1)$. *A similar property holds for finite differences along the rows also.*

Proof:
Since the first $(m-1)$ rows of D_c are linearly independent and the last row is $\mathbf{0}_m^{\mathrm{T}}$, $\operatorname{rank}(\mathsf{D}_c) = m - 1$. Since $\dot{\mathbf{Y}} = \mathsf{D}_c \mathbf{Y}$ is expressed as a product of two matrices, by

Property B.4, $\operatorname{rank}\left(\dot{\mathbf{Y}}\right) \leq \min\left(\operatorname{rank}\left(\mathbf{Y}\right), m-1\right)$. Since the null space of D_c has dimension 1, the dimension of the column space of $\mathbf{Y}$ can be reduced at most by one. Thus, $\operatorname{rank}\left(\dot{\mathbf{Y}}\right) \geq \operatorname{rank}\left(\mathbf{Y}\right) - 1$. It can be seen that the basis for the null space of D_c is the vector $\mathbf{1}_m$, since $\mathsf{D}_c\,\mathbf{1}_m = \mathbf{0}_m$. So, rank reduction is ensured if $\mathbf{1}_m$ is in the column space of $\mathbf{Y}$. Else, the rank remains intact, only to be reduced if $\operatorname{rank}\left(\mathbf{Y}\right) = m$. □

It is shown below that when $\mathbf{1}_m \in \mathcal{S}_c\left(\mathbf{Y}\right)$, subtracting the first row of matrix from all the rows is similar regarding the rank to column-mean centering and differentiation in observation direction:

Property B.8 *Let* $\mathbf{Y}_0$ *be defined as:*

$$\mathbf{Y}_0 \equiv \mathbf{Y} - \mathbf{1}_m \mathbf{y}_0^{\mathrm{T}}, \tag{B.11}$$

where $\mathbf{y}_0$ *denotes the first row of* $\mathbf{Y}$*. Iff* $\mathbf{1}_m \in \mathcal{S}_c\left(\mathbf{Y}\right)$*, then*

$$\operatorname{rank}\left(\mathbf{Y}_0\right) = \operatorname{rank}\left(\mathbf{Y}\right) - 1. \tag{B.12}$$

Proof:
⇒: If $\mathbf{1}_m \in \mathcal{S}_c\left(\mathbf{Y}\right)$, then $\mathbf{Y}$ can be expressed as

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}^{\dagger} & \mathbf{1}_m \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{y}_0^{\mathrm{T}} \end{bmatrix}, \qquad \mathbf{Y}_s = \begin{bmatrix} \mathbf{0}_n^{\mathrm{T}} \\ \mathbf{Y}^{\ddagger} \end{bmatrix}, \tag{B.13}$$

where $\mathbf{Y}^{\dagger}$, $\mathbf{Y}^{\ddagger}$, and $\mathbf{U}$ are matrices of dimension $m \times (n-1)$, $(m-1) \times (n-1)$, and $(n-1) \times n$, respectively, with $\mathbf{1}_m \notin \mathcal{S}_c\left(\mathbf{Y}^{\dagger}\right)$ and $\mathbf{y}_0 \notin \mathcal{S}_r\left(\mathbf{U}\right)$. Thus, $\mathbf{Y}_0$ becomes

$$\mathbf{Y}_0 = \mathbf{Y}^{\dagger}\mathbf{U}. \tag{B.14}$$

Thus, $\operatorname{rank}\left(\mathbf{Y}_0\right) = (n-1)$. Since $\operatorname{rank}\left(\mathbf{Y}\right) = n$, $\operatorname{rank}\left(\mathbf{Y}_0\right) = \operatorname{rank}\left(\mathbf{Y}\right) - 1$ follows.

⇐: Without loss of generality, $\mathbf{Y}$ can be decomposed as:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}^{\dagger} & \boldsymbol{\beta} \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{y}_0^{\mathrm{T}} \end{bmatrix}, \qquad \mathbf{Y}_s = \begin{bmatrix} \mathbf{0}_n^{\mathrm{T}} \\ \mathbf{Y}^{\ddagger} \end{bmatrix}, \tag{B.15}$$

where $\boldsymbol{\beta}$ is an m-dimensional vector. Assume that $\boldsymbol{\beta} \neq \mathbf{1}_m$ and $\mathbf{1}_m \notin \mathcal{S}_c\left(\mathbf{Y}^{\dagger}\right)$. Then, from (B.11), $\mathbf{Y}_0$ becomes:

$$\mathbf{Y}_0 = \mathbf{Y}^{\dagger}\mathbf{U} + \left(\boldsymbol{\beta} - \mathbf{1}_m\right)\mathbf{y}_0^{\mathrm{T}}. \tag{B.16}$$

Since $\mathbf{1}_m \notin \mathcal{S}_c\left(\mathbf{Y}^{\dagger}\right)$ and $\mathbf{y}_0 \notin \mathcal{S}_r\left(\mathbf{U}\right)$ by construction, for $\operatorname{rank}\left(\mathbf{Y}_0\right) = \operatorname{rank}\left(\mathbf{Y}\right) - 1$ to hold, $\left(\boldsymbol{\beta} - \mathbf{1}_m\right)$ must be $\mathbf{0}_m$. This is contradicted by the initial assumption and, thus, Property B.8 follows. □

In factor analysis and calibration, increasing the rank of $\mathbf{Y}$ plays an important role. In many practical applications, the rank of a data matrix can be increased by one. The corresponding rank-increasing matrix can be written as an outer product of two vectors. If these vectors do not lie in the column or row space of $\mathbf{Y}$, a rank increase of one can be observed under certain mild assumptions:

Property B.9 *Let* $\mathbf{Y}$ *be an* $m \times n$ *matrix of rank* A *that can be decomposed into two matrices* $\mathbf{T}$ *(*$m \times A$*) and* $\mathbf{P}$ *(*$n \times A$*),* $\mathbf{Y} = \mathbf{T}\,\mathbf{P}^{\mathrm{T}}$*, with* $\mathbf{T}$ *and* $\mathbf{P}$ *being matrices of full rank. Let* $\mathbf{U}$ *be a matrix constructed as:* $\mathbf{U} = \mathbf{Y} + \boldsymbol{\alpha}\,\mathbf{y}_{ref}^{\mathrm{T}}$*. If* $\mathbf{y}_{ref}$ *and* $\boldsymbol{\alpha}$ *do not lie in the column space of* $\mathbf{P}$ *and* $\mathbf{T}$*, respectively, then*

$$\operatorname{rank}(\mathbf{U}) = \min(\operatorname{rank}(\mathbf{Y}) + 1, m, n); \qquad \mathbf{y}_{ref} \notin \mathcal{S}_c(\mathbf{P}), \quad \boldsymbol{\alpha} \notin \mathcal{S}_c(\mathbf{T}). \tag{B.17}$$

Proof:
$\mathbf{U} = \mathbf{T}\,\mathbf{P}^{\mathrm{T}} + \boldsymbol{\alpha}\,\mathbf{y}_{ref}^{\mathrm{T}}$ can be written as

$$\mathbf{U} = \underbrace{\begin{bmatrix}\mathbf{T} & \boldsymbol{\alpha}\end{bmatrix}}_{\mathbf{T}^{\star}} \underbrace{\begin{bmatrix}\mathbf{P}^{\mathrm{T}} \\ \mathbf{y}_{ref}^{\mathrm{T}}\end{bmatrix}}_{(\mathbf{P}^{\star})^{\mathrm{T}}},$$

where $\mathbf{T}^{\star}$ and $\mathbf{P}^{\star}$ are matrices of dimension $m \times (A+1)$ and $n \times (A+1)$, respectively. If $\boldsymbol{\alpha} \notin \mathcal{S}_c(\mathbf{T})$ and $\mathbf{y}_{ref} \notin \mathcal{S}_c(\mathbf{P})$, then $\mathbf{T}^{\star}$ and $\mathbf{P}^{\star}$ are of full rank. Property B.9 holds trivially from Properties B.3 and B.4. □

If no information about $\mathbf{T}$ and $\mathbf{P}$ in Property B.4 is available, only an upper bound for the rank of $\mathbf{Y}$ could be provided. However, if the rank of one of these matrices is full, then the rank of $\mathbf{Y}$ is determined by the (possibly) rank-deficient matrix.

Property B.10 *Let* $\mathbf{U}$ *(*$m \times s$*) be decomposed into two matrices* $\mathbf{T}$ *(*$m \times A$*) and* $\mathbf{Q}$ *(*$s \times A$*),* $\mathbf{U} = \mathbf{T}\,\mathbf{Q}$*. Then, the following properties hold:*

(a) *If* $\operatorname{rank}(\mathbf{Q}) = A$*, then* $\operatorname{rank}(\mathbf{U}) = \operatorname{rank}(\mathbf{T})$.
(b) *If* $\operatorname{rank}(\mathbf{T}) = A$*, then* $\operatorname{rank}(\mathbf{U}) = \operatorname{rank}(\mathbf{Q})$.

Proof:
For the proof of Property B.10a, consider $\operatorname{rank}(\mathbf{U}) \leq \min(\operatorname{rank}(\mathbf{T}), \operatorname{rank}(\mathbf{Q}))$ by invoking Property B.4. Also, by Property B.3, $\operatorname{rank}(\mathbf{T}) \leq A$ and by assumption $\operatorname{rank}(\mathbf{Q}) = A$. Since $\operatorname{rank}(\mathbf{T}) \leq \operatorname{rank}(\mathbf{Q})$, it follows that $\operatorname{rank}(\mathbf{U}) \leq \operatorname{rank}(\mathbf{T})$. The situation $\operatorname{rank}(\mathbf{U}) < \operatorname{rank}(\mathbf{T})$ arises only, when some of the rows of $\mathbf{T}$ are orthogonal to the subspace spanned by the rows of $\mathbf{Q}$. Since $\mathbf{Q}$ is of full rank, the rows of $\mathbf{Q}$ span the entire A-dimensional subspace. So, no row of $\mathbf{T}$ can be orthogonal to the row space of $\mathbf{Q}$ unless it is a row with all zeros. Thus, $\operatorname{rank}(\mathbf{U}) = \operatorname{rank}(\mathbf{T})$ and Property B.10a follows. $\operatorname{rank}(\mathbf{U}) = \operatorname{rank}(\mathbf{Q})$ and Property B.10b follow from $\operatorname{rank}(\mathbf{T}) = A$ by invoking Property B.10a. □

Property B.10 can be extended directly to $\mathbf{Y}$ being composed of three matrices:

Property B.11 *Let* $\mathbf{Y}$ *(*$m \times n$*) be decomposed into three matrices* $\mathbf{T}$*,* $\mathbf{Q}$*, and* $\mathbf{P}$ *(*$n \times s$*),* $\mathbf{Y} = \mathbf{U}\,\mathbf{P}^{\mathrm{T}} = \mathbf{T}\,\mathbf{Q}\,\mathbf{P}^{\mathrm{T}}$*. If* $\operatorname{rank}(\mathbf{P}) = s$*, then the following properties hold:*

(c) $\operatorname{rank}(\mathbf{Y}) = \operatorname{rank}(\mathbf{U})$.
(d) *If* $\operatorname{rank}(\mathbf{Q}) = A$*, then* $\operatorname{rank}(\mathbf{Y}) = \operatorname{rank}(\mathbf{T})$.
(e) *If* $\operatorname{rank}(\mathbf{T}) = A$*, then* $\operatorname{rank}(\mathbf{Y}) = \operatorname{rank}(\mathbf{Q})$.

B.3 Orthogonal bases and the singular value decomposition (SVD)

In this section, singular value decomposition (SVD; Golub and Van Loan, 1983) of the $m \times n$ matrix $\mathbf{Y}$ is used to compute the orthogonal bases of the fundamental subspaces. Yet, the orthogonal basis (see Lay, 1994, p.343) and the SVD (see Lay, 1994, p.430) of $\mathbf{Y}$ are defined first.

Theorem B.12 (Orthogonal basis)
Let $\{\boldsymbol{\upsilon}_1, \ldots, \boldsymbol{\upsilon}_r\}$ *be an* orthogonal basis *for a subspace* $\mathcal{W}$ *of* $\mathbb{R}^n$. *Then, each* $\mathbf{y}$ *in* $\mathcal{W}$ *has a unique representation as a linear combination of* $\mathbf{u}_1, \ldots, \mathbf{u}_r$. *In fact, if*

$$\mathbf{y} = \alpha_1 \mathbf{u}_1 + \cdots + \alpha_r \mathbf{u}_r,$$

the weights (coordinates) α_i $(i = 1, \ldots, r)$ *can be computed as:*

$$\alpha_i = \mathbf{u}_i^{\mathrm{T}} \mathbf{y} / (\mathbf{u}_i^{\mathrm{T}} \mathbf{u}_i) \qquad \forall i = 1, \ldots, r.$$

Theorem B.13 (Singular value decomposition)
Let $\mathbf{Y}$ *be an* $m \times n$ *matrix with* $m > n$. *Then, there exists a decomposition of* $\mathbf{Y}$ *such that:*

$$\mathbf{Y} = \boldsymbol{\Upsilon}\, \boldsymbol{\Sigma}\, \boldsymbol{\Omega}^{\mathrm{T}}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix} \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n) \\ \mathbf{0}_{(m-n)\times n} \end{bmatrix}, \tag{B.18}$$

where $\boldsymbol{\Sigma}$ *is the* $m \times n$ *singular value matrix containing the singular values* σ_i *in decreasing order* $(i = 1, 2, \ldots, n)$, $\boldsymbol{\Upsilon} = [\boldsymbol{\upsilon}_1, \ldots, \boldsymbol{\upsilon}_m]$ *and* $\boldsymbol{\Omega} = [\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_n]$ *the* $m \times m$ *and* $n \times n$ *unitary matrices composed of the left and right singular vectors corresponding to* $\boldsymbol{\Sigma}$, *and* $\operatorname{diag}(a_1, \ldots, a_A)$ *an* A*-dimensional diagonal matrix containing the elements* $a_1, \ldots, a_A$.

Let $\operatorname{rank}(\mathbf{Y}) = r$. Then, the following properties of the SVD regarding the fundamental subspaces of $\mathbf{Y}$ are in order:

(P1) The set of the left singular vectors

$$\{\boldsymbol{\upsilon}_1, \ldots, \boldsymbol{\upsilon}_r\} \text{ form an orthonormal basis for } \mathcal{S}_c(\mathbf{Y}). \tag{B.19}$$

(P2) Since $\mathcal{S}_c^{\perp}(\mathbf{Y}) = \mathcal{N}(\mathbf{Y}^{\mathrm{T}})$,

$$\{\boldsymbol{\upsilon}_{r+1}, \ldots, \boldsymbol{\upsilon}_m\} \text{ is an orthonormal basis for } \mathcal{N}(\mathbf{Y}^{\mathrm{T}}). \tag{B.20}$$

(P3) Since $\mathbf{Y}\, \boldsymbol{\omega}_i = \sigma_i \upsilon_i$ for $1 \leq i \leq r$, and σ_i is 0 iff $i > r$, it is concluded that

$$\{\boldsymbol{\omega}_{r+1}, \ldots, \boldsymbol{\omega}_n\} \text{ is an orthonormal basis for } \mathcal{N}(\mathbf{Y}). \tag{B.21}$$

(P4) From (B.19) and (B.20), the orthogonal complement of $\mathcal{N}(\mathbf{Y}^{\mathrm{T}})$. Interchanging $\mathbf{Y}$ and $\mathbf{Y}^{\mathrm{T}}$, it follows that $\mathcal{N}^{\perp}(\mathbf{Y}) = \mathcal{S}_c(\mathbf{Y}^{\mathrm{T}}) = \mathcal{S}_r(\mathbf{Y})$. Thus, from (B.21),

$$\{\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_r\} \text{ is an orthonormal basis for } \mathcal{S}_r(\mathbf{Y}). \tag{B.22}$$

B.4 Pseudo-inverse

First, the Moore–Penrose (MP) pseudo-inverse is defined, which is useful because of its existence and uniqueness (Magnus *et al.*, 1988). Then, some properties of the MP pseudo-inverse are provided that are used implicitly for some proofs in this dissertation. The section concludes with the numerical construction of the MP pseudo-inverse using SVD.

Definition B.14 (Moore–Penrose pseudo-inverse) *An $n \times m$ matrix $\mathbf{U}$ is the MP pseudo-inverse of a real $m \times n$ matrix $\mathbf{Y}$, if*

$$\mathbf{Y}\,\mathbf{U}\,\mathbf{Y} = \mathbf{Y} \tag{B.23}$$

$$\mathbf{U}\,\mathbf{Y}\,\mathbf{U} = \mathbf{U} \tag{B.24}$$

$$(\mathbf{Y}\,\mathbf{U})^{\mathrm{T}} = \mathbf{Y}\,\mathbf{U} \tag{B.25}$$

$$(\mathbf{U}\,\mathbf{Y})^{\mathrm{T}} = \mathbf{U}\,\mathbf{Y} \tag{B.26}$$

The MP pseudo-inverse of $\mathbf{Y}$ is denoted as $\mathbf{Y}^{+}$.

Proposition B.15 *The following properties hold for a MP pseudo-inverse (for proof see Magnus et al., 1988):*

(a) $\mathbf{Y}^{+} = \mathbf{Y}^{-1}$ *for non-singular* $\mathbf{Y}$,
(b) $(\mathbf{Y}^{+})^{+} = \mathbf{Y}$,
(c) $(\mathbf{Y}^{\mathrm{T}})^{+} = (\mathbf{Y}^{+})^{\mathrm{T}}$,
(d) $\mathbf{Y}^{+} = \mathbf{Y}$, *if* $\mathbf{Y}$ *is symmetric and idempotent,*
(e) $\mathbf{Y}\,\mathbf{Y}^{+}$ *and* $\mathbf{Y}^{+}\,\mathbf{Y}$ *are idempotent,*
(f) $\mathbf{Y}$, $\mathbf{Y}^{+}$, $\mathbf{Y}\,\mathbf{Y}^{+}$, *and* $\mathbf{Y}^{+}\,\mathbf{Y}$ *have the same rank,*
(g) $\mathbf{Y}^{\mathrm{T}}\,\mathbf{Y}\,\mathbf{Y}^{+} = \mathbf{Y}^{\mathrm{T}} = \mathbf{Y}^{+}\,\mathbf{Y}\,\mathbf{Y}^{\mathrm{T}}$,
(h) $\mathbf{Y}^{\mathrm{T}}\,(\mathbf{Y}^{+})^{\mathrm{T}}\,\mathbf{Y}^{+} = \mathbf{Y}^{+} = \mathbf{Y}^{+}\,(\mathbf{Y}^{+})^{\mathrm{T}}\,\mathbf{Y}^{\mathrm{T}}$,
(i) $(\mathbf{Y}^{\mathrm{T}}\,\mathbf{Y})^{+} = \mathbf{Y}^{+}\,(\mathbf{Y}^{+})^{\mathrm{T}}$, $(\mathbf{Y}\,\mathbf{Y}^{\mathrm{T}})^{+} = (\mathbf{Y}^{+})^{\mathrm{T}}\,\mathbf{Y}^{+}$,
(j) $\mathbf{Y}\,(\mathbf{Y}^{\mathrm{T}}\,\mathbf{Y})^{+}\,\mathbf{Y}^{\mathrm{T}}\,\mathbf{Y} = \mathbf{Y} = \mathbf{Y}\,\mathbf{Y}^{\mathrm{T}}\,(\mathbf{Y}\,\mathbf{Y}^{\mathrm{T}})^{+}\,\mathbf{Y}$,
(k) $\mathbf{Y}^{+} = (\mathbf{Y}^{\mathrm{T}}\,\mathbf{Y})^{+}\,\mathbf{Y}^{\mathrm{T}} = \mathbf{Y}^{\mathrm{T}}\,(\mathbf{Y}\,\mathbf{Y}^{\mathrm{T}})^{+}$,
(l) $\mathbf{Y}^{+} = (\mathbf{Y}^{\mathrm{T}}\,\mathbf{Y})^{-1}\,\mathbf{Y}^{\mathrm{T}}$, *if* $\mathbf{Y}$ *has full column rank,*
(m) $\mathbf{Y}^{+} = \mathbf{Y}^{\mathrm{T}}\,(\mathbf{Y}\,\mathbf{Y}^{\mathrm{T}})^{-1}$, *if* $\mathbf{Y}$ *has full row rank,*
(n) $\mathbf{Y} = \mathbf{0}_{m\times n} \leftrightarrow \mathbf{Y}^{+} = \mathbf{0}_{n\times m}$,
(o) $\mathbf{Y}\,\mathbf{U} = \mathbf{0}_{m\times m} \leftrightarrow \mathbf{U}^{+}\,\mathbf{Y}^{+} = \mathbf{0}_{m\times m}$,
(p) $\mathbf{Y}^{+}\,\mathbf{U} = \mathbf{0}_{n\times n} \leftrightarrow \mathbf{Y}^{\mathrm{T}}\,\mathbf{U} = \mathbf{0}_{n\times n}$.

Proposition B.16 (Construction of $\mathbf{Y}^{+}$) *Let* $\operatorname{rank}(\mathbf{Y}) = r$. *Then, the MP pseudo-inverse of* $\mathbf{Y}$ *can be constructed by:*

$$\mathbf{Y}^{+} = \mathbf{\Omega}_r\,\mathbf{\Sigma}_r^{-1}\,\mathbf{\Upsilon}_r^{\mathrm{T}}, \tag{B.27}$$

where the subscript r denotes a submatrix corresponding to first r singular values.

Proof:
The proof uses the unitary properties of $\mathbf{\Omega}_r$ and $\mathbf{\Upsilon}_r$ in (B.18). □

B.5 Orthogonal projection and the least-squares problem

In this section, the relations between orthogonal projection and the least-squares solution are provided. Also, an explicit expression of the null space of a matrix is derived that complements the numerical solution by SVD in Section B.3. Also, the empirical dispersion (or cross-product), covariance, and correlation matrices are defined.

B.5.1 Orthogonal projection

Theorem B.17 (The orthonormal decomposition theorem)
Let $\mathcal{W}$ be a subspace of $\mathbb{R}^n$ that has an orthogonal basis. Then, each $\mathbf{u} \in \mathbb{R}^n$ can be written uniquely in the form

$$\mathbf{u} = \hat{\mathbf{u}} + \boldsymbol{\varepsilon},$$

where $\hat{\mathbf{u}} \in \mathcal{W}$, $\boldsymbol{\varepsilon} \in \mathcal{W}^{\perp}$, and $\hat{(\cdot)}$ denotes an estimated or predicted quantity. In fact, if the columns of $\boldsymbol{\Upsilon} = [\boldsymbol{\upsilon}_1, \ldots, \boldsymbol{\upsilon}_r]$ form an orthonormal basis of $\mathcal{W}$, then the orthogonal projection of $\mathbf{u}$ onto $\mathcal{W}$ becomes:

$$\hat{\mathbf{u}} = \boldsymbol{\mathcal{P}}\,\mathbf{u}, \qquad \boldsymbol{\mathcal{P}} \equiv \boldsymbol{\Upsilon}\,\boldsymbol{\Upsilon}^{\mathrm{T}}, \tag{B.28}$$

with the corresponding projection error:

$$\boldsymbol{\varepsilon} = \mathbf{u} - \hat{\mathbf{u}} = (\mathbf{I}_n - \boldsymbol{\mathcal{P}})\ \mathbf{u}. \tag{B.29}$$

B.5.2 Least-squares problem and its dual

Let $\mathbf{Y}$ be an $m \times n$ matrix, and $\mathbf{u}$ and $\mathbf{b}$ vectors of dimension m and n, respectively. The least-squares solution to

$$\mathbf{u} = \mathbf{Y}\,\mathbf{b} \tag{B.30}$$

is a vector $\mathbf{b}$ that makes $\mathbf{Y}\,\mathbf{b}$ as close as possible to $\mathbf{u}$ in the Euclidian norm (least squares) sense:

$$\min_{\mathbf{b}} \|\mathbf{u} - \mathbf{Y}\,\mathbf{b}\|_2, \tag{B.31}$$

where $\|\cdot\|_2$ denotes the Euclidian norm of a vector. It is shown next that the least-squares solution is closely related to the orthogonal projection of $\mathbf{u}$ on a subspace of $\mathbf{Y}$.

Theorem B.18 (The least-squares solution)
Let $\mathrm{rank}\,(\mathbf{Y}) = r$. *Then, the least-squares solution to* (B.31) *is given by*

$$\hat{\mathbf{b}} = \mathbf{Y}^{+}\mathbf{u}, \tag{B.32}$$

with $\mathbf{Y}^+$ *from* (B.27). *Thus,* $\hat{\mathbf{u}}$ *is the orthogonal projection of* $\mathbf{u}$ *onto the column space of* $\mathbf{Y}$, *spanned by the columns of* $\boldsymbol{\Upsilon}_r$:

$$\hat{\mathbf{u}} = \mathbf{Y}\,\hat{\mathbf{b}} = (\mathbf{Y}\,\mathbf{Y}^+)\,\mathbf{u} = \boldsymbol{\mathcal{P}}_u\,\mathbf{u}, \qquad \boldsymbol{\mathcal{P}}_u \equiv \boldsymbol{\Upsilon}_r\,\boldsymbol{\Upsilon}_r^{\mathrm{T}} \tag{B.33}$$

with $\boldsymbol{\mathcal{P}}_u$ *being an* $m \times m$ *projection matrix. The corresponding least-squares error is given by:*

$$\hat{\boldsymbol{\varepsilon}}_e = \mathbf{u} - \hat{\mathbf{u}} = (\mathbf{I}_m - \boldsymbol{\mathcal{P}}_u)\;\mathbf{u}. \tag{B.34}$$

The proof can be found in Lay (1994).

The dual problem to (B.30) is

$$\min_{\mathbf{u}} \|\mathbf{b} - \mathbf{Y}^{\mathrm{T}}\,\mathbf{u}\|_2 \quad \text{with} \quad \mathbf{b} = \mathbf{Y}^{\mathrm{T}}\,\mathbf{u}. \tag{B.35}$$

Then, $\hat{\mathbf{b}}$ corresponding to the least-squares solution, $\hat{\mathbf{u}} = \mathbf{Y}^{+\mathrm{T}}\mathbf{b}$, is the orthogonal projection of $\mathbf{b}$ onto the *row space of* $\mathbf{Y}$, spanned by the rows of $\boldsymbol{\Omega}_r^{\mathrm{T}}$:

$$\hat{\mathbf{b}} = \mathbf{Y}^{\mathrm{T}}\,\hat{\mathbf{u}} = (\mathbf{Y}^+\mathbf{Y})\,\mathbf{b} = \boldsymbol{\mathcal{P}}_b\,\mathbf{b}, \qquad \boldsymbol{\mathcal{P}}_b \equiv \boldsymbol{\Omega}_r\,\boldsymbol{\Omega}_r^{\mathrm{T}}, \tag{B.36}$$

with $\boldsymbol{\mathcal{P}}_b$ being an $n \times n$ projection matrix. The corresponding least-squares error is given by:

$$\hat{\boldsymbol{\varepsilon}}_e = \mathbf{b} - \hat{\mathbf{b}} = (\mathbf{I}_n - \boldsymbol{\mathcal{P}}_b)\;\mathbf{b}. \tag{B.37}$$

B.5.3 Relation between projection error and null spaces

The projection errors of the least-squares problem (see B.34) and its dual (see B.37) involve the following two projection matrices:

$$\begin{aligned} \mathbf{J}_u &\equiv \mathbf{I}_m - \boldsymbol{\mathcal{P}}_u = \mathbf{I}_m - \mathbf{Y}\,\mathbf{Y}^+ = \mathbf{I}_m - \boldsymbol{\Upsilon}_r\,\boldsymbol{\Upsilon}_r^{\mathrm{T}}, \\ \mathbf{J}_b &\equiv \mathbf{I}_n - \boldsymbol{\mathcal{P}}_b = \mathbf{I}_n - \mathbf{Y}^+\mathbf{Y} = \mathbf{I}_n - \boldsymbol{\Omega}_r\,\boldsymbol{\Omega}_r^{\mathrm{T}}. \end{aligned} \tag{B.38}$$

Since $\boldsymbol{\Upsilon}_r$ and $\boldsymbol{\Omega}_r$ are unitary matrices, $\operatorname{rank}(\mathbf{J}_u) = (m - r)$ and $\operatorname{rank}(\mathbf{J}_b) = (n - r)$.

The null space of $\mathbf{Y}$ (or $\mathbf{Y}^{\mathrm{T}}$) is an *implicit* definition. An equivalent *explicit* definition of these null spaces are provided next.

Proposition B.19 *Let* $\operatorname{rank}(\mathbf{Y}) = r$ *and* $\mathbf{J}_b$ *and* $\mathbf{J}_u$ *be defined as in* (B.38). *Then,*

$$\mathcal{N}(\mathbf{Y}) = \mathcal{S}_c(\mathbf{J}_b), \qquad \mathcal{N}(\mathbf{Y}^{\mathrm{T}}) = \mathcal{S}_c(\mathbf{J}_u). \tag{B.39}$$

Proof:
For $\mathcal{N}(\mathbf{Y}) = \mathcal{S}_c(\mathbf{J}_b)$, $\mathbf{Y}\,\mathbf{J}_b = \mathbf{0}_{m\times n}$ must be proven. Substituting $\mathbf{Y} = \boldsymbol{\Upsilon}_r\,\boldsymbol{\Sigma}_r\,\boldsymbol{\Omega}_r^{\mathrm{T}}$ and $\mathbf{J}_b = \mathbf{I}_n - \boldsymbol{\Omega}_r\,\boldsymbol{\Omega}_r^{\mathrm{T}}$, it follows that $\mathbf{Y}\,\mathbf{J}_b = \boldsymbol{\Upsilon}_r\,\boldsymbol{\Sigma}_r\,\boldsymbol{\Omega}_r^{\mathrm{T}} - \boldsymbol{\Upsilon}_r\,\boldsymbol{\Sigma}_r\,(\boldsymbol{\Omega}_r^{\mathrm{T}}\boldsymbol{\Omega}_r)\,\boldsymbol{\Omega}_r^{\mathrm{T}} = \mathbf{0}_{m\times n}$. The proof of $\mathcal{N}(\mathbf{Y}^{\mathrm{T}}) = \mathcal{S}_c(\mathbf{J}_u)$ follows the same lines and is, therefore, omitted here. □

Note that the null spaces of $\mathbf{Y}$ and $\mathbf{Y}^{\mathrm{T}}$ are zero if $\operatorname{rank}(\mathbf{Y}) = n$ or $\operatorname{rank}(\mathbf{Y}) = m$, respectively. Thus, in such cases the projection errors $\boldsymbol{\varepsilon}_e$ become trivially zero.

B.5.4 Dispersion, covariance, and correlation

The dispersion (cross-product), empirical covariance, and empirical correlation matrices of an $m \times q$ matrix $\mathbf{U}$ and an $m \times n$ matrix $\mathbf{Y}$ are defined as:

$$\begin{aligned}
\mathcal{D}(\mathbf{U}, \mathbf{Y}) &\equiv \mathbf{U}^{\mathrm{T}}\mathbf{Y} \\
\mathcal{C}(\mathbf{U}, \mathbf{Y}) &\equiv \frac{1}{m-1}\bar{\mathbf{U}}_c^{\mathrm{T}}\bar{\mathbf{Y}}_c = \frac{1}{m-1}\mathbf{U}^{\mathrm{T}}\mathbf{J}_c\mathbf{Y} \\
\mathcal{R}(\mathbf{U}, \mathbf{Y}) &\equiv \operatorname{diag}\left(\mathcal{C}(\mathbf{U}, \mathbf{Y})\right)^{-1/2} \mathcal{C}(\mathbf{U}, \mathbf{Y}) \operatorname{diag}\left(\mathcal{C}(\mathbf{U}, \mathbf{Y})\right)^{-1/2},
\end{aligned} \tag{B.40}$$

with $\operatorname{diag}(\mathbf{G})$ being a diagonal matrix containing the diagonal elements of $\mathbf{G}$. Note that, since $\mathbf{J}_c$ is a projection matrix, it follows that $\mathbf{J}_c^2 = \mathbf{J}_c$. For simplicity of notation, $\mathcal{D}(\mathbf{U}) = \mathbf{U}^{\mathrm{T}}\mathbf{U}$, $\mathcal{C}(\mathbf{U}) = \frac{1}{m-1}\mathbf{U}^{\mathrm{T}}\mathbf{J}_c\mathbf{U}$, and $\mathcal{R}(\mathbf{U}) = \operatorname{diag}\left(\mathcal{C}(\mathbf{U})\right)^{-1/2} \mathcal{C}(\mathbf{U}) \operatorname{diag}\left(\mathcal{C}(\mathbf{U})\right)^{-1/2}$.

C

Statistical methods

The following methods are briefly reviewed: (i) principal component analysis (PCA), (ii) multivariate linear calibration, (iii) factor analysis (FA), and (iv) methods for estimating the optimal structures of (a) PCA and FA models for "best" fit, and (b) PCA, calibration, and FA models for "best" prediction performance.

C.1 Principal component analysis (PCA)

PCA is a useful tool for data compression and information extraction of collinear or near-collinear data by finding combinations of *factors* that describe major trends in the data. This collinearity means that a data matrix will have some dominating types of variability that carry most of the available information. Redundancy and smaller noise variability can be removed by PCA.

Subsection C.1.1 formulates the problem of PCA, and Subsection C.1.2 provides a numerical solution using singular value decomposition (SVD).

C.1.1 Problem formulation

Let $\tilde{\mathbf{Y}}$ be an $m \times n$ noisy matrix with $\tilde{\mathbf{Y}} = [\tilde{\mathbf{y}}_1, \ldots, \tilde{\mathbf{y}}_n]$. PCA expresses the main information in the variables $\tilde{\mathbf{y}}_i$ $(i = 1, \ldots, n)$ by a lower number of new variables $\hat{\mathbf{t}}_a$ $(a = 1, \ldots, A)$, the so-called *principal components* (PCs) or *scores* of $\tilde{\mathbf{Y}}$, with the number of PCs (factors) A being $A < n$. The scores contain useful information on how the observations relate to each other. Mathematically, PCA relies upon a spectral decomposition (Lay, 1994) of the dispersion, empirical covariance or correlation of $\tilde{\mathbf{Y}}$ (see B.40). For simplicity of notation, only the decomposition based on the dispersion matrix (see B.40) of $\tilde{\mathbf{Y}}$ is studied next. The formulation of PCA using the covariance or the correlation of $\tilde{\mathbf{Y}}$ matrices (see B.40) are straightforward and, therefore, omitted here.

PCA relies on a linear model that decomposes $\tilde{\mathbf{Y}}$ into a sum of outer products of the scores $\mathbf{t}_a$ and the *loadings* $\mathbf{p}_a$ to give:

$$\tilde{\mathbf{Y}} = \mathbf{T}\,\mathbf{P}^{\mathrm{T}} + \tilde{\mathbf{F}}, \tag{C.1}$$

where $\tilde{\mathbf{F}}$ is an $m \times n$ random noise matrix including measurement noise and model inadequacies, $\mathbf{T}$ the $m \times A$ scores matrix, and $\mathbf{P}$ the $n \times A$ loadings matrix. Thus, with only A PCs collected, the data matrix $\tilde{\mathbf{Y}}$ can be *approximated* by

$$\hat{\mathbf{Y}} = \hat{\mathbf{T}}\,\hat{\mathbf{P}}^{\mathrm{T}}, \qquad \hat{\mathbf{F}} = \tilde{\mathbf{Y}} - \hat{\mathbf{Y}}, \tag{C.2}$$

where $\hat{\mathbf{T}}$ and $\hat{\mathbf{P}}$ are the $m \times A$ estimated scores matrix and the $n \times A$ estimated loading matrix, respectively, and $\hat{\mathbf{F}}$ the estimated residuals.

The (estimated) loadings contain useful information on how the variables relate to each other. Furthermore, they represent the first A orthonormal eigenvectors of the dispersion matrix $\tilde{\mathbf{Y}}^{\mathrm{T}}\tilde{\mathbf{Y}}$, which can be calculated successively by solving the following constrained optimization problem for $a = 1, \ldots, A$:

$$\hat{\mathbf{p}}_a = \arg\max_{\mathbf{p}_a} \mathbf{p}_a^{\mathrm{T}}(\tilde{\mathbf{Y}}^{\mathrm{T}}\tilde{\mathbf{Y}})\mathbf{p}_a, \tag{C.3}$$

subject to the constraints

$$\begin{aligned} \mathbf{p}_a^{\mathrm{T}}\mathbf{p}_a &= 1, \\ \mathbf{t}_a^{\mathrm{T}}\mathbf{t}_j &= 0, \quad \forall j > a, \\ \mathbf{t}_a &= \tilde{\mathbf{Y}}\,\mathbf{p}_a, \end{aligned} \tag{C.4}$$

i.e., the scores are forced to be mutually orthogonal to each other and the loadings are normalized to 1. From the maximum principle of real symmetric matrices (Keener, 1988), it follows from (C.3)–(C.4) that

$$(\tilde{\mathbf{Y}}^{\mathrm{T}}\tilde{\mathbf{Y}})\,\hat{\mathbf{p}}_a = \hat{\sigma}_a^2\,\hat{\mathbf{p}}_a \quad \text{with} \quad \hat{\sigma}_a^2 \equiv \hat{\mathbf{t}}_a^{\mathrm{T}}\hat{\mathbf{t}}_a, \qquad a = 1, \ldots, A, \tag{C.5}$$

where $\hat{\sigma}_a^2$ is the ath eigenvalue of the dispersion matrix with $\hat{\sigma}_1^2 \geq \hat{\sigma}_2^2 \geq \cdots \geq \hat{\sigma}_A^2$. Likewise, it can be shown that the estimated scores represent the first A eigenvectors of $\tilde{\mathbf{Y}}\,\tilde{\mathbf{Y}}^{\mathrm{T}}$, scaled to length $\hat{\sigma}_a$ ($a = 1, \ldots, A$).

Since the scores are related to the eigenvalues of the dispersion matrix by $\hat{\sigma}_a^2 \equiv \hat{\mathbf{t}}_a^{\mathrm{T}}\hat{\mathbf{t}}_a$, PCA explains parts of the data variability by the eigenvalues of the dispersion matrix. Because the eigenvalues are in descending order, the first pair $\{\hat{\mathbf{t}}_1, \hat{\mathbf{p}}_1\}$ captures the greatest amount of information (variability) of any pair of the decomposition in the sense of the maximization problem (C.3).

For $\tilde{\mathbf{Y}}$ and any pair $\{\hat{\mathbf{t}}_a, \hat{\mathbf{p}}_a\}$, due to $\hat{\mathbf{P}}^{\mathrm{T}}\hat{\mathbf{P}} = \mathbf{I}_A$, it follows from (C.2) that

$$\hat{\mathbf{T}} = \tilde{\mathbf{Y}}\,\hat{\mathbf{P}}, \tag{C.6}$$

i.e., the scores vector $\hat{\mathbf{t}}_a$ is a linear combination of the rows of $\tilde{\mathbf{Y}}$ weighted by the loading vector $\hat{\mathbf{p}}_a$. In other words, the loadings $\hat{\mathbf{p}}_a$ ($a = 1, \ldots, A$) represent the new orthogonal basis vectors for the row space of $\hat{\mathbf{Y}}$, and the elements of $\hat{\mathbf{t}}_a$ are the new coordinates (or projections) of the rows of $\mathbf{X}$ on the new basis vectors $\hat{\mathbf{p}}_a$ ($a = 1, \ldots, A$). This property gives rise to special algorithms for estimating eigenvalues and eigenvectors (see also next subsection).

C.1.2 Numerical solution by SVD

There exist several algorithms to perform PCA of $\tilde{\mathbf{Y}}$ such as the nonlinear iterative partial least-squares (NIPALS) algorithm (Wold, 1966) or SVD (see Appendix B.3). The latter is summarized next.

From (B.18), $\tilde{\mathbf{Y}}$ can be approximated by:

$$\hat{\mathbf{Y}} = \boldsymbol{\Upsilon}_A \, \boldsymbol{\Sigma}_A \, \boldsymbol{\Omega}_A^{\mathrm{T}}, \qquad \boldsymbol{\Sigma}_A = \operatorname{diag}\left(\sigma_1, \sigma_2, \ldots, \sigma_A\right), \tag{C.7}$$

where $\hat{\mathbf{Y}}$ is the $m \times n$ data matrix reconstructed by retaining the first A significant factors, $\boldsymbol{\Sigma}_A$ the A-dimensional diagonal singular value matrix composed of the first A singular values, and $\boldsymbol{\Upsilon}_A$ and $\boldsymbol{\Omega}_A$ unitary matrices of dimension the $m \times A$ and $n \times A$, respectively, corresponding to $\boldsymbol{\Sigma}_A$.

(C.7) can be related to (C.2) by

$$\begin{aligned} &\hat{\mathbf{Y}} = \hat{\mathbf{T}} \, \hat{\mathbf{P}}^{\mathrm{T}} \quad \text{with} \quad \hat{\mathbf{T}} = \boldsymbol{\Upsilon}_A \, \boldsymbol{\Sigma}_A, \quad \hat{\mathbf{P}} = \boldsymbol{\Omega}_A^{\mathrm{T}}, \\ &\hat{\mathbf{F}} = \tilde{\mathbf{Y}} - \hat{\mathbf{Y}} = \sum_{j=A+1}^{\min(m,n)} \sigma_j \boldsymbol{\upsilon}_j \boldsymbol{\omega}_j^{\mathrm{T}}. \end{aligned} \tag{C.8}$$

Since $\hat{\mathbf{T}}^{\mathrm{T}}\hat{\mathbf{T}} = \boldsymbol{\Sigma}_A^2$, the eigenvalues of the approximated dispersion matrix are the first A squared singular values of $\tilde{\mathbf{Y}}$. Thus, equations (B.18) and (C.8) can be used to perform PCA of $\tilde{\mathbf{Y}}$.

C.2 Calibration

Subsection C.2.1 reviews calibration based on a bilinear model, and Subsection C.2.2 presents two regression methods that are used herein to solve the calibration problem, namely PCR and SIMPLS.

C.2.1 Prediction

Let a *bilinear model* (Martens and Naes, 1989) link an n-dimensional input vector (e.g., an spectrum with n channels) with and a q-dimensional output vector (e.g., a concentration vector of q species) for m observations, and let $\tilde{\mathbf{Y}}$ be the corresponding $m \times n$ noisy input matrix and $\tilde{\mathbf{U}}$ the corresponding $m \times q$ noisy output matrix. Then,

$$\begin{aligned} \tilde{\mathbf{Y}} &= \mathbf{T} \, \mathbf{P}^{\mathrm{T}} + \tilde{\mathbf{F}}_Y, \\ \tilde{\mathbf{U}} &= \mathbf{T} \, \mathbf{Q}^{\mathrm{T}} + \tilde{\mathbf{F}}_U, \end{aligned} \tag{C.9}$$

with $\tilde{\mathbf{F}}_Y$ and $\tilde{\mathbf{F}}_U$ being random noise matrices including measurement noise and model inadequacies of dimension $m \times n$ and $m \times q$, respectively, $\mathbf{T}$ the $m \times A$ scores matrix, and $\mathbf{P}$ and $\mathbf{Q}$ loading matrices of dimension $(n \times A)$ and $(q \times A)$, respectively. Also, let $\mathcal{F}(\cdot)$ be a matrix function that guarantees

$$\mathbf{W} = \mathcal{F}(\tilde{\mathbf{Y}}, \tilde{\mathbf{U}}), \qquad \mathbf{T} = \tilde{\mathbf{Y}} \, \mathbf{W}, \tag{C.10}$$

where $\mathbf{W}$ is an $n \times A$ weighting matrix. Then, the closed-form calibration model (C.9) becomes

$$\tilde{\mathbf{U}} = \tilde{\mathbf{Y}}\,\mathbf{B} + \tilde{\mathbf{G}} \tag{C.11}$$

with $\mathbf{B}$ being a $q \times n$ regressor matrix and $\tilde{\mathbf{G}}$ an $m \times q$ matrix including measurement errors and model inadequacies.

Theorem C.1
Let the pair $\{\tilde{\mathbf{Y}}, \tilde{\mathbf{U}}\}$*, following the bilinear model* (C.9)–(C.10)*, be available for calibration. Then, the prediction of* $\mathbf{u}_n$ *for a given* $\tilde{\mathbf{y}}_n$ *following* (C.9)–(C.10) *is:*

$$\hat{\mathbf{u}}_n^{\mathrm{T}} = \tilde{\mathbf{y}}_n^{\mathrm{T}}\,\hat{\mathbf{B}}, \quad \textit{where} \quad \hat{\mathbf{B}} = \mathcal{I}(\tilde{\mathbf{Y}}, \tilde{\mathbf{U}})\,\tilde{\mathbf{U}} \quad \textit{with} \quad \mathcal{I}(\tilde{\mathbf{Y}}, \tilde{\mathbf{U}}) = \hat{\mathbf{W}}\,\hat{\mathbf{T}}^{+},\ \hat{\mathbf{T}} = \tilde{\mathbf{Y}}\,\hat{\mathbf{W}}, \tag{C.12}$$

where $\mathcal{I}(\tilde{\mathbf{Y}}, \tilde{\mathbf{U}})$ *indicates an inverse function that depends on the regression method used (see Subsection C.2.2 below).*

Proof:
Note that $\hat{\mathbf{W}}\,\mathbf{P}^{\mathrm{T}} \approx \mathbf{I}_n$, i.e., $\hat{\mathbf{W}}$ is an estimate of the pseudo–inverse of $\mathbf{P}^{\mathrm{T}}$, i.e., an estimate of $\mathbf{P}$. Then $\hat{\mathbf{T}} = \tilde{\mathbf{Y}}\,\hat{\mathbf{W}}$, and using the standard least-squares solution for $\mathbf{Q}$ gives $\hat{\mathbf{Q}}^{\mathrm{T}} = \hat{\mathbf{T}}^{+}\,\tilde{\mathbf{U}}$, where $\hat{\mathbf{T}}^{+} = (\hat{\mathbf{T}}^{\mathrm{T}}\,\hat{\mathbf{T}})^{-1}\,\hat{\mathbf{T}}^{\mathrm{T}}$. The internal prediction of $\tilde{\mathbf{U}}$ is

$$\hat{\mathbf{U}} = \hat{\mathbf{T}}\,\hat{\mathbf{Q}}^{\mathrm{T}} = \tilde{\mathbf{Y}}\,\left(\hat{\mathbf{W}}\,\hat{\mathbf{Q}}^{\mathrm{T}}\right) \equiv \tilde{\mathbf{Y}}\,\hat{\mathbf{B}} \tag{C.13}$$

where $\hat{\mathbf{B}} = \hat{\mathbf{W}}\,\hat{\mathbf{Q}}^{\mathrm{T}} = \hat{\mathbf{W}}\,\hat{\mathbf{T}}^{+}\,\tilde{\mathbf{U}}$.

Since a new measurement $\tilde{\mathbf{y}}_n$ satisfies (C.9), $\tilde{\mathbf{y}}_n^{\mathrm{T}} = \mathbf{t}_n^{\mathrm{T}}\,\mathbf{P}^{\mathrm{T}} + \mathbf{f}_{A,n}^{\mathrm{T}}$. Then, (C.12) can be used for prediction of $\hat{\mathbf{u}}_n$ from $\tilde{\mathbf{y}}_n$, since

$$\tilde{\mathbf{y}}_n^{\mathrm{T}}\,\hat{\mathbf{B}} \approx \mathbf{t}_n^{\mathrm{T}}\,\mathbf{P}^{\mathrm{T}}\,\hat{\mathbf{W}}\,\hat{\mathbf{Q}}^{\mathrm{T}} = \mathbf{t}_n^{\mathrm{T}}\,\hat{\mathbf{Q}}^{\mathrm{T}} = \hat{\mathbf{u}}_n^{\mathrm{T}}. \tag{C.14}$$

□

The assumption that both the calibration and the new data follow the bilinear model (C.9)–(C.10) implies that $\tilde{\mathbf{y}}_n \in \mathcal{S}_r\left(\tilde{\mathbf{Y}}\right)$. Note that condition $\tilde{\mathbf{y}}_n \in \mathcal{S}_r\left(\tilde{\mathbf{Y}}\right)$ can be checked by computing the least-squares error $\varepsilon_p(\tilde{\mathbf{y}}_n, \tilde{\mathbf{Y}})$ of $\tilde{\mathbf{y}}_n$ on the row space of $\tilde{\mathbf{Y}}$:

$$\varepsilon_p(\tilde{\mathbf{y}}_n, \tilde{\mathbf{Y}}) \equiv \left\| \tilde{\mathbf{y}}_n^{\mathrm{T}}\,\left(\mathbf{I}_m - \tilde{\mathbf{Y}}^{+}\,\tilde{\mathbf{Y}}\right) \right\|_2. \tag{C.15}$$

ε_p^2 is also known as the squared-sum-of-errors or residuals used to test on lack of fit. For the simultaneous test of K_n new spectra, (C.15) becomes

$$\varepsilon_p(\tilde{\mathbf{Y}}_n, \tilde{\mathbf{Y}}) \equiv \left\| \tilde{\mathbf{Y}}_n\,\left(\mathbf{I}_m - \tilde{\mathbf{Y}}^{+}\,\tilde{\mathbf{Y}}\right) \right\|_{\mathrm{fro}}, \tag{C.16}$$

where $\|\mathbf{G}\|_{\mathrm{fro}}$ is the Frobenius norm of $\mathbf{G}$, defined as $\|\mathbf{G}\|_{\mathrm{fro}} = (\mathrm{tr}\,(\mathbf{G}^{\mathrm{T}}\mathbf{G}))^{1/2}$ with $\mathrm{tr}\,(\cdot)$ being the trace of a matrix.

Remark C.2 (Additive nonlinearities)
Additive nonlinearities for the input matrix can be handled in calibration. Consider

$$\tilde{\mathbf{Y}}_{nl} = \mathbf{Y} + \mathbf{T}_{nl}\mathbf{P}_{nl}^{\mathrm{T}} + \tilde{\mathbf{F}}_Y \tag{C.17}$$

with $\tilde{\mathbf{Y}}_{nl}$ being the $m \times n$ measured input matrix, $\mathbf{Y}$ the $m \times n$ matrix containing useful signal, $\mathbf{T}_{nl}$ and $\mathbf{P}_{nl}$ the $m \times A_{nl}$ scores matrix and the $n \times A_{nl}$ loading matrix, respectively, corresponding to the additive nonlinearities, and A_{nl} the number of factors of the additive nonlinearities. Theorem C.1 holds also for such data with the redefinition $\tilde{\mathbf{Y}} := \tilde{\mathbf{Y}}_{nl}$. Thus, as long as the condition $\tilde{\mathbf{y}}_n \in \mathcal{S}_r\left(\tilde{\mathbf{Y}}_{nl}\right)$ is guaranteed, the additive nonlinearities will not interfere with the prediction of $\mathbf{u}_n$ from a new input $\tilde{\mathbf{y}}_n$.

C.2.2 Regression methods

Calibration in Theorem C.1 is addressed under the assumption that an estimate $\hat{\mathbf{W}}$ obeying (C.10) is available. In the next two paragraphs, $\hat{\mathbf{W}}$ and $\mathcal{I}(\tilde{\mathbf{Y}}, \tilde{\mathbf{U}})$ are studied for PCR and SIMPLS, the two regression methods that are used throughout this dissertation.

C.2.2.1 PCR

PCR consists of two steps (Jackson, 1991): (S1) PCA of $\tilde{\mathbf{Y}}$ by retaining A factors, and (S2) multiple regression of the columns of $\tilde{\mathbf{U}}$ on the scores matrix $\hat{\mathbf{T}}$. The two steps are summarized for the calibration pair $\{\tilde{\mathbf{Y}}, \tilde{\mathbf{U}}\}$:

(S1) From the solution of the constrained optimization problem (C.3)–(C.4) for all $a = 1, \ldots, A$:

$$\begin{aligned} &\max_{\mathbf{p}_a} \mathbf{p}_a^{\mathrm{T}}(\tilde{\mathbf{Y}}^{\mathrm{T}}\tilde{\mathbf{Y}})\,\mathbf{p}_a \\ &\text{s.t. } \|\mathbf{p}_a\|_2 = 1 \\ &\qquad \mathbf{t}_a^{\mathrm{T}}\mathbf{t}_j = 0 \quad \forall j > a. \end{aligned} \tag{C.18}$$

$\hat{\mathbf{W}}$ is identified as $\hat{\mathbf{W}}(\tilde{\mathbf{Y}}) = \hat{\mathbf{P}} = \boldsymbol{\Omega}$.

(S2) The regression step is given in (C.12). Since $\hat{\mathbf{T}}^{+} = \boldsymbol{\Sigma}_A^{-1}\boldsymbol{\Upsilon}_A^{\mathrm{T}}$, the inverse function becomes $\mathcal{I}(\tilde{\mathbf{Y}}) = \hat{\mathbf{W}}\,\hat{\mathbf{T}}^{+} = \boldsymbol{\Omega}_A\,\boldsymbol{\Sigma}_A^{-1}\,\boldsymbol{\Upsilon}_A^{\mathrm{T}}$, which is the pseudo-inverse of $\hat{\mathbf{Y}}$ (see also Appendix B.4).

C.2.2.2 SIMPLS

In SIMPLS, $\mathcal{I}(\tilde{\mathbf{Y}}, \tilde{\mathbf{U}})$ is obtained from an iterative algorithm (de Jong, 1993) that solves the following constrained optimization problem for all $a = 1, \ldots, A$:

$$\begin{aligned} &\max_{\mathbf{w}_a, \mathbf{q}_a} \mathbf{w}_a^{\mathrm{T}}(\tilde{\mathbf{Y}}^{\mathrm{T}}\tilde{\mathbf{U}})\,\mathbf{q}_a \\ &\text{s.t. } \|\mathbf{w}_a\|_2 = 1 \\ &\qquad \|\mathbf{q}_a\|_2 = 1 \\ &\qquad \mathbf{t}_a^{\mathrm{T}}\mathbf{t}_j = 0 \quad \forall j > a, \end{aligned} \tag{C.19}$$

where $\mathbf{t}_a = \tilde{\mathbf{Y}}\mathbf{w}_a$. It can be seen from (C.19) that the cross-product matrix $\tilde{\mathbf{Y}}^{\mathrm{T}}\tilde{\mathbf{U}}$ is maximized. Note that without the constraint $\mathbf{t}_a^{\mathrm{T}}\mathbf{t}_j = 0$, there is only one solution: $\mathbf{w}_1$ and $\mathbf{q}_1$ are the first left and right singular vectors of the $n \times q$ cross-product matrix $\tilde{\mathbf{Y}}^{\mathrm{T}}\tilde{\mathbf{U}}$. However, with the constraint a set of orthogonal factors of $\tilde{\mathbf{Y}}$ is obtained.

If there is only one output (univariate output), then $\mathbf{q}_a = 1$ and, thus, the constraint $\|\mathbf{q}_a\|_2 = 1$ in (C.19) is trivially satisfied. In this case, the solution to (C.19) is non-iterative.

C.3 Factor analysis

The term "factor analysis" (FA) is quite ambiguous in the literature. In chemometrics, FA or *multivariate curve-resolution method* means a procedure consisting of two steps: (i) PCA, (ii) rotation of scores and loading matrices into physically-meaningful ones using prior knowledge. In statistics, FA means the investigation of correlations of random variables (Hamer, 1989). The statistical FA is sometimes called "orthodox FA" or "nonlinear FA" by physicists or chemists (Paatero and Tapper, 1993). Here, the chemometric definition of FA is used.

The two FA methods considered in this dissertation are target factor analysis (TFA; Malinowski, 1991; Bonvin and Rippin, 1990) and AR. First, PCA is explained in the light of FA. Then, TFA and AR are briefly reviewed.

C.3.1 PCA

Let $\mathbf{Y}$ be an $m \times n$ noise-free matrix of rank A that is described by a linear factorization, i.e., it can be described by two physical matrices, $\mathfrak{T}$ and $\mathfrak{P}$, of dimension $m \times A$ and $n \times A$, respectively:

$$\mathbf{Y} = \mathfrak{T}\,\mathfrak{P}^{\mathrm{T}}. \tag{C.20}$$

In the noisy case, PCA (see Appendix C.1) estimates the (approximate) spaces for $\mathfrak{T}$ and $\mathfrak{P}$:

$$\begin{aligned} \tilde{\mathbf{Y}} &\overset{\text{PCA}}{=} \mathbf{T}\,\mathbf{P}^{\mathrm{T}} + \tilde{\mathbf{F}} \\ \hat{\mathbf{Y}} &= \hat{\mathbf{T}}\,\hat{\mathbf{P}}^{\mathrm{T}}. \end{aligned} \tag{C.21}$$

The column space of $\hat{\mathbf{Y}}$ corresponds to that of $\hat{\mathbf{T}}$. The row space of $\hat{\mathbf{Y}}$ corresponds to the column space of $\hat{\mathbf{P}}$.

C.3.2 Rotation

First, the problem of rotation is formulated, and the concepts of intensity and rotational ambiguities are introduced. In Subsection C.3.2.2, the concepts of selectivity and zero-factor windows are revisited. In Subsection C.3.2.3, TFA will be used to resolve the rotational ambiguity, and in many cases the intensity ambiguity. Subsection C.3.2.4 presents AR. Depending on the constraints available, the intensity ambiguity can be resolved in addition to the rotational ambiguity.

C.3.2.1 Problem formulation

In the second step of FA, an $A \times A$ rotation matrix, $\mathbf{R}$, is determined so as to estimate the physically-meaningful quantities $\hat{\mathfrak{T}} = \hat{\mathbf{T}}\,\mathbf{R}$ and $\hat{\mathfrak{P}}^{\mathrm{T}} = \mathbf{R}^{-1}\hat{\mathbf{P}}^{\mathrm{T}}$.

There are two classes of ambiguities associated with curve-resolution methods, namely *intensity* and *rotational ambiguities* (see Tauler *et al.*, 1995, and the references therein). Most curve resolution methods do not resolve the intensity ambiguity, that is, each column of the estimated matrices $\hat{\mathfrak{T}}$ and $\hat{\mathfrak{P}}$ will be scaled by some unknown factor. This is not a serious problem in qualitative identification (e.g., spectral identification, finger printing), but is a serious problem in quantification.

However, a somehow more serious problem is the rotational ambiguity, which generally occurs when two or more linearly-independent factors are present. Each factor or column of $\hat{\mathbf{T}}$ and $\hat{\mathbf{P}}$ will be an unknown linear combination of the true columns of $\mathfrak{T}$ and $\mathfrak{P}$, respectively.

C.3.2.2 Selectivity and zero-factor windows

Let assume that the complete data matrix $\mathbf{Y}$ is described (globally) by A factors. No rotational ambiguity arises for any submatrix of $\mathbf{Y}$ that contains a single factor (*selectivity*) since the submatrix is of rank one. Thus, it can be uniquely decomposed into an outer product of two vectors. Alternatively, when a particular factor is absent in a submatrix of $\mathbf{Y}$ and present in the remaining part of $\mathbf{Y}$ (*zero-factor window*), then the rotational ambiguity for that particular factor can also be resolved. Selectivity and zero-factor windows can be viewed as those regions where one or several of the A factors become locally linearly dependent. These regions can be determined by evolving factor analysis (see Appendix C.4.1.3 below).

For those factors where neither a zero-factor window nor a selective region is present, the extent of the rotational ambiguity can be considerably reduced if appropriate constraints are used. Note that these constraints are dependent on the type of data under investigation.

C.3.2.3 Target factor analysis (TFA)

In TFA, A targets $\mathfrak{p}_{tar}$ of dimension n (i.e., possible rows of $\mathfrak{P}^{\mathrm{T}}$ that are assumed to be available from prior knowledge) are tested individually to check whether or not they lie in the row space of $\hat{\mathbf{P}}^{\mathrm{T}}$. Let $\mathbf{R}_i \equiv \mathbf{R}^{-1}$. Then, the matrices $\mathbf{R}$ and $\mathfrak{T}$ are estimated as:

$$\hat{\mathbf{R}}_i = \mathfrak{P}^{\mathrm{T}}\,\hat{\mathbf{P}}, \quad \hat{\mathbf{R}} = \hat{\mathbf{R}}_i^{-1}, \quad \hat{\mathfrak{T}} = \hat{\mathbf{T}}\,\hat{\mathbf{R}}. \tag{C.22}$$

Note that owing to orthonormality of the columns of $\hat{\mathbf{P}}$, $(\hat{\mathbf{P}}^{\mathrm{T}})^{+}$ was written as $\hat{\mathbf{P}}$ in (C.22). Herein, the projection error (see C.15) is used as the acceptance criterion in the target testing step (Bonvin and Rippin, 1990), i.e., ε_p is the projection error (C.15) of $\mathfrak{p}_{tar}$ on the row space of $\hat{\mathbf{P}}^{\mathrm{T}}$:

$$\varepsilon_p(\mathfrak{p}_{tar}, \tilde{\mathbf{Y}}) = \left\| \mathfrak{p}_{tar}^{\mathrm{T}} \left(\mathbf{I}_n - \tilde{\mathbf{Y}}^{+}\tilde{\mathbf{Y}} \right) \right\|_2 \approx \left\| \mathfrak{p}_{tar}^{\mathrm{T}} \left(\mathbf{I}_n - \hat{\mathbf{P}}\,\hat{\mathbf{P}}^{\mathrm{T}} \right) \right\|_2 . \tag{C.23}$$

Let δ be a threshold to be defined. If $\varepsilon_p(\mathfrak{p}_{tar}, \tilde{\mathbf{Y}}) < \delta$, then the target $\mathfrak{p}_{tar}$ is accepted. Note that in the noise-free case,

$$\varepsilon_p(\mathfrak{p}_{tar}, \mathbf{Y}) = 0 \quad \Leftrightarrow \quad \mathfrak{p}_{tar} \in \mathcal{S}_r\left(\hat{\mathbf{P}}^{\mathrm{T}}\right) \quad \Leftrightarrow \quad \mathfrak{p}_{tar} \notin \mathcal{N}\left(\hat{\mathbf{P}}^{\mathrm{T}}\right). \tag{C.24}$$

C.3.2.4 Alternating regression (AR)

AR consists of iteratively solving a constrained multivariate linear least-squares problem and its dual (see Appendix B.5.2 for the solution to unconstrained multiple linear least-squares problem).

Given the linear statistical model:

$$\mathcal{U} = \mathcal{Y}\mathbf{B} + \tilde{\mathbf{G}}, \tag{C.25}$$

with $\mathcal{U}$ and $\mathcal{Y}$ being matrices of dimension $m \times n$ and $m \times A$, respectively, $\mathbf{B}$ the $A \times n$ regressor matrix, and $\tilde{\mathbf{G}}$ the $m \times n$ random matrix including model inadequacies. Then, the constrained multivariate linear least-squares problem can be formulated as:

$$\begin{aligned} &\min_{\mathbf{B}} \|\hat{\mathcal{U}}(\mathbf{B}) - \mathcal{U}\|_2 \\ &\text{s.t. } \mathcal{C}(\mathbf{B}), \end{aligned} \tag{C.26}$$

where $\mathcal{C}(\cdot)$ denotes the constraints on $\mathbf{B}$, $\hat{\mathcal{U}}(\mathbf{B})$ the estimate of $\mathcal{U}$ based on (C.25), and $\mathcal{U}$ the $m \times n$ approximated data matrix. Note that $\mathcal{U}$ and $\mathcal{Y}$ are known, and $\mathbf{B}$ and $\tilde{\mathbf{G}}$ unknown.

Let $\mathfrak{T}^0$ denote the initial estimate of $\mathfrak{T}$. Then, for the first iteration step of AR ($i = 1$), the first least-squares problem is computed with $\mathcal{U} := \hat{\mathbf{Y}}$, $\mathcal{Y} := \mathfrak{T}^0$, and $\mathbf{B} := (\mathfrak{P}^{\mathrm{T}})^1$, where $\mathfrak{P}^1$ is the estimate of $\mathfrak{P}$ in the first iteration step. Thus, the solution of

$$\begin{aligned} &\max_{\mathfrak{P}} \|\hat{\mathcal{U}}(\mathfrak{P}^1) - \mathcal{U}\|_2 \\ &\quad \text{s.t. } \mathcal{U} = \mathfrak{T}^0 \left(\mathfrak{P}^{\mathrm{T}}\right)^1 + \tilde{\mathbf{G}} \\ &\qquad \mathcal{C}(\mathfrak{P}^1) \end{aligned} \tag{C.27}$$

is $\hat{\mathfrak{P}}^1$. The second least-squares problem of the first iteration step is computed with $\mathcal{U} := \hat{\mathbf{Y}}^{\mathrm{T}}$, $\mathcal{Y} := \hat{\mathfrak{P}}^1$, and $\mathbf{B} := (\mathfrak{T}^{\mathrm{T}})^1$, where $\mathfrak{T}^1$ is the estimate of $\mathfrak{T}$ in the first iteration step. Thus, the solution of

$$\begin{aligned} &\max_{\mathfrak{T}} \|\hat{\mathcal{U}}(\mathfrak{T}^1) - \mathcal{U}\|_2 \\ &\quad \text{s.t. } \mathcal{U} = \hat{\mathfrak{P}}^1 \left(\mathfrak{T}^{\mathrm{T}}\right)^1 + \tilde{\mathbf{G}} \\ &\qquad \mathcal{C}(\mathfrak{T}^1) \end{aligned} \tag{C.28}$$

is $\hat{\mathfrak{T}}^1$. For the ith iteration step of AR, the matrices $\mathcal{U}$, $\mathcal{Y}$, and $\mathbf{B}$ corresponding to the two constrained multivariate linear least-squares problems are given in Table C.1.

Table C.1. Alternating regression (AR) at the ith iteration step: matrices corresponding to (α) a constrained multivariate linear least-squares problem and (β) its dual.

Problem	$\mathcal{U}$	$\mathcal{Y}$	$\mathbf{B}$
(α)	$\hat{\mathbf{Y}}$	$\hat{\mathfrak{T}}^{i-1}$	$(\mathfrak{P}^{\mathrm{T}})^i$
(β)	$\hat{\mathbf{Y}}^{\mathrm{T}}$	$\hat{\mathfrak{P}}^i$	$(\mathfrak{T}^{\mathrm{T}})^i$

C.4 Structure selection

In this section, methods for estimating the optimal structures of (a) PCA and FA models for "best" fit, and (b) PCA, calibration, and FA models for "best" prediction performance are examined. These include methods for pseudo-rank estimation (see Subsection C.4.1 below) and prediction-based methods (see Subsection C.4.2 below).

C.4.1 Methods for pseudo-rank estimation

Methods for pseudo-rank estimation are important in PCA and FA of a matrix $\mathbf{Y}$. These methods can be applied on the complete matrix $\mathbf{Y}$ (see Appendix C.4.1.2 below), or on submatrices of $\mathbf{Y}$ (local method; see Appendix C.4.1.3).

C.4.1.1 Problem formulation

The pseudo-rank A of a matrix is defined as the rank of a matrix in the absence of measurement error. Pseudo-rank estimation is a major problem in multivariate data analysis (Faber *et al.*, 1994). Malinowski (1991) classifies the methods for pseudo-rank estimation into two categories:

(1) Methods based on the knowledge of the experimental error. These methods rely on the "external" comparison of the eigenvalues with experimental error. They can be further divided into methods based on the knowledge of the mean and/or the variance of the experimental error.
(2) Methods requiring no knowledge of the experimental error. These methods rely on the fact that the signal-to-noise ratio is significantly large and, therefore, the relative magnitude between the signal and noise eigenvalues is high ("internal" comparison). Empirical rules are usually applied (empirical methods).

The experimental error usually includes measurement noise and model inadequacies. Since no information about the model inadequacies is available, empirical methods for pseudo-rank estimation are used in this study. A clear drawback of these methods is that they all depend on a threshold value, ε_A, that must be defined properly based on either empirical or statistical criteria. In the absence of such a threshold value,

it is common practice to compare pseudo-rank estimates using various pseudo-rank estimation methods (Malinowski, 1978; Hopke, 1989).

Let $\tilde{\mathbf{Y}}$ denote an $m \times n$ noisy data matrix. Most of the pseudo-rank estimation methods rely on deleting some of the smallest singular values of $\tilde{\mathbf{Y}}$, assuming that they contribute only to noise in $\tilde{\mathbf{Y}}$. However, this assumption is only true in some special cases since the measurement error[1] (ME; $\tilde{\mathbf{F}}$ in C.21) affects both the embedded error (IE) associated with the first A dominant singular values of $\tilde{\mathbf{Y}}$ and the extracted error (EE) associated with the remaining $(\underline{A} - A)$ singular values, where $\underline{A} \equiv \min(m, n)$. Under the assumption of equal error variance associated with each element of $\tilde{\mathbf{Y}}$, the following theoretical relationship exists between the standard deviations of the various error terms (Malinowski, 1991):

$$(\text{ME})^2 = (\text{IE})^2 + (\text{EE})^2, \qquad \text{IE} = \sqrt{A/\underline{A}}\ \text{ME},\ \text{EE} = \sqrt{(\underline{A} - A)/\underline{A}}\ \text{ME}. \tag{C.29}$$

From this equation, it can be seen that the embedded error is negligible only for $A \ll \underline{A}$. Otherwise, analysis results based on the reconstructed data matrix $\hat{\mathbf{Y}}$ in (C.2) are distorted.

For data exhibiting smooth responses such as concentration profiles from chemical reaction systems, the pseudo-rank can be estimated by looking at the scores or/and loadings: scores or/and loadings that represent real signal will be smooth, whereas those representing noise will be scattered.

C.4.1.2 Pseudo-rank estimation

Only those methods that are used in this work are shortly revisited. They include methods based on: (i) the singular values using visual inspection, the Scree test, or the reduced F-test, and (ii) the singular vectors using the autocorrelation.

Let $\text{prank}\left(\tilde{\mathbf{Y}}\right)$ denote the pseudo-rank of $\tilde{\mathbf{Y}}$. The estimation procedure of the pseudo-rank A of $\tilde{\mathbf{Y}}$ is generally formulated as: Find A such that

$$\Delta(a, \tilde{\mathbf{Y}}) \leq \varepsilon_A, \quad \forall a = 1, \ldots, A, \tag{C.30}$$

where Δ is a measure that depends on the ath factor.

Visual Inspection. By plotting the singular values σ_a of $\tilde{\mathbf{Y}}$ (i.e., $\Delta(a, \tilde{\mathbf{Y}}) = \sigma_a$) with respect to the increasing index, the pseudo-rank is visually found/estimated as the cut-off point for the Ath singular value with $\sigma_A \leq \varepsilon_A$. It should be noted that the magnitude and the spacing of the singular values is strongly scale-dependent.

Scree Test. According to Cattell (1966), the pseudo-rank is based on the residual percent variance (RPV) $\text{RPV}(a, \tilde{\mathbf{Y}})$ of $\tilde{\mathbf{Y}}$ (i.e., $\Delta(a, \tilde{\mathbf{Y}}) = \text{RPV}(a, \tilde{\mathbf{Y}})$) computed as:

$$\text{RPV}(a, \tilde{\mathbf{Y}}) \equiv 100\, \frac{\sum_{j=a}^{\underline{A}} \sigma_j^2(\tilde{\mathbf{Y}})}{\sum_{j=1}^{\underline{A}} \sigma_j^2(\tilde{\mathbf{Y}})}, \tag{C.31}$$

[1] 'Error' and 'uncertainty' are used as interchangeable terms.

where σ_j is the jth singular value of $\tilde{\mathbf{Y}}$.

Reduced factor F-test. The reduced factor F-test from Malinowski (1988) is originally based on the assumption of normally distributed measurement noise. However, it is also often used as empirical method: the pseudo-rank is based on the reduced F-statistics $F(1, \underline{A} - a)$ of $\tilde{\mathbf{Y}}$ (i.e., $\Delta(a, \tilde{\mathbf{Y}}) = F(1, \underline{A} - a)$) computed as:

$$F(1, \underline{A} - a) = \frac{\sum_{j=a+1}^{\underline{A}} (\max(m, n) - j + 1)\,(\underline{A} - j + 1)}{(\max(m, n) - a + 1)(\underline{A} - a + 1)} \frac{\sigma_a^2}{\sum_{j=a+1}^{\underline{A}} \sigma_j^2} \tag{C.32}$$

with $\varepsilon_A := 1 - \alpha$ and α the quantile to be specified.

Autocorrelation. According to Shrager and Hendler (1982), the pseudo-rank is based on the autocorrelation $\mathrm{AUTO}(\boldsymbol{y}_a)$ of the scores or loading vectors $\boldsymbol{y}_a$ of $\tilde{\mathbf{Y}}$ (i.e., $\Delta(a, \tilde{\mathbf{Y}}) = \mathrm{AUTO}(\boldsymbol{y}_a)$) computed as:

$$\mathrm{AUTO}(\boldsymbol{y}_a) = \sum_{k=2}^{\check{m}} \mathfrak{y}_a(k)\, \mathfrak{y}_a(k-1), \tag{C.33}$$

where $\mathfrak{y}_a(k)$ is the kth element of $\boldsymbol{y}_a$, and $\check{m}$ the corresponding dimension of $\boldsymbol{y}_a$ (either m or n). Shrager and Hendler (1982) proposed $\varepsilon_A = 0.5$ as a threshold value to test on smoothness.

It is well known that the autocorrelation is the energy of the low-frequency part of $\boldsymbol{y}_a$ and, thus, it is ideally suited to measure smoothness. Since singular-vector matrices are unitary matrices, the maximal value of $\mathrm{AUTO}(\boldsymbol{y}_a)$ is 1, a case that occurs when $\boldsymbol{y}_a$ is a horizontal line (the extreme case of a smooth signal).

C.4.1.3 Local methods

The original version of evolving factor analysis (EFA; Gampp *et al.*, 1985) is based on repetitive SVD analysis of submatrices of a data matrix that is gathered from an ordered (evolutionary) process. Usually, one looks at the evolution of the singular values to comment on the local pseudo-rank (an integer). However, EFA can generally be based on a repetitive *pseudo-rank analysis* rather than SVD analysis, where any pseudo-rank estimation method can be used.

In *forward EFA*, pseudo-rank analyses are performed on a series of matrices constructed by successively adding observations to the previous matrix in an ordered sequence. When a significant factor appears, the pseudo-rank will increase. *Backward EFA* is initiated by starting the pseudo-rank analysis with the last few observations and systematically adding observations in the reverse order of collection. The resulting pseudo-rank analysis maps out the disappearance of a significant factor (i.e., appearance in the reverse order). In window evolving factor analysis (WEFA; Keller and Massart, 1991; Malinowski, 1992), pseudo-rank analysis is performed on a window of $\mathcal{w}$ observations that slides over the entire matrix. WEFA can also be performed simultaneously in both dimensions of $\tilde{\mathbf{Y}}$ with a two-dimensional window of size $\mathcal{w}_m \times \mathcal{w}_n$. The results can be visualized in a contour plot or *local rank map* (Geladi and Wold, 1987).

C.4.2 Prediction-based structure selection

C.4.2.1 Prediction error

A prediction error can be defined in various ways depending on the available data and the linear statistical model used (PCA, calibration formulation, regression method, etc.). Common error measures include:

(1) Root mean squared error (RMS). It is defined for the ith output as

$$\mathrm{RMS}_i \equiv \frac{1}{m} \|\hat{\mathbf{f}}_{U,i}\|_2 \quad \forall i = 1, \ldots, q, \tag{C.34}$$

where $\hat{\mathbf{f}}_{U,i}$ are the residuals of the ith output for m observations defined as:

$$\hat{\mathbf{f}}_{U,i} \equiv \tilde{\mathbf{u}}_i - \hat{\mathbf{u}}_i \quad \forall i = 1, \ldots, q, \tag{C.35}$$

with $\hat{\mathbf{u}}_i$ being the m-dimensional vector of the ith output predicted from the model.

(2) Average additive prediction error or *bias* estimated by:

$$\widehat{\mathrm{EE}}_i \equiv \frac{1}{m} \sum_{k=1}^{m} \hat{f}_{U,i}(k), \tag{C.36}$$

(3) Prediction intervals for each output.

The first two error measures are usually estimated by performing predictions on observations for which the output values are known. Both measures give an *average prediction error* that is characteristic for the calibration model rather than for an individual observation (Faber and Kowalski, 1997). This is in contrast to methods that estimate the prediction intervals for *individual* observations and that do not require the knowledge of output values. Such methods include Monte–Carlo techniques such as the bootstrap (Efron and Tibshirani, 1986). For PCR and PLS, analytical expressions for the prediction intervals have been derived by Phatak *et al.* (1993); Denham (1997); Faber and Kowalski (1997). Since in Section 5.3.1 the validity of the calibration model is the central focus rather than the prediction intervals for the individual observations, it suffices to consider the first two error measures (the RMS and the bias), which are the subject of the next paragraph.

The estimated prediction (RMS) and interference errors (underlying bias) are linked *via* the following relationship (Martens and Naes, 1989):

$$\mathrm{RMS}_i^2 = \widehat{\mathrm{IE}}_i^2 + \widehat{\mathrm{EE}}_i^2, \tag{C.37}$$

where $\widehat{\mathrm{IE}}_i^2$ is the squared estimation error or residual variance of the ith output defined as:

$$\widehat{\mathrm{IE}}_i^2 = \frac{1}{m} \sum_{k=1}^{m} \left(\hat{f}_{U,i}(k) - \widehat{\mathrm{EE}}_i \right)^2. \tag{C.38}$$

Note that, since the extracted (EE) and embedded errors (IE) in PCA (see C.29) are closely related to the bias and the residual standard deviation in regression, the same symbols are used. Furthermore, $\widehat{\mathrm{IE}}_i$ is also known as the standard error of prediction (SEP) where m^{-1} is replaced by $(m^{-1}-1)$.

It is important to note that the RMS in (C.34) is an estimate of the true RMS, since the predicted output values $\hat{\mathbf{u}}_i$ are compared in (C.35) with the noisy reference output values $\tilde{\mathbf{u}}_i$ instead of the noise-free ones $\mathbf{u}_i$. According to DiFoggio (1995), the RMS defined in (C.34) is called the *apparent* RMS (see C.35) as opposed to the *actual* RMS obtained for noise-free reference outputs.

To facilitate the comparison between (i) outputs with values of different magnitude and (ii) different data sets, the normalized RMS, bias, and estimation error are used herein:

$$\mathrm{NRMS}_i \equiv \mathrm{RMS}_i/\|\mathbf{u}_i\|_2, \quad \widehat{\mathrm{NEE}}_i \equiv \widehat{\mathrm{EE}}_i/\|\mathbf{u}_i\|_2, \quad \widehat{\mathrm{NIE}}_i \equiv \widehat{\mathrm{IE}}_i/\|\mathbf{u}_i\|_2 \quad \forall i = 1,\ldots,q. \tag{C.39}$$

Internal and external validation. The two most commonly used procedures for the determination of the RMS and the bias are *cross validation* and *prediction testing* on an independent data set (Martens and Naes, 1989).

M-block cross validation estimates statistical models by leaving out distinct segments of M observations for which the output values are predicted. This step is repeated M times. From the differences between the predicted and the output observations left out, prediction errors are estimated ($\mathrm{NRMS}_{\mathrm{CV},i}$ and $\mathrm{NEE}_{\mathrm{CV},i}$). Then, all calibration observations are used to estimate the final statistical model and prediction errors ($\mathrm{NRMS}_{\mathrm{C},i}$ and $\mathrm{NEE}_{\mathrm{C},i}$). Since cross validation estimates prediction errors based on observations present in the model-building set (e.g., the calibration set in calibration), it is also called *internal validation*. For data with correlated observations (e.g. data from chemical reaction systems), common practice has shown that M can be chosen larger than 1 (Wise, 1991).

Note that the segments in internal validation must be chosen such that each segment and the remaining model-building observations span the same space. Similarly, the test set must be chosen such that test set and the model-building set span the same space.

For a set of observations *not present* in the calibration set, prediction errors $\mathrm{NRMS}_{\mathrm{N},i}$ and $\mathrm{NEE}_{\mathrm{N},i}$ are obtained from the statistical model (*external validation*). Since external validation more closely assesses the predictive ability of the calibration model for new observations, it is obviously better than internal validation. However, external validation is rather wasteful, since the observations used for testing are not available to build the calibration model.

C.4.2.2 Model structure selection

In PCA or regression techniques such as PCR or SIMPLS, one free parameter, the number of factors A retained in the model, must be chosen based on stopping rules for model structure selection (Martens and Naes, 1989). Since each factor models an

independent contribution to the data (e.g., an interfering phenomenon in the spectral data), the number of factors required depends on the number of independent contributions to the data. If the goal is "best" prediction, the stopping rule must find an optimal number of factors, A^*, for which the prediction error of the calibration model reaches a minimum.

According to (C.37), the prediction error ($\mathrm{RMS_n}$ or $\mathrm{RMS_{cv}}$) is composed of two main contributions: the bias ($\widehat{\mathrm{EE}}_i$) and the estimated variance ($\widehat{\mathrm{IE}}_i^2$). The bias can be interpreted as the systematic error due to unmodeled useful signal, and the estimated variance is caused by random measurement noise. The two contributions to the prediction error have opposite trends with increasing number of factors. This is illustrated schematically in Figure C.1 which represents the *variance–bias tradeoff* in PCA, PCR, and PLS: the modeling error decreases with increasing number of factors, while the statistical uncertainty (estimated variance) increases with increasing model complexity. Note that discarding relevant factors leads to *underfitting*, while modeling too many phenomena leads to *overfitting* (fitting the noise).

The following procedure for the choice of the optimal number of factors A^* is adopted (Martens and Naes, 1989):

(1) Computation of the prediction error for different number of factors based on either internal or external test sets (internal or external validation).
(2) Selection of A^* with the lowest (N)RMS.

In practice, a shallow valley can occur in the (N)RMS curve as a function of the number of factors. Then, following the parsimony principle (Osten, 1988), the statistical model with the lowest number of factors should be chosen.

In the case of spectral data, when the pure-component spectra and/or the concentration profiles are expected to be smooth, then, alternatively to the choice of A^* based on the (N)RMS curve, A^* could be determined by visual inspection or the autocorrelation of

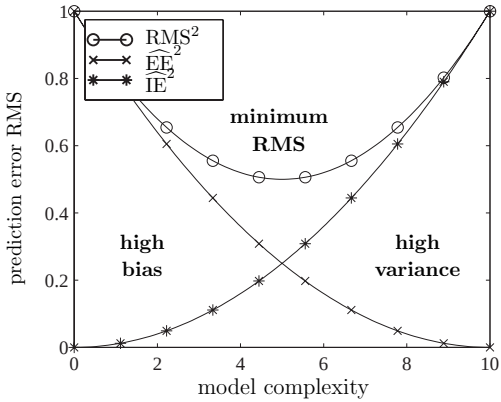

Figure C.1. Variance–bias tradeoff as a function of model complexity.

the loadings, weights, or scores of the statistical candidate models. This is equivalent to pseudo-rank estimation of $\tilde{\mathbf{Y}}$ (see Subsection C.4.1).

Example C.3 (Model selection based on internal or external validation)
Consider non-reacting mixtures involving $S = 4$ absorbing species. Let $\mathcal{X}_i$ denote the ith absorbing species ($i = 1, \ldots, S$). The pure-component spectra of the four species are given in Figure C.2 ($L = 120$).
Three issues are illustrated here for the prediction of the concentrations of $\mathcal{X}_3$ from spectra based on a calibration model: (i) bias-variance-tradeoff, (ii) internal vs. external validation for model structure selection, and (iii) model structure selection based on visual inspection of the loadings. For this purpose, a calibration set is designed using a three-level full-factorial plan with the lower and upper concentration limits for the four species being {[1, 2.5]; [0.2, 2]; [0.5, 1]; [0.8, 3]}. This results in $K_{\mathrm{c}} = 81$ mixtures. A test set with $K_{\mathrm{n}} = 40$ mixtures is generated randomly for concentrations between the given lower and upper limits. $\mathbf{A}$ of both calibration and test sets are corrupted by noise generated by multiplying a number from a uniform distribution in the range $[-0.05, 0.05]$ by the signal. Also, $\mathbf{C}_k$ is corrupted by noise generated from the Gaussian distribution with standard deviation being 10% of the maximum of the column-means of $\mathbf{C}_k$. Since PCR and SIMPLS give similar results, only those for SIMPLS are shown.
It is desirable that the calibration model with minimum RMS also has a negligible bias. This leads to (almost) unbiased prediction intervals (Faber and Kowalski, 1997). However, this goal is difficult to achieve in a situation described schematically by Figure C.1, where low bias can only be obtained by increasing the model complexity beyond the minimum prediction error $\mathrm{NRMS_n}$. However, here, the bias essentially goes to zero for four or more factors retained, while the estimated variance increases slowly (see Figure C.3a). Fortunately, this often happens for spectral data from chemical applications where the number of true factors is rather small compared to the number of observations and channels.
Internal validation is performed by leave-one-out-cross validation. Figure C.3b illustrates that the fit ($\mathrm{NRMS_c}$) improves steadily with increasing number of factors and, therefore, $\mathrm{NRMS_c}$ cannot be used for choosing the optimal number of factors A^*. Alternatively, both internal and external validation procedures determine an optimal

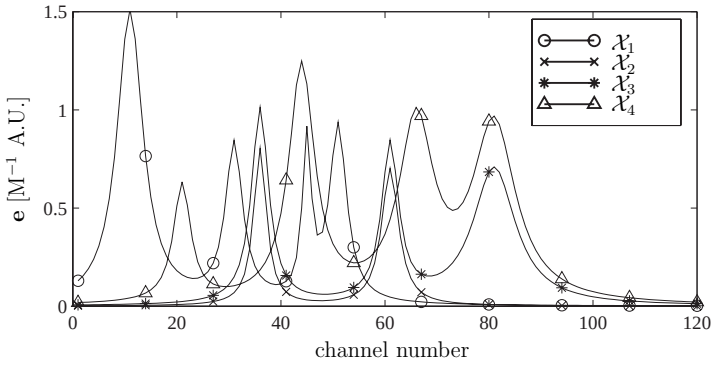

Figure C.2. Pure-component spectra $\mathbf{E}$ for Example C.3.

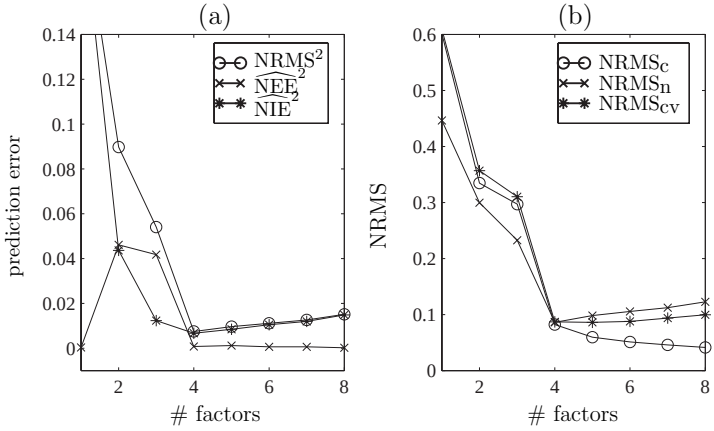

Figure C.3. (a) Relation between normalized prediction error (NRMS_n) and number of factors (model complexity), (b) fit (NRMS_c), external validation (NRMS_n), and internal validation (NRMS_cv) as a function of factors (model complexity).

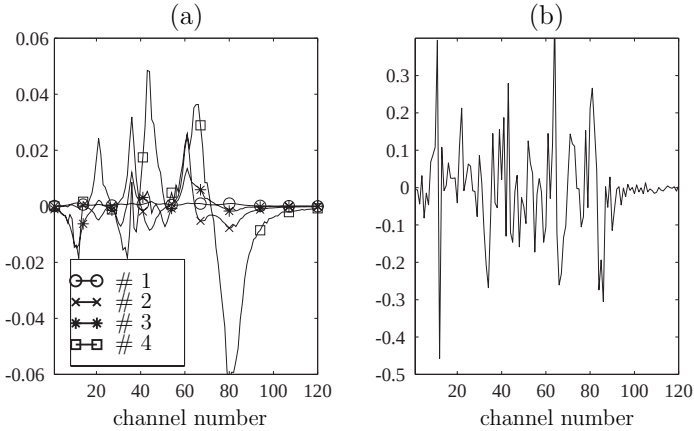

Figure C.4. SIMPLS weights $\hat{\mathbf{W}}$: (a) weight vectors #1–4 and (b) weight vector #5.

number of factors at $A^* = 4 = S$ as expected. This nicely illustrates the versatility of internal validation for model structure selection, which has the additional advantage over external validation that it does not require an external test set. The prediction error at $A^* = 4$ is NRMS = 8.7 %, which is in good agreement with the added Gaussian noise level of 10 %. Furthermore, $\mathrm{NRMS_n}$ is slightly larger than $\mathrm{NRMS_{cv}}$ owing to the fact that, contrary to the internal test sets used in CV, the external test set is not used for calibration.

Figure C.4 shows that the first four weight vectors $\hat{\mathbf{w}}_i$ $(i = 1, \ldots, 4)$ are smooth. This can be attributed to the pure-component spectra of the four species being smooth. However, the fifth one is already scattered due to noise. Consequently, alternatively to internal or external validation, the optimal number of $A^* = 4$ can also be deduced from visual inspection of the weight vectors.

D

Assumptions for the various representations

The following assumptions generally hold:

(A1) $r_i(\mathbf{c}) \geq 0 \; \forall i$, i.e., the reaction rates are nonnegative rational functions that do not depend explicitly on time.

(A2) $r_i(\mathbf{c}) = 0$ if $c_j = 0$ for some $j \in I^{ri}$. This condition is due to the physical fact that Reaction i only takes place if all its reactants are present or, equivalently, it does not take place if one of its reactants is absent.

(A3) The reaction system is modeled in terms of R independent reactions and p inlet streams with constant, linearly-independent inlet concentrations and volumetric flowrates. Thus, $\text{rank}\,(\mathbf{N}) = R$ and $\text{rank}\,(\mathbf{C}_{in}) = p$. Furthermore, mixing is ideal and the concentrations in the outlet(s) are equal to that inside the reactor.

For the basic model, it is further assumed that:

(A4) Reactor and inlet densities are constant and equal and, thus, $\rho = \rho_{in}^i$ $(i = 1, \ldots, p)$,

(A5) Reactor and inlet temperatures are constant and equal and, thus, $T = T_{in}^i$ $(i = 1, \ldots, p)$,

For the factorization of concentration data, the following assumption is usually made in addition to Assumptions A1–5:

(A6) $K \geq \max(S, R + p + 1)$,

(A7) Each of the R independent reactions is active on a time interval during the K observations.

For the factorization of reacting spectral data, the following assumptions are usually made in addition to Assumptions A1–7:

(A8) The factorization of general spectral data (5.2) holds.

(A9) $L \geq \max(S, R+p+1)$. Without loss of generality, assume channels to be ordered in increasing order.

(A10) $\mathbf{E}$ is independent of temperature, pressure, and pH.

(A11) $\text{rank}\,(\mathbf{E}) = S$, i.e., the pure-component spectra of the S absorbing species are linearly independent.

(A12) All species both react and absorb.

E

Alternate derivation of (4.3) and alternate data pre-treatment

E.1 Alternate derivation of (4.3)

In Subsection 4.1.1, the factorization of concentration data $\mathbf{c}(k)$ (see equation 4.3) was derived from the transformation to reaction-variant form. This led to the interpretation of $\mathbf{x}$, $\mathbf{z}$, and V_0 being the reaction-variant vector, reaction-invariant and flow-variant vector, and reaction and flow invariant, respectively. Alternatively, it is shown next that (4.3) can also be derived directly from (2.38) without the use of the transformation to reaction-variant form. However, the latter derivation lacks the interpretation of factorization of concentration data derived in Theorem 4.1.

The direct integration of (2.38) with respect to time from $t = 0$ to t_k leads to:

$$\mathbf{c} = \mathbf{N}^{\mathrm{T}}\,\mathbf{x}^{\star}/V + \mathbf{C}_{in}\,\mathbf{z}^{\star}/V - \mathbf{c}^{\star}_{out}/V + V_0\,\mathbf{c}_0/V \tag{E.1}$$

where the R-, p-, and S-dimensional vectors $\mathbf{x}^{\star}$, $\mathbf{z}^{\star}$, and $\mathbf{c}^{\star}_{out}$ are defined as $\dot{\mathbf{x}}^{\star} \equiv V\,\mathbf{r}(\mathbf{c})$, $\dot{\mathbf{z}}^{\star} \equiv \mathbf{q}_{in}$, and $\dot{\mathbf{c}}^{\star}_{out} \equiv q_{out}\,\mathbf{c}$ with initial conditions arbitrarily set to zero. The main problem with formulation (E.1) is that $\mathbf{c}$ depends on its integral $\mathbf{c}^{\star}_{out}$. To avoid this, an auxiliary state λ is introduced that is as defined in (4.7). Combining (2.38) and (4.7) gives:

$$\frac{d}{dt}\left(\frac{V\,\mathbf{c}}{\lambda}\right) = \mathbf{N}^{\mathrm{T}}\,\mathbf{r}(\mathbf{c})\,V/\lambda + \mathbf{C}_{in}\,\mathbf{q}_{in}/\lambda \tag{E.2}$$

Let the R- and p-dimensional vectors $\mathbf{x}$ and $\mathbf{z}$ be defined as in (4.7). From (4.7), the extent of the ith reaction reads:

$$x_i(t_k) = \int_0^{t_k} \frac{V(t)}{\lambda(t)}\,r_i(\mathbf{c}(t))\,dt. \tag{E.3}$$

Then, integrating (E.2) and (4.7) with respect to time from $t = 0$ to t_k:

$$\mathbf{c}\,V/\lambda - \mathbf{c}_0\;V_0 = \mathbf{N}^{\mathrm{T}}\,\mathbf{x} + \mathbf{C}_{in}\,\mathbf{z}. \tag{E.4}$$

Equation (4.3) is obtained after some algebraic manipulations of (E.4).

E.2 Alternate data pre-treatment

If $\mathbf{C}_{in}$, $\mathbf{c}_0$, $\mathbf{c}(k)$, $\mathbf{q}_{in}(k)$, $q_{out}(k)$, and $V(k)$ are either known or measured, then an alternative way of computing the data vector $\mathbf{d}^\star(k)$, similar to $\mathbf{d}(k)$, exists. By rearranging (E.1), $\mathbf{d}^\star(k)$ can be computed using (Harmon *et al.*, 1995):

$$\mathbf{d}^\star(k) \equiv V(k)\,\mathbf{c}(k) - V_0\,\mathbf{c}_0 - \mathbf{C}_{in}\,\mathbf{z}^\star(k) + \mathbf{c}^\star_{out}(k) = \mathbf{N}^{\mathrm{T}}\mathbf{x}^\star(k). \tag{E.5}$$

For K observations, $\mathbf{D}^\star = \mathbf{X}^\star\,\mathbf{N}$ is obtained with $\mathbf{X}^\star$ being a $K \times (R + p + 1)$ matrix.

Note that $\mathbf{c}(k)$ is used explicitly and implicitly for the computation of $\mathbf{c}^\star_{out}(k)$. It is important to compare the precision with which $\mathbf{d}(k)$ and $\mathbf{d}^\star(k)$ can be computed. Usually, the number of observations of $\mathbf{q}_{in}$ and q_{out} is large and, thus, guarantees a high precision of the states $\tilde{V}(k)$, $\mathbf{z}(k)$, and $\mathbf{z}^\star(k)$ by numerical integration. In contrast, the number of observations of $\mathbf{c}(k)$ is small, leading to imprecision in the computation of $\mathbf{c}^\star_{out}$. The larger the variation of q_{out} between subsequent observations of $\mathbf{c}$, the larger the error in $\mathbf{c}^\star_{out}$. Thus, compared to $\mathbf{d}^\star(k)$, $\mathbf{d}(k)$ has the advantages of higher precision and lower computational cost ($(p+1)$ vs. $(p+S)$ differential equations to integrate).

F

Spectral measurements and calibration

F.1 Spectral measurements

F.1.1 Instrument

The instrumental response a is often the result of the pre-linearized instrumental output O using a pre-linearizing transform $f_L(O)$ such as the natural logarithm of the inverse of the transmission in absorption spectroscopy (Burns and Ciurzak, 1992):

$$a = f_L(O). \tag{F.1}$$

Following the tensorial terminology introduced by Sánchez and Kowalski (1988*a*,*b*) into the field of calibration, a scalar instrumental response is a zeroth-order tensor, a vector instrumental response a first-order tensor, and an instrumental matrix response a second-order tensor.

Here, only responses from first-order instruments are considered. Let $a(k, \ell)$ and $\mathbf{a}(k)$ be the response at the ℓth channel and the complete L-dimensional response vector at L channels, respectively, for the kth observation instant:

$$a(k, \ell) = \sum_{s=1}^{S} \mathfrak{c}_s(k)\, e_s(\ell) \quad \text{or} \quad \mathbf{a}^{\mathrm{T}}(k) = \mathbf{c}^{\mathrm{T}}(k)\, \mathbf{E}, \tag{F.2}$$

where $\mathbf{c} = [\mathfrak{c}_1, \ldots, \mathfrak{c}_S]^{\mathrm{T}}$ is the S-dimensional concentration vector, $e_s(\ell)$ is the instrumental coefficient of the sth species at the ℓth channel, $\mathbf{E}$ the $S \times L$ matrix of instrumental coefficients (pure-component spectra or instrumental matrix). The response a is said to be *linear*, if it is a linear combination of the S (time-invariant) pure-component spectra weighted by the concentrations of the S species $\mathbf{c}$.

For K observations, (F.2) reads:

$$\mathbf{A} = \mathbf{C}\, \mathbf{E} \tag{F.3}$$

with $\mathbf{A}$ being the $K \times L$ response matrix and $\mathbf{C}$ the $K \times S$ concentration matrix. Equation (F.3) describes the *factorization of general (linear) spectral data.*

Three different types of concentration measures $\mathfrak{c}_j(k)$ for the jth species at observation k are considered here.

(1) The molar concentration $c_j(k)$ is defined as in (2.11).
(2) The weight fraction $w_j(k)$ is defined as in (2.41).
(3) The mole fraction $w_{n,j}(k)$ is defined as:

$$w_{n,j}(k) \equiv \frac{c_j(k)}{\mathbf{1}_S^{\mathrm{T}}\mathbf{c}(k)}, \quad \forall j = 1, \ldots, S, \qquad \mathbf{1}_S^{\mathrm{T}}\mathbf{w}_n(k) = 1, \tag{F.4}$$

where $\mathbf{c}$ is the S-dimensional molar concentration vector, and $\mathbf{w}_n$ the S-dimensional mole fraction vector. Note that $\mathbf{1}_S^{\mathrm{T}}\mathbf{c}(k)$ is equivalent to $\sum_{s=1}^{S} c_s(k)$ and the mole fractions are closed, since $\mathbf{1}_S^{\mathrm{T}}\mathbf{w}_n(k) = 1$. The unit is usually given in [n-%].

A channel number can correspond to: (i) a wavelength λ e.g. in [nm], (ii) a wavenumber or frequency ν e.g. in [cm^{-1}], (iii) the number $\mathfrak{p}$ of an integrated peak, (iv) the mass/electron charge (m/e) peak, (v) the cross-polarization contact time t_{cp}, or (vi) the retention time t_r. An observation instant k can be a sample number or more specifically a time instant t_k.

For various types of first-order instruments such as absorption (NIR, MIR, UV/VIS), emission, optical rotary dispersion (ORD), mass spectroscopy (MS), nuclear magnetic resonance (NMR), gas chromatography (GC), liquid chromatography (LC), table F.1 presents the corresponding response $a(k, \ell)$ as a function of the instrumental output O and the pre-linearizing transform, the concentration $\mathfrak{c}_s(k)$, the instrumental coefficient $e_s(\ell)$, and the channel type ℓ. For more details about the particular measurements it is referred to the literature listed in Table F.1. Note that excitation/emission fluorescence spectroscopy, GC/MS or LC/UV are second-order instruments. However, to use these types of measurements as data from first-order instruments, it is assumed that only one column or row of the measured matrix is analyzed.

F.1.2 Data pre-treatment

The types of data pre-treatment considered herein are: (i) channel selection, (ii) removing of outliers, (iii) filtering, (iv) mean centering, (v) differentiation, (vi) normalization, (vii) auto-scaling, and (viii) baseline elimination. The last four pre-treatment types ii–vii can be carried out either in the observation or in the channel direction.

Channel selection
Let $\mathbf{A}^\star$ denote the $K \times L^\star$ spectral matrix, where $L^\star$ is the total number of available channels. Then,

$$\mathbf{A} = \mathbf{A}^\star \boldsymbol{\mathcal{P}}_L, \tag{F.5}$$

where $\mathbf{A}$ is the $K \times L$ spectral matrix, $\boldsymbol{\mathcal{P}}_L$ an $L^\star \times L$ matrix, and L the number of selected channels.

Removing outliers
Observations/variables showing some types of departure from the bulk of the data are usually called *outliers* or abnormal observations/variables. In both calibration and factor analysis, outliers may be detrimental to the quality of the predictor or the

Table F.1. The measurement type, the response $a(k,\ell)$ as a function of the instrumental output O and the pre-linearizing transform, the concentration $\mathfrak{c}_s(k)$, the instrumental coefficient $e_s(\ell)$, and channel type ℓ for various types of first-order instruments: Absorption (NIR, MIR, UV/VIS), emission, optical rotary dispersion (ORD), mass spectroscopy (MS), nuclear magnetic resonance (NMR), gas chromatography (GC), liquid chromatography (LC). Notation: (i) wavelength λ, (ii) wavenumber or frequency ν, (iii) number $\mathfrak{p}$ of an integrated peak, (iv) cross-polarization contact time t_{cp}, (v) retention time t_r.

Type	$a(k,\ell)$	$\mathfrak{c}_s(k)$	$e_s(\ell)$	ℓ	References
Absorption	transmittance absorbance[a]	c	molar extinction/ pathlength	NIR, UV/VIS: λ MIR: ν, p	**NIR:** Burns and Ciurzak (1992); **MIR:** Pretsch *et al.* (1981); Griffiths and de Haseth (1986); **UV/VIS:** Pretsch *et al.* (1981)
	reflectance absorbance[b]	w or c	extinction/ pathlength	NIR: λ	Birth and Hecht (1987)
Emission	Raman intensity	c	molar intensity	ν, p	Colthup *et al.* (1990)
	fluorescence intensity	c	[c]	λ_j or λ_i[d]	Chang *et al.* (1993)
ORD	rotation angle	c	specific rotation	λ	McMullen *et al.* (1967)
MS	signal intensity	w_n	[e]	m/e peak	Malinowski and McCue (1977)
NMR	signal intensity	c	[f]	t_{cp}[d]	Kormos and Waugh (1983)
GC/MS	signal intensity	w_n	[e]	t_r[d]	Knorr *et al.* (1981)
LC/UV	absorbance	c	molar extinction/ pathlength	t_r[d]	Frans *et al.* (1985); Weiss (1991)

[a] According to Beer's law, the absorbance is calculated as: $\ln(1/O)$.

[b] According to Kubelka–Munk's transform (Kortum, 1969), the absorbance is calculated as: $\frac{(1-O)^2}{2O}$.

[c] Quantity proportional to the number of photons absorbed per unit concentration of Species s at excitation wavelength λ_j or fraction of fluorescence emitted by Species s at emission wavelength λ_i.

[d] Assumptions. Fluorescence emission: $\lambda_i = cst.$ or $\lambda_j = cst.$. NMR: cst. spinning speed. GC/MS: only one mass selected. LC/UV: only one wavelength selected.

[e] Height of pure Species s at normal pressure of Species s × mass discrimination factor × total pressure/normal pressure corresponding to Species s.

[f] Cross-polarization efficiency × line shape function for Species s.

interpretability, respectively. In prediction, it is important to have some type of error warning to avoid individual mistakes (Martens and Naes, 1989). Outliers may occur for many different reasons such as laboratory errors, observations from another population, and instrument errors. Thus, outlier detection and possible rejection (removing) is important.

A way to deal with outliers is to use robust statistical methods (Huber, 1981). Based on certain statistical assumptions about the data, most robust methods downweight observations with high influence (outliers). However, outliers may sometimes be highly informative and not at all erroneous. In calibration, for example, an observation may be an outlier only because it alone spans a certain type of important variability in the data. Therefore, in this dissertation the approach is adopted that only detects the outliers and let the user decide what to do with it, rather than implicitly treating them using robust methods. Various methods for outlier detection are available in the literature. For example, Martens and Naes (1989) (Chap. 5) present methods for outlier detection in regression problems, Steinhauser (1996) addresses the impact of influential observations and the masking problem of *multiple* outliers in factor analysis, and Walczak and Massart (1995) presents a robust method as a detection tool for outliers.

Filtering
For the elimination of high-frequency noise in spectral data with localized signal details such as sharp bands in MIR spectral data, denoising using wavelets have shown to be a powerful filtering technique (Vetterli and Kovačević, 1995). Denoising can be applied in channel or observation direction.

Mean centering
The column-mean-centered matrix of $\mathbf{A}$ denoted by $\bar{\mathbf{A}}_c$, is given by

$$\bar{\mathbf{A}}_c = \mathbf{A} - \mathbf{1}_K\,\bar{\mathbf{a}}_c^{\mathrm{T}} = \mathbf{J}_c\,\mathbf{A} \quad \text{with} \quad \bar{\mathbf{a}}_c^{\mathrm{T}} = \frac{1}{K}\,\mathbf{1}_K^{\mathrm{T}}\,\mathbf{A} \tag{F.6}$$

the column mean. From (F.3), column-mean centering leads to

$$\bar{\mathbf{A}}_c = \bar{\mathbf{C}}_c\,\mathbf{E}, \tag{F.7}$$

Row-mean centering leads to $\bar{\mathbf{A}}_r = \mathbf{C}\,\bar{\mathbf{E}}_r$, where $\mathbf{A}_r$ and $\mathbf{E}_r$ are row-mean-centered matrices.

Differentiation
Let

$$\mathsf{D}_c \equiv \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix} \tag{F.8}$$

be the $K \times K$ backward finite difference matrix, and $\dot{\mathbf{A}} = \mathsf{D}_c\,\mathbf{A}$ denote the spectral matrix obtained by taking finite differences along the columns of $\mathbf{A}$. Since $\mathbf{E}$ is assumed

to be time-invariant, it follows from (F.3):

$$\dot{\mathbf{A}} = \dot{\mathbf{C}}\,\mathbf{E}, \tag{F.9}$$

When differentiated with respect to channel number,

$$\frac{d\mathbf{A}}{d\lambda} = \mathbf{C}\,\frac{d\mathbf{E}}{d\lambda} \tag{F.10}$$

since $\mathbf{C}$ is independent of the channel number.

Normalization and auto-scaling
Normalization is performed on either the columns (rows) of $\mathbf{A}$ such that each column (row) has a Euclidian norm of 1. Auto-scaling of $\mathbf{A}$ in column direction is a data pre-treatment composed of two sequential steps: (1) column-mean centering and (2) division of each column by its standard deviation.

Baseline elimination
One specificity of spectral data compared to concentration data is that an instrumental shift can cause baseline problems. A nonzero reference spectrum is always present (e.g. spectrum of air or of the solvent). This causes the rank of $\mathbf{A}$ to be larger than rank $(\mathbf{C})$, since rank analysis considers the air/solvent as an extra absorbing species. Here, the removal of the additional nonzero singular value(s) caused by the baseline and its time variation is considered.

Time-invariant baseline:
Usually, the reference spectrum is subtracted from the spectral matrix before any analysis is performed. However, if the reference spectrum is not subtracted, then it can be considered as the spectrum of a substance whose concentration remains unchanged. Mathematically stated, one column of the $K \times (S+1)$ concentration matrix $\mathbf{C}_t$ remains constant with

$$\begin{aligned} \mathbf{C}_t &\equiv \left[\begin{array}{c|c} \mathbf{c}_{base} & \mathbf{C} \end{array}\right] = \begin{bmatrix} \mathbf{1}_K & \mathbf{C} \end{bmatrix} \left[\begin{array}{c|c} c_{base} & \mathbf{0}_S^{\mathrm{T}} \\ \hline \mathbf{0}_S & \mathbf{I}_S \end{array}\right], \\ \mathbf{A} &= \mathbf{C}_t\,\mathbf{E}_t, \quad \mathbf{E}_t = \left[\begin{array}{c} \mathbf{e}_{base}^{\mathrm{T}} \\ \hline \mathbf{E} \end{array}\right], \end{aligned} \tag{F.11}$$

where it was assumed, without loss of generality, that the first column of $\mathbf{C}$ remains constant, and c_{base} is a constant pseudo-concentration for the baseline. As can be seen from (F.11), a pure-component spectrum $\mathbf{e}_{base}$ can be attributed to the baseline.

By (i) subtracting a reference spectrum, (ii) column-mean centering, or (iii) differentiation with respect to time, any time-invariant contribution of the measurement device to the spectrum can be eliminated since $\mathbf{1}_K \in \mathcal{S}_c\,(\mathbf{C}_t)$ and Property B.6. However, time-dependent changes in the baseline will not be corrected this way.

Baseline with time-varying abscissa:
If $\mathbf{A}$ can be decomposed into

$$\mathbf{A} = \mathbf{a}_{base}\,\mathbf{1}_L^{\mathrm{T}} + \mathbf{A}_1, \tag{F.12}$$

then differentiating with respect to the channel number leads to

$$\frac{d\mathbf{A}}{d\lambda} = \frac{d\mathbf{A}_1}{d\lambda}, \tag{F.13}$$

where $\mathbf{a}_{base}$ is the abscissa of the baseline $\mathbf{a}_{base}$ signifying the time variation of a baseline parallel to the channel axis.

Analogous to a time-invariant contribution of the measurement device, any channel-independent contribution $\mathbf{a}_{base}$ can be eliminated by differentiating once with respect to channel number, while differentiation with respect to time will not help.

Among time-varying baselines, the case where the baseline is parallel to the channel axis, has been discussed above. Alternatively, if the baseline varies linearly with the channel number, the second derivative with respect to channel number has to be used to eliminate the effect.

Baseline with time-varying abscissa and slope:
If the baseline varies linearly with the channel number, then

$$\mathbf{A} = \mathbf{a}_{base}\,\mathbf{1}_L^{\mathrm{T}} + \mathbf{m}\,\boldsymbol{\lambda}^{\mathrm{T}} + \mathbf{A}_1, \tag{F.14}$$

where $\mathbf{m}$ is the baseline slope, $\boldsymbol{\lambda}$ the channel vector, and $\mathbf{a}_{base}$ and $\mathbf{m}$ contain the variation of the abscissa and slope of the baseline with respect to time, respectively. When the spectral matrix is differentiated twice along the channel direction, then

$$\frac{d^2\mathbf{A}}{d\lambda^2} = \frac{d^2\mathbf{A}_1}{d\lambda^2}. \tag{F.15}$$

F.2 Advantages of multivariate forward calibration

Especially in the case when the spectrum of the analyte is overlapped with other species (e.g. in NIR spectra), multivariate calibration methods have been shown to improve analysis precision, accuracy, reliability, and applicability of spectral analyses relative to the more conventional univariate methods of data analysis (Haaland, 1992). Rather than attempting to find and use only an isolated spectral feature, multivariate methods derive their power from the simultaneous use of multiple regions in each spectral measurement. Thus, the problem of spectral interferences can be eliminated with the use of any one of the numerous multivariate calibration methods. For practical purposes, it is important to quantify the concentrations of only the species for which calibration concentrations are available (*known* species or species of interest) in a complex mixture irrespective of the presence of other species (*unknown* species or interferents).

The difference between the various multivariate calibration methods are a result of the model relating spectral intensities and concentrations and the assumptions made in the models about the errors. The present distinction of the various calibration methods concerns predictive vs. causal modeling (see Martens and Naes, 1989, for other ways of classification). Two calibration model types are distinguished (Krutschkoff, 1967): the *forward* (or *inverse*) and the *reverse* (or *classical*) calibration model. For spectral data, the forward (predictive model) and the reverse (causal model) are given by

$$\tilde{\mathbf{C}} \;=\; \mathbf{A}\,\mathbf{B} + \tilde{\mathbf{G}}, \tag{F.16}$$

$$\tilde{\mathbf{A}} \;=\; \mathbf{C}\,\mathbf{E} + \tilde{\mathbf{F}}, \tag{F.17}$$

respectively, where $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{F}}$ are matrices of random errors or model inadequacies. From (5.11) and (F.17), it can be seen that forward and reverse calibration means regressing $\tilde{\mathbf{C}}$ on $\mathbf{A}$ and $\tilde{\mathbf{A}}$ on $\mathbf{C}$, respectively. The forward and reverse calibration methods give predictors with different statistical properties since the error structures are different. However, the difference is small if the noise levels in $\tilde{\mathbf{C}}$ and/or $\tilde{\mathbf{A}}$ are low (Martens and Naes, 1989).

The reverse formulation requires the columns of $\mathbf{C}$ to be linearly independent (or, equivalently, to be of full column rank; Krutschkoff, 1967). This assumption is a major disadvantage, since all interfering sources (chemical or physical) need to be known and included in the calibration. Furthermore, all the species need to be linearly independently varied; an assumption that can usually be verified for non-reacting mixtures by the use of experimental planning methods (Box *et al.*, 1978).

The *original* forward formulation requires full column rank of $\mathbf{A}$. For the verification of this assumption, channel selection methods must be applied (Gemperline, 1989; Brown, 1992, 1993; Lucasius *et al.*, 1994). This represents a major disadvantage since suboptimal channel selection can introduce problems such as poor baseline modeling and noise inflation from the collinearity problem (Haaland and Thomas, 1988). However, this assumption can be relaxed using the framework of *bilinear models* together with e.g. PCR or PLS (see Appendix C.2; Martens and Naes, 1989). Owing to the disadvantage of the reverse formulation, only the (multivariate) forward formulation is studied.

Remark F.1 (Necessary conditions for $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$)
For the specification of *necessary* conditions for $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$, note that $\mathbf{a}_n \notin \mathcal{S}_r(\mathbf{A})$ and $\operatorname{rank}(\mathbf{E}) = S$ imply the corresponding (complete) concentration vector $\mathbf{c}_n \notin \mathcal{S}_r(\mathbf{C})$. With $\mathbf{c}_n^{\mathrm{T}} = \mathbf{t}_n^{\mathrm{T}}\mathbf{C}$ and $\mathbf{c}_{k,n}^{\mathrm{T}} = \mathbf{C}\,\mathbf{J}_k$,

$$\mathbf{c}_n^{\mathrm{T}} = \mathbf{t}_n^{\mathrm{T}}\mathbf{C} + \mathbf{c}^{\perp\mathrm{T}}, \quad \mathbf{c}_{k,n}^{\mathrm{T}} = \mathbf{t}_n^{\mathrm{T}}\mathbf{C}\,\mathbf{J}_k + \mathbf{c}^{\perp\mathrm{T}}\mathbf{J}_k, \tag{F.18}$$

where $\mathbf{c}^{\perp}$ is the orthogonal complement to $\mathcal{S}_r(\mathbf{C})$, i.e., $\mathbf{c}^{\perp} \notin \mathcal{S}_r(\mathbf{C})$.

It can easily be verified that $\mathbf{Q}^{+\mathrm{T}}\mathbf{Q}^{\mathrm{T}}$ and $\mathbf{E}\,\mathbf{P}^{+\mathrm{T}}\mathbf{Q}$ span the same row space and $\mathbf{c}^{\perp\mathrm{T}}\mathbf{E}\,\mathbf{P}^{+\mathrm{T}}\mathbf{Q} = \mathbf{0}_S^{\mathrm{T}}$. Then,

$$\hat{\mathbf{c}}_{k,n}^{\mathrm{T}} = \mathbf{a}_n^{\mathrm{T}}\hat{\mathbf{B}}_k = \mathbf{c}_n^{\mathrm{T}}\mathbf{E}\,\hat{\mathbf{B}}_k = (\mathbf{t}_n^{\mathrm{T}}\mathbf{C} + \mathbf{c}^{\perp\mathrm{T}})\,\mathbf{E}\,\mathbf{P}^{+\mathrm{T}}\mathbf{Q}\mathbf{J}_k = \mathbf{t}_n^{\mathrm{T}}\mathbf{C}\,\mathbf{J}_k. \tag{F.19}$$

Comparing (F.18) with (F.19) shows that $\hat{\mathbf{c}}_{k,n}$ is offset with respect to $\mathbf{c}_{k,n}$ by $\mathbf{c}^{\perp\mathrm{T}}\,\mathbf{J}_k$, i.e., the component that lies outside $\mathcal{S}_r(\mathbf{C})$ cannot be predicted. However, if S_k, $\mathbf{Q}$ and $\mathbf{c}^{\perp\mathrm{T}}$ are chosen so that $\mathbf{c}^{\perp\mathrm{T}}\,\mathbf{J}_k = \mathbf{0}_{S_k}$ (necessary conditions), then the prediction is correct.

G

Local rank

Local rank analysis is used for the detection of local rank changes (decreasing or increasing) in time or channel direction. It will be shown that local rank analysis in time direction can provide valuable information about (i) chemical reactions such as start or termination of one or more reactions, or/and (ii) a change in operation such as start and stop of feeding. Alternatively, it will be shown that only under restrictive assumptions, local rank analysis in channel direction of spectral data from reacting mixtures determines regions where one or several species do or do not absorb.

It is assumed that: (i) R is the total number of independent reactions that occur during the complete time considered, (ii) each reacting species absorbs in at least one of L channels, and (iii) the complete pure-component spectra of all the S absorbing species are linearly independent, i.e., $\text{rank}\,(\mathbf{E}) = S$.

The following definitions are useful in this section.

Definition G.1 (Local rank change) *Let an $m \times n$ noise-free matrix $\mathbf{Y}$ be partitioned into*

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_- \\ \mathbf{Y}_+ \end{bmatrix}, \tag{G.1}$$

where $\mathbf{Y}_-$ and $\mathbf{Y}_+$ are matrices of dimension $m_- \times n$ and $m_+ \times n$, respectively, with $m = m_- + m_+$ and $m_+, m_- \geq n$. Let

$$d \equiv \text{rank}\,(\mathbf{Y}_+) - \text{rank}\,(\mathbf{Y}_-)\,. \tag{G.2}$$

If $d > 0$, it is said that $\mathbf{Y}$ exhibits a local rank increase of d, *and if $d < 0$, it is said that $\mathbf{Y}$ exhibits a* local rank reduction of $(-d)$.

The pseudo-rank of a matrix (denote as $\text{prank}\,(\cdot)$) was introduced in Appendix C.4.1.2 as the rank of a matrix in the absence of measurement error. Here, the concept of approximate-rank is introduced for *noise-free* matrices. If a signal becomes arbitrarily small, its contribution to the the variability of the matrix may become negligible. Thus, since most of the pseudo-rank estimation methods rely on some variance criterion of the residuals, it is proposed to use them also for the determination of the approximate-rank.

Definition G.2 (Local approximate-rank reduction) *Let an $m \times n$ noise-free matrix $\mathbf{Y}$ be partitioned as in* (G.1)*. $\mathbf{Y}$ is said to have a* local approximate-rank reduction of d *with*

$$d \equiv \operatorname{prank}(\mathbf{Y}_+) - \operatorname{prank}(\mathbf{Y}_-), \tag{G.3}$$

where $\operatorname{prank}(\cdot)$ *is redefined as the approximate-rank of a matrix.*

Definition G.3 (Linear dependence, approximate-rank, and rank deficiency) *If* $\operatorname{rank}(\mathbf{Y}) < n$*, then the columns of an $m \times n$ matrix $\mathbf{Y}$ are linearly dependent. If* $\operatorname{prank}(\mathbf{Y}) < n$*, then the columns of an $m \times n$ matrix $\mathbf{Y}$ are linearly dependent in the approximate-rank sense. For both cases, it is said that $\mathbf{Y}$ is rank deficient.*

First, local rank changes in time direction due to the underlying chemical reaction system and, then local rank changes in channel direction due to selectivity are studied.

G.1 Local rank in time direction

Rank changes in the concentration matrices $\mathbf{C}$ and $\mathbf{D}$ mainly occur in time direction due to operating or kinetic changes. A rank change can occur in the sense of both mathematical rank and approximate-rank. It will be shown that a mathematical rank change is due to a discrete event in the generalized inlets (e.g. additional batch runs, start/stop of feeding, addition of a mixture), i.e., the rank changes abruptly in $\mathbf{Z}$. Such rank changes were discussed in Subsections 5.2.2.1 and 5.2.2.2. In this section, only local rank changes in $\mathbf{X}$ are considered.

It will be shown that the discrete event in the generalized inlets can also cause a discrete (mathematical) rank increase in $\mathbf{X}$ (e.g. start of a reaction). Alternatively, it will be shown that a rank change in the approximate-rank sense can be due to the reaction kinetics, i.e., the approximate-rank changes locally in $\mathbf{X}$, which is not directly linked to a discrete event in the generalized inlets.

Practical situations for local mathematical and approximate-rank changes in $\mathbf{X}$ are studied in the next two subsections. Since only local rank changes in $\mathbf{X}$ are considered, it suffices to analyze the local rank of the RV-spectral matrix $\mathbf{H}$ and the corresponding RV-concentration matrix $\mathbf{D}$.

G.1.1 Local mathematical rank change in X

The column dimension of $\mathbf{X}$ is determined by the total number of independent reactions R that take place during K observations. An example for local mathematical rank change of $\mathbf{X}$ is provided next.

For data from reaction systems with different stoichiometries, consider $\mathbf{H}$ to be composed of 2 runs exhibiting different linearly-independent stoichiometries with R_j independent reactions ($j = 1, 2$). Locally, the rank is R_1 in the observation interval $[1, K_1]$, and R_2 in $]K_1, K_1 + K_2]$. A local mathematical rank change of $d = |R_2 - R_1|$ occurs at the $(K_1 + 1)$st time instant. This can easily be generalized to the case of B runs

exhibiting different linearly-independent stoichiometries. Note that the stoichiometries of the different runs can be mutually linearly dependent of each other. A special case is the activation of additional reactions.

Example 2.4e (Activation of an additional reaction; cont'd from page 12**)**
Consider the reversible parallel reaction system (2.9). Assume power-law kinetics, i.e., $r_1 = \kappa_1 c_1\, c_2 - \kappa_2 c_3$ and $r_2 = \kappa_3 c_1^2$. It is assumed that all four species absorb ($S = 4$). One batch run is conducted with $\mathbf{c}_0 = [3,\ 0,\ 0,\ 0]^{\mathrm{T}}$ M. At $t = 0.33$ h, a mixture of volume $V_\Delta = 0.2$ l is added with $\mathbf{c}_\Delta = [0,\ 10,\ 0,\ 0]^{\mathrm{T}}$ M. Similarly to Example 2.4c on page 28, $\mathbf{c}_0$ was chosen such that only the second reaction is active in the beginning, whereas the inlet concentrations $\mathbf{c}_{in}$ activates the second reaction (see Appendix A.4 for the parameters). Figure G.1 presents the concentrations and RV-concentrations for $K = 60$ observations. Figure G.2 depicts the pure-component spectra of all absorbing species ($L = 80$), which are used to construct $\mathbf{A}$. Gaussian noise with standard deviation of 0.0066 is added to $\mathbf{A}$. Figure G.1 shows that before the addition, the concentrations and RV-concentrations of Species $\mathcal{X}_2$ and $\mathcal{X}_3$ are zero, since the first reaction is not activated. Since noise is added, mathematical rank must be replaced by pseudo-rank. The pseudo-rank of $\check{\mathbf{H}}$ is $2 = R$ as expected. Figure G.3 illustrates the reaction variants and EFA forward in time applied to submatrices of $\check{\mathbf{H}}$ with pseudo-rank estimation based on visual inspection of the singular values. Figure G.3b shows that the second singular value becomes significant after the addition and, therefore, it correctly predicts the start of a new reaction induced by the addition (compare also with Figure G.3a). The identity of the reaction, however, cannot be determined from EFA, and must be deduced from physical insight. Thus, the global stoichiometric matrix is given by $\mathbf{N} = \begin{bmatrix} \mathbf{N}_1 \\ \mathbf{N}_2 \end{bmatrix}$ with $\mathbf{N}_1 = [-2,\ 0,\ 0,\ 1]$ and $\mathbf{N}_2 = [-1,\ -1,\ 1,\ 0]$.

Note that with this design of the experimental concentration matrix, i.e., $\mathbf{C}_x = \begin{bmatrix} \mathbf{c}_\Delta^1 \\ \mathbf{c}_0^{\mathrm{T}} \end{bmatrix}$, it was possible to decouple the first reaction from the second one. Such an approach will be useful in the resolution of the concentrations from spectral data using factor-analytical methods (see Subsection 5.3.2.4).

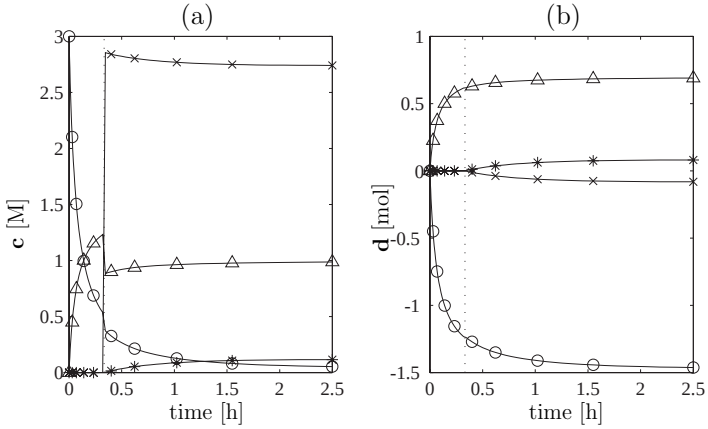

Figure G.1. (a) The concentrations $\mathbf{c}$, and (b) the RV-concentrations $\mathbf{d}$ of Example 2.4e. See Figure 4.12 on page 86 for legend. Addition at $t = 0.33$ h (dotted vertical line).

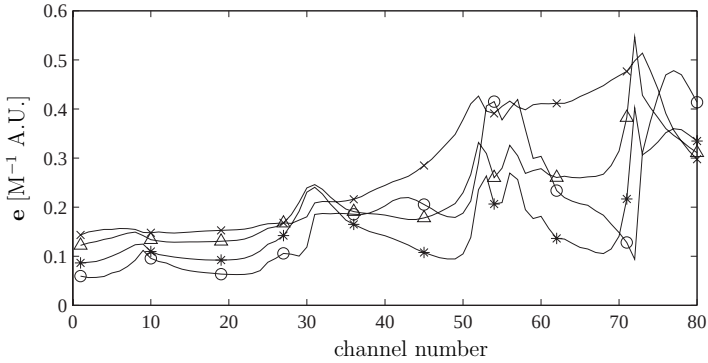

Figure G.2. Pure-component spectra **E** for Example 2.4e. For legend, see Figure 4.12 on page 86.

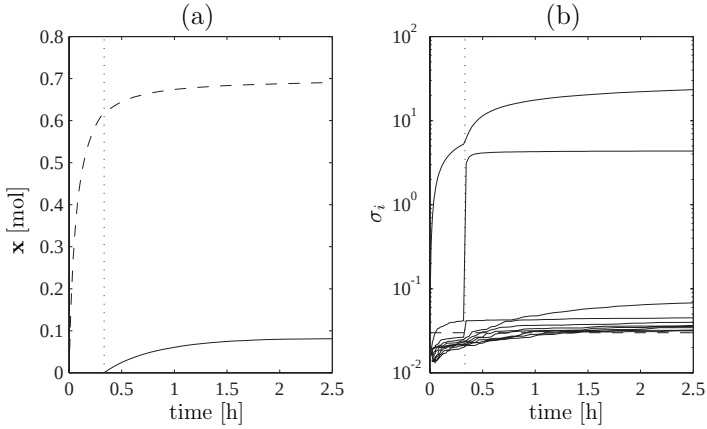

Figure G.3. Example 2.4e: (a) The reaction variants **x**, and (b) EFA forward in time applied to submatrices of $\tilde{\mathbf{H}}$ with pseudo-rank estimation based on visual inspection of the singular values. Legend for (a): (–) first reaction, (– –) second reaction. Addition at $t = 0.33$ h (dotted vertical line) and noise level at $\varepsilon_A = 0.03$ (dashed horizontal line).

G.1.2 Local approximate-rank change in X

Local approximate-rank change can occur due to the reaction rate, r_b, becoming linearly dependent/independent on the remaining $(R-1)$ reaction rates for $t \to \infty$. In practice, owing to noise, it is assumed that the reaction rate $r_b(t)$ is linearly dependent on the remaining ones in the sense of Definition G.3 for $t \geq \acute{t}$:

$$r_b(t) \approx \sum_{j \neq b} \alpha_j r_j(t), \qquad \forall t \geq \acute{t}, \tag{G.4}$$

where α_j is a coefficient, and $\acute{t}$ the time instant for which $r_b(t)$ is linearly dependent.

Let $\operatorname{rank}(\mathbf{E}) = S$ and $\mathbf{N}$ be constant, and the reaction variant $x_l(k)$ $(l = 1, \ldots, R)$ be decomposed as:

$$x_l(k) = \acute{x}_l + \Delta x_l(k), \quad \acute{x}_l \equiv \int_{t=0}^{\acute{t}_l} V(t)\, r_l(t)/\lambda(t)\, dt, \quad \forall k \in [\acute{k},\, K], \tag{G.5}$$

where $\acute{x}_l$ is the reaction rate integrated up to time instant $\acute{t}$ (with the discrete-time equivalent $\acute{k}$), and $\Delta x_l(k)$ the reaction variant change relative to $\acute{x}_l$ with $\Delta x_l(k) \equiv x_l(k) - \acute{x}_l$. Thus, (G.4) in terms of the reaction variants becomes with (G.5):

$$\Delta x_b(k) \approx \sum_{j \neq b} \alpha_j \Delta x_j(k), \quad \forall k \in [\acute{k},\, K]. \tag{G.6}$$

With $\Delta x_l(k) = (x_l(k) - \acute{x}_l)$, (G.6) becomes:

$$x_b(k) \approx \sum_{j \neq b} \alpha_j x_j(t) + \beta, \quad \beta = -\sum_{j=1}^{R} \alpha_j \acute{x}_j,\ \alpha_b \equiv -1, \quad \forall k \in [\acute{k},\, K]. \tag{G.7}$$

(G.7) shows that although the reaction rate r_b is linearly dependent on the remaining ones, the reaction variant x_b is usually independent of the remaining ones due to the (unknown) constant β.

Let $\mathbf{X}$ and $\mathbf{H}$ be partitioned as:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_- \\ \mathbf{X}_+ \end{bmatrix}, \qquad \mathbf{H} = \begin{bmatrix} \mathbf{H}_- \\ \mathbf{H}_+ \end{bmatrix}, \tag{G.8}$$

with $\mathbf{X}_-$ and $\mathbf{X}_+$ being matrices of dimension $k \times R$ and $(K-k) \times R$, respectively, and $\mathbf{H}_-$ and $\mathbf{H}_+$ matrices of dimension $k \times L$ and $(K-k) \times L$, respectively.

Without loss of generality, assume that r_b is the reaction rate of the first reaction. Then, for the last $(K - \acute{k})$ observations, the $(K - \acute{k}) \times R$ matrix of reaction variants, $\acute{\mathbf{X}}$, (approximately) becomes with (G.7):

$$\acute{\mathbf{X}} = \left[\, \mathbf{X}_i \;\middle|\; \mathbf{1}_{K-\acute{k}} \,\right] \left[\begin{array}{cc} \boldsymbol{\alpha} & \mathbf{I}_{R-1} \\ \hline \beta & \mathbf{0}_{R_1}^{\mathrm{T}} \end{array}\right] = \mathbf{X}_{i,e}\, \mathbf{R}, \quad \boldsymbol{\alpha} \equiv \begin{bmatrix} \alpha_2 & \alpha_3 & \cdots & \alpha_R \end{bmatrix}, \tag{G.9}$$

where $\mathbf{X}_i$ is an $(K - \acute{k}) \times (R - 1)$ matrix containing the linearly-independent reaction variants r_j $(j \neq b)$ on each row, $\mathbf{X}_{i,e}$ and $\mathbf{R}$ matrices of dimension $(K - \acute{k}) \times R$ and $R \times R$, respectively, and $\boldsymbol{\alpha}$ an $(R - 1)$-dimensional vector. Thus, $\acute{\mathbf{H}}$ becomes with (G.9):

$$\acute{\mathbf{H}} = \mathbf{X}_{i,e}\,\mathbf{R}\,\mathbf{N}\,\mathbf{E}. \tag{G.10}$$

From (G.9) and (G.10), it can be seen that $\mathbf{1}_{K-\acute{k}}$ lies (approximately) in $\mathcal{S}_c\left(\acute{\mathbf{H}}\right)$. For the elimination of $\mathbf{1}_{K-\acute{k}}$, either of the following three data pre-treatment methods of $\acute{\mathbf{H}}$ can be applied according to Properties B.6–P-Rank-5b: (D1) column-mean centering, (D2) differentiation with respect to time, and (D3) subtraction of the first row of $\acute{\mathbf{H}}$. Thus, $\acute{k}$ (i.e., the start of that region where r_b becomes linearly dependent/independent) can be determined from backward/forward EFA (see Appendix C.4.1.3) based on the approximate-rank of the pre-treated $\mathbf{H}_+$.

Two practical situations are depicted where one or several reaction rates become linearly dependent on the remaining ones in the sense of Definition G.3:

(S1) *Presence of limiting species in a set of irreversible reactions:* Assume a set of R irreversible reactions. Then, in the absence of feeding after a certain time instant, if the consumption of a reactant $\mathcal{X}_j$ is greater than its production rate, the concentration and the R_b consumption rates $\mathbf{r}_b(c_1, \ldots, c_j, \ldots, c_S)$ containing $\mathcal{X}_j$ as a reactant become asymptotically zero. This is a special case of linear dependence on the remaining $(R - R_b)$ reaction rates with the coefficients α_j being zero.

(S2) *Reversible reactions in equilibrium:* Assume that a reversible reaction asymptotically reaches equilibrium for $k \geq \acute{k}$. Then, the corresponding reaction rate r_b becomes asymptotically zero. Similarly to Situation S1, this is a special case of linear dependence on the remaining $(R - 1)$ reaction rates with the coefficients α_j being zero.

One practical situation where one or several reaction rates become linearly independent of the remaining ones in the sense of Definition G.3 is depicted. Assume that a reaction is auto-catalyzed and its reaction rate r_b is negligibly small for $k < \acute{k}$. Similarly to Situations S1 and S2, this is a special case of linear dependence on the remaining $(R - 1)$ reaction rates with the coefficients α_j being zero. For $k \geq \acute{k}$, however, r_b becomes significant or, equivalently, r_b becomes linearly independent of the remaining ones.

Four examples are presented next.

Example 2.4f (Limiting concentration of species $\mathcal{X}_2$ in irreversible reactions; cont'd from page 12**)**
One batch run is conducted with $\mathbf{c}_0 = [3, 0.8, 0, 0]^{\mathrm{T}}$ M and $V_0 = 0.8$ l. $\mathbf{c}_0$ was chosen such that the concentration of $\mathcal{X}_1$ is in excess, and the concentration of $\mathcal{X}_2$ is the limiting concentration for the first reaction (see Appendix A.4 for the parameters). Figure G.4 presents the concentrations and RV-concentrations for $K = 200$ observations. The pure-component spectra are the same as in Figure G.2. Uniform noise in the range [-0.0046, 0.0046] is added to $\mathbf{A}$.

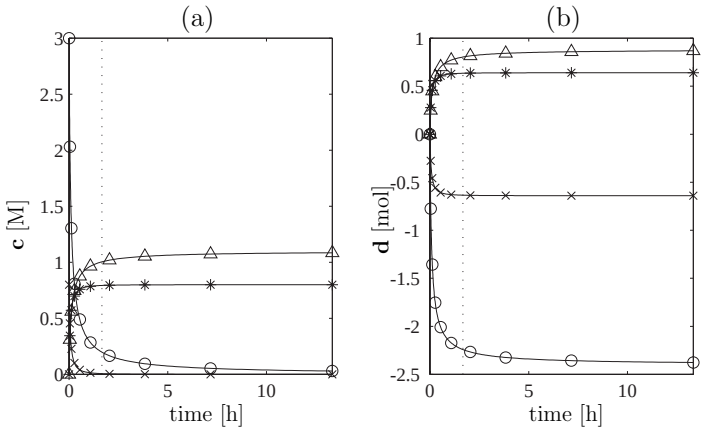

Figure G.4. (a) The concentrations **c**, and (b) the RV-concentrations **d** of Example 2.4f. See Figure 4.12 on page 86 for legend. Termination of the first reaction at $\acute{t} = 1.7$ h (dotted vertical line) according to Figure G.5b.

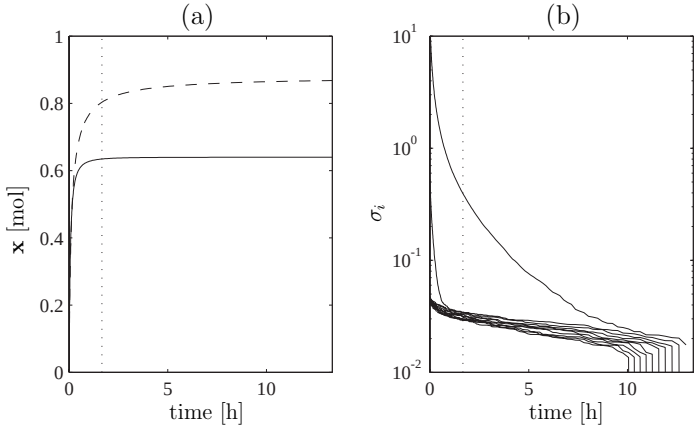

Figure G.5. Example 2.4f: (a) The reaction variants **x**, and (b) EFA backward in time applied to column-mean-centered submatrices of $\tilde{\mathbf{H}}$ with approximate-rank estimation based on visual inspection of the singular values. Legend for (a): (–) first reaction, (– –) second reaction. Termination of the first reaction at $\acute{t} = 1.7$ h (dotted vertical line) under the assumption of noise level $\varepsilon_A = 0.02$.

Figure G.4 shows that the concentration of Species $\mathcal{X}_2$ rapidly approaches 0 (limiting concentration) and the concentration of Species $\mathcal{X}_3$ reaches its final value.
The pseudo-rank of $\tilde{\mathbf{H}}$ is $2 = R$ as expected. Figure G.5 illustrates the reaction variants and EFA backward in time applied to column-mean-centered submatrices of $\tilde{\mathbf{H}}$ in time direction with approximate-rank estimation based on visual inspection of the singular values. As can be seen from Figure G.5b, the second largest singular value decreases rapidly and disappears in the noise at level $\varepsilon_A = 0.02$. Thus, a local approximate-rank reduction of 1 occurs at time instant $\acute{t} = 1.7$ h. Under the assumption of irreversible reactions and the presence of limiting species concentrations, this can be interpreted as the termination of a reaction. Another rank drop can be observed at 11.7 h indicating the termination of the second reaction. In summary, it can be concluded under the assumption of irreversible reactions, the presence of limiting species concentrations, and the knowledge of $\mathbf{N}$, that the first reaction ($\mathcal{X}_1 + \mathcal{X}_2 \rightarrow \mathcal{X}_3$) terminates before the second one ($2\,\mathcal{X}_1 \rightarrow \mathcal{X}_4$). This is confirmed in Figure G.5a, where the reaction extent of the first reaction stays approximately constant after $\acute{t} = 1.7$ h, and that of the second reaction after 11.7 h.

Example 2.2g (Reversible reaction in equilibrium; cont'd from page 18**)**
One batch run is performed with $\mathbf{c}_0 = [5, 5, 0, 0]^{\mathrm{T}}$ M. The initial concentration of the sulfuric acid is 0.23 M (see Appendix A.1 for the numerical values of the parameters). The concentration and RV-concentration profiles are presented in Figure G.6. Figure 5.9 depicts the pure-component spectra of all absorbing species ($L = 101$), which are used to construct $\mathbf{A}$. Gaussian noise with standard deviation of 0.038 is added to $\mathbf{A}$. Figure G.6 shows that the concentrations and RV-concentrations of all the species approximately reach equilibrium after about 1 h.
The pseudo-rank of $\tilde{\mathbf{H}}$ is $1 = R$ as expected. Figure G.7 illustrates the reaction variant and EFA backward in time applied to column-mean-centered submatrices of $\tilde{\mathbf{H}}$ in time direction with approximate-rank estimation based on visual inspection of the singular

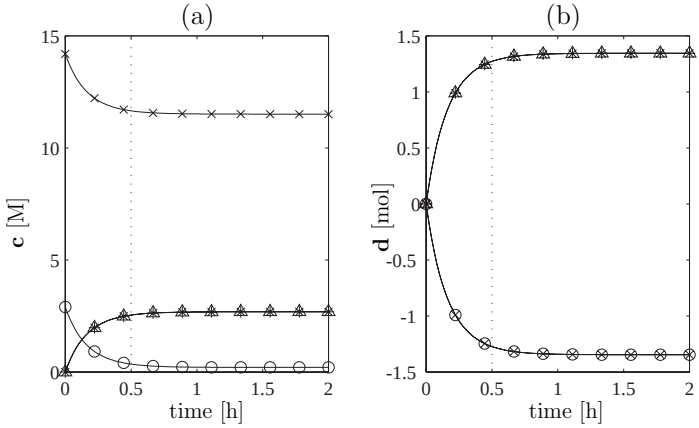

Figure G.6. (a) The concentrations $\mathbf{c}$, and (b) the RV-concentrations $\mathbf{d}$ of Example 2.2g. Termination of the first reaction at $\acute{t} = 0.5$ h (dotted vertical line) according to Figure G.7b. See Figure 5.8 for legend.

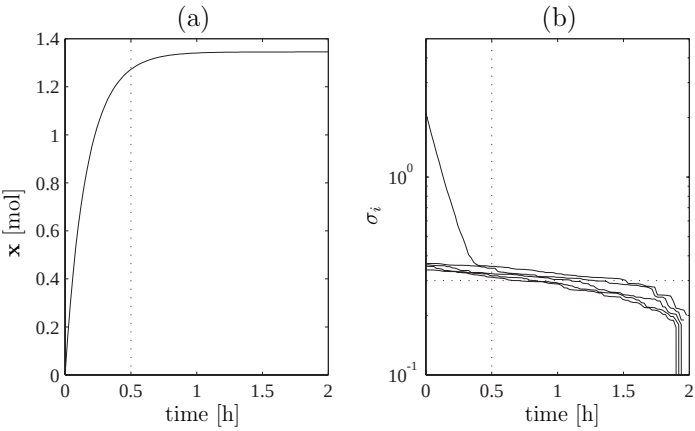

Figure G.7. Example 2.2g: (a) The reaction variants $\mathbf{x}$, and (b) EFA backward in time applied to column-mean-centered submatrices of $\tilde{\mathbf{H}}$ with approximate-rank estimation based on visual inspection of the singular values. Termination of the first reaction at $\acute{t} = 0.5$ h (dotted horizontal and vertical lines) under the assumption of noise level $\varepsilon_A = 0.03$.

values. Figure G.7b shows that the largest singular value decreases rapidly and disappears in the noise at level $\varepsilon_A = 0.03$. Thus, a local approximate-rank reduction of 1 can be estimated at time instant $\acute{t} = 0.5$ h. Under the assumption of reversible reactions, the local approximate-rank reduction can be interpreted as the reaction having reached the equilibrium. Owing to the bad signal/noise ratio, the local approximate-rank reduction takes place before the time instant determined visually from Figure G.6 (1 h).

Example 2.13e (Linearly-dependent reaction rates; cont'd from page 31**)**
Consider Run R3 of Example 2.13 on page 31 (constant-density semibatch reaction system). It is assumed that all reacting species absorb except for the solvent ($S = 5$). Figure G.8 shows the profiles of the concentrations and RV-concentrations, and Figure G.9 depicts the pure-component spectra of all absorbing species ($L = 80$), which are used to construct $\mathbf{A}$. Uniform noise in the range [-0.04, 0.04] is added to $\mathbf{A}$. The pseudo-rank of $\tilde{\mathbf{H}}$ is $2 = R$ as expected. Figure G.10a shows that the two (unknown) reaction variants become asymptotically parallel after about 5 h, meaning that the corresponding reaction rates become linearly dependent in the sense of Definition G.3. Thus, for $t \geq 5$ h, the produced species $\mathcal{X}_2$ is immediately consumed by the second reaction, and a pathway equivalent to (2.54) can be found as $2\mathcal{X}_1 \rightarrow \mathcal{X}_3 + \mathcal{X}_4 + 2\,\mathcal{X}_5$ (see also Example 2.4a on page 19).
Figure G.10b illustrates EFA backward in time applied to column-mean-centered submatrices of $\tilde{\mathbf{H}}$ with approximate-rank estimation based on visual inspection of the singular values. The second largest singular value decreases rapidly and disappears in the noise at level $\varepsilon_A = 7.6$. Thus, a local approximate-rank reduction of 1 occurs at time instant $\acute{t} = 3$ h. Under the assumption of an irreversible reaction system and the absence of limiting species concentration, this rank drop can be interpreted as the reaction having reached linearly-dependent reaction rates. Owing to the bad

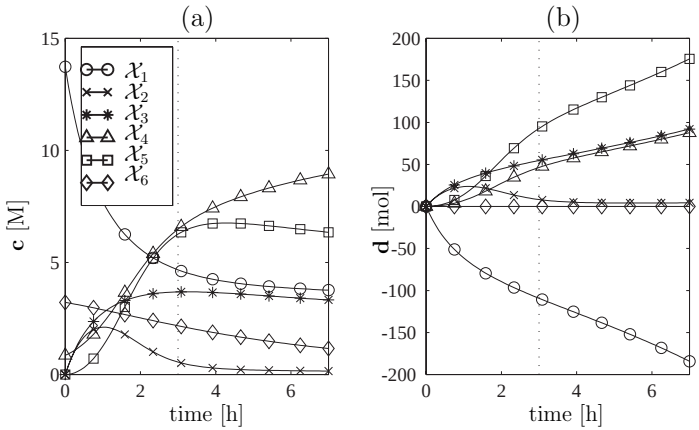

Figure G.8. (a) The concentrations **c**, and (b) the RV-concentrations **d** of Example 2.13e. Linearly-dependent reaction rates at $\acute{t} = 3$ h (dotted vertical line) according to Figure G.10b.

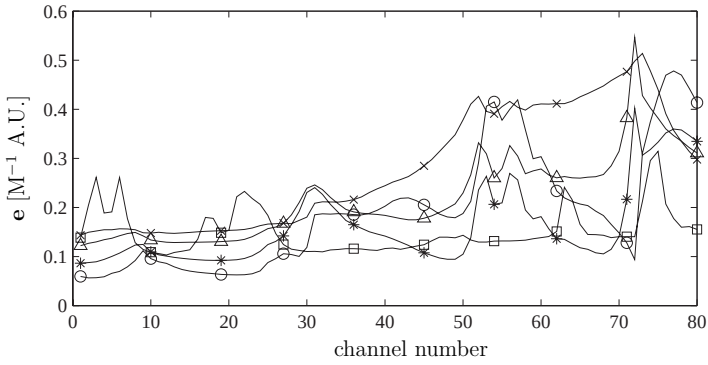

Figure G.9. Pure-component spectra **E** for Example 2.13e. For legend, see Figure G.8.

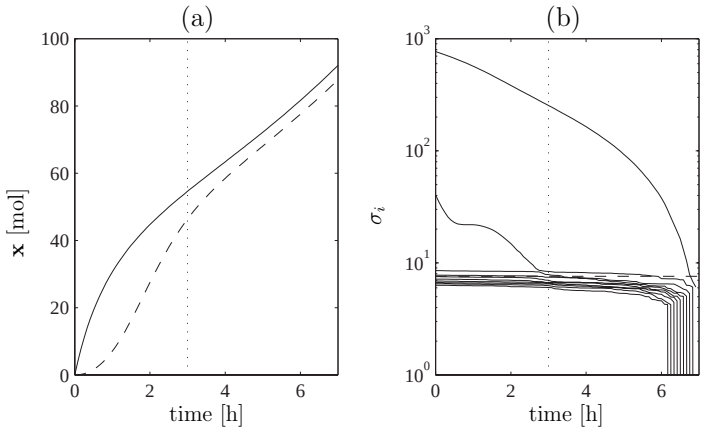

Figure G.10. Example 2.13e: (a) The reaction variants $\mathbf{x}$, and (b) EFA backward in time applied to column-mean-centered submatrices of $\tilde{\mathbf{H}}$ with approximate-rank estimation based on visual inspection of the singular values. Legend for (a): (–) first reaction, (– –) second reaction. Linearly-dependent reaction rates at $\acute{t} = 3$ h (dotted horizontal and vertical lines) under the assumption of noise level $\varepsilon_A = 7.6$.

signal/noise ratio, the local approximate-rank reduction takes place before the time instant determined visually Figure G.10a (5 h).

Example 2.13f (Auto-catalyzed reaction system; cont'd from page ??)
The second reaction is auto-catalyzed by $\mathcal{X}_4$. Since the concentration of $\mathcal{X}_4$ is small in the beginning, the corresponding (unknown) reaction variant is also small. With increasing concentration of $\mathcal{X}_2$ produced by the first reaction, more $\mathcal{X}_4$ is produced, which, in turn, increases the reaction rate of the second reaction.
Figure G.11b illustrates EFA forward in time applied to column-mean-centered submatrices of $\tilde{\mathbf{H}}$ with approximate-rank estimation based on visual inspection of the singular values. The second largest singular value becomes significant for time instants $t \geq \acute{t} = 0.6$ h. Under the assumption that the reaction system is irreversible and auto-catalyzed, this rank increase can be interpreted as the time instant where the second reaction rate becomes significant or, equivalently where both reaction rates become linearly independent.

G.2 Local rank in channel direction

In practical applications, the reacting species do not absorb in the *all* spectral regions of interest. There exist spectral regions where only a single reacting species absorbs such as discrete bands in FTIR, mass spectroscopy, Raman. These *selective* regions play an important role in the resolution of concentrations from spectral data using FA methods (see Tauler *et al.*, 1995, and the references therein). For non-reacting mixtures, local rank analysis in channel direction can detect such regions as rank-one

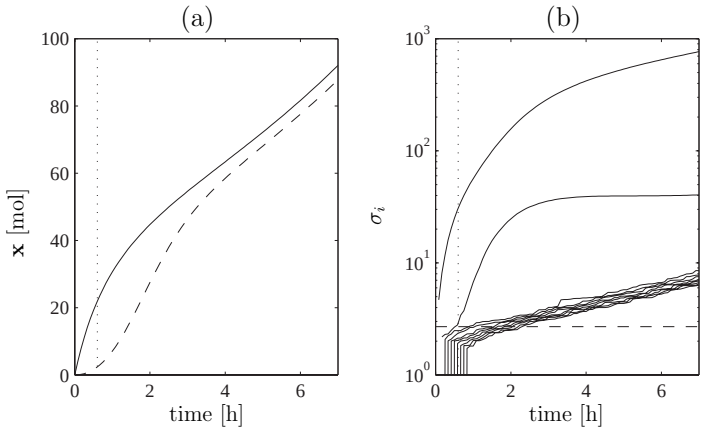

Figure G.11. Example 2.13f: (a) The reaction variants $\mathbf{x}$, and (b) EFA forward in time applied to column-mean-centered submatrices of $\hat{\mathbf{H}}$ with approximate-rank estimation based on visual inspection of the singular values. Legend for (a): (–) first reaction, (– –) second reaction. Linearly-independent reaction rates at $\acute{t} = 0.6$ h (dotted horizontal and vertical lines) under the assumption of noise level $\varepsilon_A = 2.7$.

regions (Geladi and Wold, 1987). For reacting mixtures, however, it will be shown that rank-one or, more generally, rank-A regions in channel direction can also occur for regions with more than one or A absorbing species, respectively. Thus, in contrast to non-reacting mixtures, prior knowledge is necessary to interpret rank-A regions as spectral regions with A absorbing species.

WEFA with window size of w channels is a useful tool for the detection of local rank regions in channel direction (see Appendix C.4.1.3). Let $\mathbf{A}(\ell, w)$ denote the corresponding spectral matrix in the spectral region $[\ell,\ \ell + w[$. Let $S_m(\ell, w)$ denote the number of species that locally absorb, and the subscript m the quantities related to these S_m species.

Thus, the factorization of spectral data for the local region becomes:

$$\mathbf{A}(\ell, w) = \mathbf{C}\,\boldsymbol{\mathcal{P}}_m(\ell, w)\,\mathbf{E}_m(\ell, w) = \mathbf{X}_e\,\mathbf{N}_e\,\boldsymbol{\mathcal{P}}_m(\ell, w)\,\mathbf{E}_m(\ell, w), \tag{G.11}$$

where $\boldsymbol{\mathcal{P}}_m$ is an $S \times S_m$ matrix that selects those S columns which correspond to the S_m species, and $\mathbf{E}_m$ the corresponding $S_m \times w$ pure-component spectra matrix. It is assumed that $\operatorname{rank}(\mathbf{E}_m(\ell, w)) = S_m(\ell, w)$, i.e., the pure-component spectra of the $S_m(\ell, w)$ in the spectral region are linearly independent, implying that $w \geq S_m(\ell, w)$. Note that $\operatorname{rank}(\boldsymbol{\mathcal{P}}_m(\ell, w)) = S_m(\ell, w)$. The following proposition is formulated.

Proposition G.4 *Let Assumptions A1–10 in Appendix D be verified, and* $\operatorname{rank}(\mathbf{E}_m(\ell, w)) = S_m(\ell, w)$. *If* $S_m = 1$, *then* $\operatorname{rank}(\mathbf{A}(\ell, w)) = 1$. *For reacting mixtures, if* $S_m > 1$, *then*

$$A = \operatorname{rank}(\mathbf{A}(\ell, w)) \quad \textit{does not imply} \quad S_m = A. \tag{G.12}$$

Proof:
For simplicity of notation, the arguments (ℓ, w) are suppressed for the proof. Since

$\text{rank}\,(\mathbf{E}_m) = S_m$, by invoking Property B.10a, it follows:

$$\text{rank}\,(\mathbf{A}) = \text{rank}\,(\mathbf{C}\,\boldsymbol{\mathcal{P}}_m) \leq \min(\text{rank}\,(\mathbf{X}_e)\,, \mathbf{N}_{e,m}\boldsymbol{\mathcal{P}}_m) \leq (R + p + 1, S_m). \tag{G.13}$$

If $S_m = 1$, then $\text{rank}\,(\mathbf{A}) = \text{rank}\,(\mathbf{c}_m) = 1$. Alternatively, if $S_m > 1$ and $(R + p + 1) < S_m$, then $\text{rank}\,(\mathbf{A}) < S_m$ follows from (G.13). □

For reacting and non-reacting mixtures, Proposition G.4 says that if only one species absorbs in region $[\ell,\, \ell+\mathcal{w}[$, then the rank of $\mathbf{A}(\ell, \mathcal{w})$ is also one. For reacting mixtures, however, if more than one species absorbs in a spectral region, then the rank of $\mathbf{A}(\ell, \mathcal{w})$ does not necessarily correspond to the number of absorbing species (see Section 5.2.1). More importantly, the rank of $\mathbf{A}(\ell, \mathcal{w})$ can even be 1. Thus, selective regions cannot unambiguously be determined from the local rank analysis of $\mathbf{A}$ in channel direction.

WEFA can also be performed simultaneously in both dimensions of $\mathbf{A}$. Then, Proposition G.4 still holds with the arguments $(\ell, \mathcal{w})$ replaced by $(\{k, \ell\}, \{\mathcal{w}_m, \mathcal{w}_n\})$, i.e., a two-dimensional window of size $\mathcal{w}_m \times \mathcal{w}_n$ slides over the entire two-dimensional matrix $\mathbf{A}$ with $\{k, \ell\}$ being the coordinates of the upper left corner of the window.

H

Proofs

H.1 Modeling chemical reaction systems

Proposition 2.7 on page 16:
The following properties result from (2.18)*:*

(a) The number of independent reactions obeys $R_N < S$.

(b) $(S - R_N)$ *linearly-independent columns of* $\mathbf{\Gamma}$ *are required to span the complete null space of* $\mathbf{N}$*, and thus,* $N \geq S - R_N$.

(c) For any $\gamma \times S$ *concentration matrix* $\check{\mathbf{C}}$*, if* $\operatorname{rank}\left(\check{\mathbf{C}}\,\mathbf{\Gamma}\right) = \gamma$*, then* $\operatorname{rank}\left(\left[\begin{smallmatrix}\mathbf{N}\\ \check{\mathbf{C}}\end{smallmatrix}\right]\right) = R_N + \gamma$.

(d) For any nonzero concentration vector $\mathbf{c}$*,* $\mathbf{c} \notin \mathcal{S}_r(\mathbf{N})$*, where* $\mathcal{S}_r(\cdot)$ *denotes the subspace spanned by the rows of a matrix (row space).*

Proof:
(a) Since $\mathbf{N}\,\mathbf{\Gamma} = \mathbf{0}_{R\times N}$, $\mathbf{\Gamma}$ lies in the null space of $\mathbf{N}$. Thus, $\operatorname{rank}(\mathbf{N}) + \operatorname{rank}(\mathbf{\Gamma}) \leq S$ (Waller and Mäkilä, 1981). Since $\operatorname{rank}(\mathbf{\Gamma}) > 0$ and $\operatorname{rank}(\mathbf{N}) = R_N$, $R_N < S$.

(b) Since $\operatorname{rank}(\mathbf{N}) = R_N$, $\dim(\mathcal{N}(\mathbf{N})) = S - R_N$, where $\dim(\mathcal{N}(\cdot))$ denotes the dimension of the null space of a matrix. Thus, $(S - R_N)$ linearly-independent vectors of dimension S are required to span the null space of $\mathbf{N}$. Since $\mathbf{\Gamma}$ lies in the null space of $\mathbf{N}$ (see 2.18), these vectors must lie in $\mathcal{S}_c(\mathbf{\Gamma})$.

(c) Proposition 2.7c will be shown by contradiction. Assume $\operatorname{rank}\left(\left[\begin{smallmatrix}\mathbf{N}\\ \check{\mathbf{C}}\end{smallmatrix}\right]\right) < R_N + \gamma$. Then, there exists a row of $\check{\mathbf{C}}$, $\mathbf{c}_i$, such that

$$\begin{aligned}\mathbf{c}_i^{\mathrm{T}} &= \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{N} + \sum_{j\neq i}\beta_j\mathbf{c}_j^{\mathrm{T}}\\ \text{or } \mathbf{c}_i^{\mathrm{T}}\mathbf{\Gamma} &= \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{N}\,\mathbf{\Gamma} + \sum_{j\neq i}\beta_j\mathbf{c}_j^{\mathrm{T}}\mathbf{\Gamma}.\end{aligned}$$

Since $\mathbf{N}\,\mathbf{\Gamma} = \mathbf{0}_{R\times N}$, $\mathbf{c}_i^{\mathrm{T}}\mathbf{\Gamma} = \sum_{j\neq i}\beta_j\mathbf{c}_j^{\mathrm{T}}\mathbf{\Gamma}$, which implies $\operatorname{rank}\left(\check{\mathbf{C}}\,\mathbf{\Gamma}\right) < \gamma$ contradicting the assumption of c.

(d) It is easy to verify that there exists a column of $\mathbf{\Gamma}$, $\boldsymbol{\gamma}$, that has no zero element (e.g. the molecular weights). Since $\mathbf{c}$ and $\boldsymbol{\gamma}$ have only nonnegative entries and, by assumption, $\mathbf{c} \neq \mathbf{0}_S$, $\mathbf{c}^{\mathrm{T}}\boldsymbol{\gamma} \neq 0$ follows immediately. Thus, $\operatorname{rank}(\mathbf{c}^{\mathrm{T}}\boldsymbol{\gamma}) = 1$ and, from c, $\operatorname{rank}\left(\left[\begin{smallmatrix}\mathbf{N}\\ \mathbf{c}^{\mathrm{T}}\end{smallmatrix}\right]\right) = R_N + 1$, implying $\mathbf{c} \notin \mathcal{S}_r(\mathbf{N})$. □

H.2 Reaction and flow variants/invariants in the dynamic model

Theorem 3.2 on page 39:
Let $\varsigma = \operatorname{rank}([\mathbf{N}^{\mathrm{T}}, \mathbf{C}_{in}]) = (R+p)$. *Then, a diffeomorphism* $\mathcal{T} : \left[\begin{smallmatrix}\boldsymbol{n}\\ V\end{smallmatrix}\right] \leftrightarrow \left[\begin{smallmatrix}\boldsymbol{x}\\ \lambda\end{smallmatrix}\right]$ *exists that transforms model* (2.38) *into:*

$$\begin{aligned}
\dot{\boldsymbol{x}}_1 &= \mathsf{h}(\boldsymbol{x}_2)\,\mathbf{r}(\boldsymbol{x}), & \boldsymbol{x}_1(0) &= \eta\,\mathbf{N}^{+\mathrm{T}}(\mathbf{I}_S - \mathbf{C}_{in}\,\mathbf{\mathcal{M}}^{\mathrm{T}})\boldsymbol{n}_0,\\
\dot{\boldsymbol{x}}_2 &= \mathbf{q}_{in}/\lambda, & \boldsymbol{x}_2(0) &= \eta\,\mathbf{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}_0,\\
\dot{\boldsymbol{x}}_3 &= \mathbf{0}_{S-\varsigma}, & \boldsymbol{x}_3(0) &= \eta\,\mathbf{\Omega}^{\mathrm{T}}\,\boldsymbol{n}_0,\\
\dot{\lambda} &= -q_{out}/\mathsf{h}(\boldsymbol{x}_2), & \lambda(0) &= 1/\eta,
\end{aligned} \tag{H.1}$$

where

$$\mathsf{h}(\boldsymbol{x}_2) = \eta\,(V_0 - \mathbf{1}_p^{\mathrm{T}}\mathbf{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}_0) + \mathbf{1}_p^{\mathrm{T}}\boldsymbol{x}_2, \tag{H.2}$$

and $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, *and* $\boldsymbol{x}_3$ *are vectors of dimension* R, p, *and* $(S-\varsigma)$, *respectively,* η *a nonzero arbitrary constant,* $\mathbf{r}$ *the* R*-dimensional reaction rate vector expressed in terms of* $\boldsymbol{x}$, $\mathbf{\mathcal{M}} = \mathbf{L}(\mathbf{C}_{in}^{\mathrm{T}}\,\mathbf{L})^{-1}$, $\mathbf{L} \in \mathbb{R}^{S\times(\varsigma-R)}$, *and* $\mathbf{\Omega} \in \mathbb{R}^{S\times(S-\varsigma)}$ *matrices with orthonormal columns that satisfy: (i)* $\operatorname{rank}([\mathbf{N}^{\mathrm{T}}, \mathbf{L}, \mathbf{\Omega}]) = S$, *(ii)* $\mathbf{N}^{\mathrm{T}}$, $\mathbf{L}$ *and* $\mathbf{\Omega}$ *are mutually orthogonal, (iii)* $\mathbf{\Omega}^{\mathrm{T}}\mathbf{C}_{in} = \mathbf{0}_{S-\varsigma\times p}$, *(iv)* $\mathbf{L}^{\mathrm{T}}\mathbf{C}_{in}$ *is invertible.*

The transformation to normal form $\mathcal{T}$ *is one-to-one and can be written as follows using* $\mathsf{g}(\boldsymbol{n}, V) = \eta\,(V_0 - \mathbf{1}_p^{\mathrm{T}}\mathbf{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}_0)/(V - \mathbf{1}_p^{\mathrm{T}}\mathbf{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n})$*:*

$$\begin{bmatrix}\boldsymbol{n}\\ V\end{bmatrix} \to \begin{bmatrix}\boldsymbol{x}\\ \lambda\end{bmatrix} \;:\quad \begin{bmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\\ \boldsymbol{x}_3\\ \lambda\end{bmatrix} = \begin{bmatrix}\mathsf{g}(\boldsymbol{n}, V)\,\mathbf{N}^{+\mathrm{T}}(\mathbf{I}_S - \mathbf{C}_{in}\,\mathbf{\mathcal{M}}^{\mathrm{T}})\boldsymbol{n}\\ \mathsf{g}(\boldsymbol{n}, V)\,\mathbf{\mathcal{M}}^{\mathrm{T}}\boldsymbol{n}\\ \mathsf{g}(\boldsymbol{n}, V)\,\mathbf{\Omega}^{\mathrm{T}}\,\boldsymbol{n}\\ 1/\mathsf{g}(\boldsymbol{n}, V)\end{bmatrix} \tag{H.3}$$

$$\begin{bmatrix}\boldsymbol{x}\\ \lambda\end{bmatrix} \to \begin{bmatrix}\boldsymbol{n}\\ V\end{bmatrix} \;:\quad \begin{bmatrix}\boldsymbol{n}\\ V\end{bmatrix} = \begin{bmatrix}\lambda\,(\mathbf{N}^{\mathrm{T}}\,\boldsymbol{x}_1 + \mathbf{C}_{in}\,\boldsymbol{x}_2 + \mathbf{\Omega}\boldsymbol{x}_3)\\ \lambda\,\mathsf{h}(\boldsymbol{x}_2)\end{bmatrix}. \tag{H.4}$$

Proof:
The construction of the transformation to normal form was shown in the text. However, (i) the global nature of the transformation and (ii) the reconstruction of the original states $\boldsymbol{n}$ and V from $\boldsymbol{x}$ and λ remain to be proven.

(i) To proof the global nature of $\mathcal{T}$, it will be proven that $\mathsf{g}(\boldsymbol{n}, V)$ is always defined. Let $a \equiv V - \mathbf{1}_p^{\mathrm{T}} \boldsymbol{\mathcal{M}}^{\mathrm{T}} \boldsymbol{n}$ and $\zeta \equiv V_0 - \mathbf{1}_p^{\mathrm{T}} \boldsymbol{\mathcal{M}}^{\mathrm{T}} \boldsymbol{n}_0$. Since from (H.3), $a = \eta\, \zeta\, \lambda$. Differentiating the latter equation with respect to time, and noting $V = \mathsf{h}(\boldsymbol{x}_2)\, \lambda$, it follows that

$$\begin{aligned} \dot{a} = \eta\, \zeta\, \dot{\lambda} &= -\eta\, \zeta\, q_{out}/\mathsf{h}(\boldsymbol{x}_2) = (\eta\, \zeta \lambda)\, q_{out}/(\mathsf{h}(\boldsymbol{x}_2)\, \lambda) \\ &= -\frac{q_{out}}{V}\, a, \qquad a(0) = V_0 - \mathbf{1}_p^{\mathrm{T}} \boldsymbol{\mathcal{M}}^{\mathrm{T}} \boldsymbol{n}_0. \end{aligned} \tag{H.5}$$

Since q_{out} and V are positive, $a(t) = V(t) - \mathbf{1}_p^{\mathrm{T}} \boldsymbol{\mathcal{M}}^{\mathrm{T}} \boldsymbol{n}(t) \neq 0$ for all t if $a(0) = V_0 - \mathbf{1}_p^{\mathrm{T}} \boldsymbol{\mathcal{M}}^{\mathrm{T}} \boldsymbol{n}_0 \neq 0$. If $a(0) = V_0 - \mathbf{1}_p^{\mathrm{T}} \boldsymbol{\mathcal{M}}^{\mathrm{T}} \boldsymbol{n}_0 = 0$, then from (H.5), $a(t) = V(t) - \mathbf{1}_p^{\mathrm{T}} \boldsymbol{\mathcal{M}}^{\mathrm{T}} \boldsymbol{n}(t) = 0$ for all t. In such a case, g is defined by invoking de l'Hospital's rule.

(ii) The relationship $\left[\begin{smallmatrix} \boldsymbol{z} \\ V \end{smallmatrix}\right] \rightarrow \left[\begin{smallmatrix} \boldsymbol{x} \\ \lambda \end{smallmatrix}\right]$ is given in (3.8). The inverse can be obtained after a few algebraic manipulations as:

$$\begin{bmatrix} \boldsymbol{z} \\ V \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}\, \lambda \\ \mathsf{h}(\boldsymbol{x}_2)\, \lambda \end{bmatrix}. \tag{H.6}$$

According to (3.6), the number of moles $\boldsymbol{n}$ satisfies:

$$\boldsymbol{z} = \begin{bmatrix} \mathbf{N}^{+\mathrm{T}} \left(\mathbf{I}_S - \mathbf{C}_{in} \boldsymbol{\mathcal{M}}^{\mathrm{T}}\right) \\ \boldsymbol{\mathcal{M}}^{\mathrm{T}} \\ \boldsymbol{\mathcal{Q}}^{\mathrm{T}} \end{bmatrix} \boldsymbol{n}. \tag{H.7}$$

Multiplying both sides of (H.7) by $[\mathbf{N}^{\mathrm{T}},\, \mathbf{C}_{in},\, \boldsymbol{\mathcal{Q}}]$ and noting that $\mathbf{N}^{\mathrm{T}}\, \mathbf{N}^{+\mathrm{T}} + \mathbf{L}\, \mathbf{L}^{\mathrm{T}} + \boldsymbol{\mathcal{Q}}\, \boldsymbol{\mathcal{Q}}^{\mathrm{T}} = \mathbf{I}_S$,

$$\begin{aligned} \begin{bmatrix} \mathbf{N}^{\mathrm{T}} & \mathbf{C}_{in} & \boldsymbol{\mathcal{Q}} \end{bmatrix} \boldsymbol{z} &= \left(\mathbf{N}^{\mathrm{T}}\, \mathbf{N}^{+\mathrm{T}} \left(\mathbf{I}_S - \mathbf{C}_{in} \boldsymbol{\mathcal{M}}^{\mathrm{T}}\right) + \mathbf{C}_{in}\, \boldsymbol{\mathcal{M}}^{\mathrm{T}} + \boldsymbol{\mathcal{Q}}\, \boldsymbol{\mathcal{Q}}^{\mathrm{T}}\right) \boldsymbol{n} \\ &= \left(\mathbf{N}^{\mathrm{T}}\, \mathbf{N}^{+\mathrm{T}} + \left(\mathbf{L}\, \mathbf{L}^{\mathrm{T}} + \boldsymbol{\mathcal{Q}}\, \boldsymbol{\mathcal{Q}}^{\mathrm{T}}\right) \mathbf{C}_{in} \boldsymbol{\mathcal{M}}^{\mathrm{T}} + \boldsymbol{\mathcal{Q}}\, \boldsymbol{\mathcal{Q}}^{\mathrm{T}}\right) \boldsymbol{n} \\ &= \left(\mathbf{N}^{\mathrm{T}}\, \mathbf{N}^{+\mathrm{T}} + \mathbf{L}\, \mathbf{L}^{\mathrm{T}} + \boldsymbol{\mathcal{Q}}\, \boldsymbol{\mathcal{Q}}^{\mathrm{T}}\right) \boldsymbol{n} = \boldsymbol{n}. \end{aligned} \tag{H.8}$$

Substituting $\boldsymbol{z} = \boldsymbol{x}\, \lambda$ into (H.8) leads to $\boldsymbol{n} = \lambda \left(\mathbf{N}^{\mathrm{T}}\, \boldsymbol{x}_1 + \mathbf{C}_{in}\, \boldsymbol{x}_2 + \boldsymbol{\mathcal{Q}} \boldsymbol{x}_3\right)$ and, hence, the theorem follows. □

Corollary 3.3 on page 40:
For $\varsigma < R + p$, the transformed model (3.9) *represents* (2.38)*, where* $\boldsymbol{\mathcal{M}} := \mathbf{L}\, (\mathbf{C}_{in}^{\mathrm{T}}\, \mathbf{L})^{+}$. *The relation* $\left[\begin{smallmatrix} \boldsymbol{x} \\ \lambda \end{smallmatrix}\right] \rightarrow \left[\begin{smallmatrix} \boldsymbol{n} \\ V \end{smallmatrix}\right]$ *is defined by* (3.12).

Proof:
The only difference between Corollary 3.3 and Theorem 3.2 is that $\boldsymbol{\mathcal{M}} = \mathbf{L}(\mathbf{C}_{in}^{\mathrm{T}}\, \mathbf{L})^{+}$. Since $\mathbf{L}\mathbf{L}^{\mathrm{T}} \mathbf{C}_{in} \boldsymbol{\mathcal{M}}^{\mathrm{T}} = \mathbf{L}(\mathbf{L}^{\mathrm{T}} \mathbf{C}_{in})(\mathbf{L}^{\mathrm{T}} \mathbf{C}_{in})^{+} \mathbf{L}^{\mathrm{T}} = \mathbf{L}\mathbf{L}^{\mathrm{T}}$, however, the equation $\mathbf{N}^{\mathrm{T}} \mathbf{N}^{+\mathrm{T}} + \mathbf{L}\mathbf{L}^{\mathrm{T}} + \boldsymbol{\mathcal{Q}}\, \boldsymbol{\mathcal{Q}}^{\mathrm{T}} = \mathbf{I}_S$ in the proof of Theorem 3.2 still holds and, hence, the corollary follows. □

H.3 Reaction and flow variants/invariants in the factorization of concentration data

H.3.1 Factorization of concentration data

Theorem 4.1 on page 55:
Let Assumptions A1–7 in Appendix D be verified. The concentrations $\mathbf{c}(k)$ *from reaction systems described by* (2.38) *are factorized as*

$$\mathbf{c}^{\mathrm{T}}(k) = \frac{1}{\mathrm{h}(\mathbf{z}(k))} \left[\begin{array}{c|c|c} \mathbf{x}^{\mathrm{T}}(k) & \mathbf{z}^{\mathrm{T}}(k) & V_0 \end{array}\right] \left[\begin{array}{c} \mathbf{N} \\ \hline \mathbf{C}_{in}^{\mathrm{T}} \\ \hline \mathbf{c}_0^{\mathrm{T}} \end{array}\right] \equiv \mathbf{x}_e^{\mathrm{T}} \mathbf{N}_e \tag{H.9}$$

with $\mathrm{h}(\mathbf{z}(k)) \equiv V_0 + \mathbf{1}_p^{\mathrm{T}} \mathbf{z}(k)$,

where $\mathbf{x}_e$ *is an* $(R+p+1)$*-dimensional vector and* $\mathbf{N}_e$ *is the* extended stoichiometric matrix *of dimension* $(R+p+1) \times S$. (4.3) *is termed* the factorization of concentration data*. For* K *observations, the factorization of concentration data becomes:*

$$\mathbf{C} = \left\{ \mathcal{H}^{-1} \left[\begin{array}{c|c|c} \mathbf{X} & \mathbf{Z} & V_0 \, \mathbf{1}_K \end{array}\right] \right\} \left[\begin{array}{c} \mathbf{N} \\ \hline \mathbf{C}_{in}^{\mathrm{T}} \\ \hline \mathbf{c}_0^{\mathrm{T}} \end{array}\right] = \mathbf{X}_e \, \mathbf{N}_e \tag{H.10}$$

with $\mathcal{H} \equiv V_0 \, \mathbf{I}_K + \mathrm{diag}\,(\mathbf{Z}\,\mathbf{1}_p)$

and $\mathbf{X}$ *being the* $K \times R$ *matrix of reaction variants,* $\mathbf{Z}$ *the* $K \times p$ *matrix of reaction invariants* and *flow variants,* $V_0 \mathbf{1}_K$ *the* K*-dimensional vector of reaction and flow invariants,* $\mathbf{X}_e$ *an* $K \times (R+p+1)$ *matrix, and* $\mathcal{H}$ *a* K*-dimensional diagonal scaling matrix computed from* (4.5) *for* K *observations.*

Proof:
For simplicity of notation, the dependency on time instant k is suppressed in the proof.
With $\eta = 1$, $\mathrm{h}(\boldsymbol{x}_2) = \lambda/V$, and (4.2) substituted into (4.1),

$$\mathbf{c} = \left\{ \mathbf{N}^{\mathrm{T}} \mathbf{x} + \mathbf{C}_{in} \mathbf{z} + V_0 \left[\mathbf{N}^{\mathrm{T}} \mathbf{N}^{+\mathrm{T}} (\mathbf{I}_S - \mathbf{C}_{in} \mathcal{M}^{\mathrm{T}}) + \mathbf{C}_{in} \mathcal{M}^{\mathrm{T}} + \mathcal{Q} \mathcal{Q}^{\mathrm{T}} \right] \mathbf{c}_0 \right\} / \mathrm{h}(\mathbf{z}).$$

From (H.8), $(\mathbf{N}^{\mathrm{T}} \mathbf{N}^{+\mathrm{T}} (\mathbf{I}_S - \mathbf{C}_{in} \mathcal{M}^{\mathrm{T}}) + \mathbf{C}_{in} \mathcal{M}^{\mathrm{T}} + \mathcal{Q} \mathcal{Q}^{\mathrm{T}}) = \mathbf{I}_S$, and thus, (H.9) holds. (H.10) is obtained from (H.9) for K observations. □

Theorem 4.2 on page 60:
Let Assumptions A1–7 in Appendix D be verified. If $\mathbf{C}_{in}$*,* V_0*,* $\mathbf{c}(k)$*,* $\mathbf{q}_{in}(k)$*, and* $q_{out}(k)$ *are known/measured, then an* S*-dimensional concentration vector in reaction-variant form (RV-concentration vector),* $\mathbf{d}(k)$*, that relates directly to* $\mathbf{x}(k)$ *and* $\mathbf{N}$ *can be defined from* (4.3)*:*

$$\mathbf{d}(k) \equiv \mathrm{h}(\mathbf{z}(k)) \, \mathbf{c}(k) - \mathbf{C}_{in} \, \mathbf{z}(k) - V_0 \, \mathbf{c}_0 = \mathbf{N}^{\mathrm{T}} \, \mathbf{x}(k). \tag{H.11}$$

The operation to this effect is termed data pre-treatment to reaction-variant form. *For K observations, a $K \times S$ concentration matrix in reaction-variant form (RV-concentration matrix),* $\mathbf{D}$*, can be constructed from* (4.10) *that relates directly to* $\mathbf{X}$ *and* $\mathbf{N}$*:*

$$\mathbf{D} \equiv \mathcal{H}\mathbf{C} - \mathbf{Z}\,\mathbf{C}_{in}^{\mathrm{T}} - V_0\,\mathbf{1}_K\,\mathbf{c}_0^{\mathrm{T}} = \mathcal{H}\mathbf{C} - \mathbf{Z}_x\,\mathbf{C}_x = \mathbf{X}\,\mathbf{N}. \tag{H.12}$$

(4.11) *is termed* factorization of concentration data in reaction-variant form (factorization of RV-concentration data). *Proof:*
Since V_0, $\mathbf{q}_{in}(k)$, and $q_{out}(k)$ are assumed to be known, $V(k)$ can be computed from (2.38). $\mathbf{z}(k)$ and $\lambda(k)$ are computed from (4.7). Then, (H.11) is derived from (H.9) by simple algebraic manipulations. □

H.3.2 Rank

Theorem 4.3 on page 64:
Let Assumptions A1–7 in Appendix D be verified. Then, the rank of $\mathbf{C}$ *is bounded by:*

$$R + 1 \leq \mathrm{rank}\,(\mathbf{C}) \leq \min(R + p + 1, S). \tag{H.13}$$

Proof: According to Property B.4 and Assumption A2.6–7, $\mathrm{rank}\,(\mathbf{C}) \leq \min(\mathrm{rank}\,(\mathbf{X}_e), \mathrm{rank}\,(\mathbf{N}_e)) = \min(R + p + 1, S)$.

Let $\check{\mathbf{C}} \equiv \begin{bmatrix} \mathbf{X} & V_0\mathbf{1}_K \end{bmatrix} \begin{bmatrix} \mathbf{N} \\ \mathbf{c}_0^{\mathrm{T}} \end{bmatrix}$. For the proof of $\mathrm{rank}\,(\mathbf{C}) \geq R{+}1$, note that: (i) $\mathbf{c}_0 \notin \mathcal{S}_r\,(\mathbf{N})$ (see Proposition 2.7d), and (ii) $\mathbf{1}_K \notin \mathcal{S}_c\,(\mathbf{X})$ due to the first of $\mathbf{X}$ being zero. From (i) follows that $\mathrm{rank}\,(\mathbf{N}_e) \geq R{+}1$, and from (ii) $\mathrm{rank}\,(\mathbf{X}_e) \geq R{+}1$. Thus, $\mathrm{rank}\,\big(\check{\mathbf{C}}\big) = R{+}1$ by invoking Property B.10. Since $\mathrm{rank}\,(\mathbf{C}) \geq \mathrm{rank}\,\big(\check{\mathbf{C}}\big)$, $\mathrm{rank}\,(\mathbf{C}) \geq R + 1$ follows. □

Proposition 4.5 on page 65:
Let Assumptions A1–7 in Appendix D be verified. If the p generalized inlets are of the type

(a) additional batch runs,
(b) impulse additions at different time instants, or
(c) inlet flowrates with abrupt changes at different time instants,

then $\mathrm{rank}\,(\mathbf{X}_e) = R + p + 1$.

Proof:
The conditions for $\mathrm{rank}\,(\mathbf{X}_e) = (R{+}p{+}1)$ are proven individually for each generalized inlet type.
(a) *additional batch runs:* For p additional batch runs with an optional outlet, K^i the number of observations of the ith batch run ($i = 1, \ldots, p + 1$), and $K = \sum_{i=1}^{p+1} K_i$, it

can be easily verified that

$$\mathbf{X}_e^\star = \mathcal{H}\mathbf{X}_e = \begin{bmatrix} \mathbf{X}^1 & \mathbf{0}_{K^1} & \mathbf{0}_{K^1} & \cdots & \mathbf{0}_{K^1} & V_0^1\,\mathbf{1}_{K^1} \\ \mathbf{X}^2 & V_0^2\,\mathbf{1}_{K^2} & \mathbf{0}_{K^2} & \cdots & \mathbf{0}_{K^2} & \mathbf{0}_{K^2} \\ \mathbf{X}^3 & \mathbf{0}_{K^3} & V_0^3\,\mathbf{1}_{K^3} & \cdots & \mathbf{0}_{K^3} & \mathbf{0}_{K^3} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{X}^{p+1} & \mathbf{0}_{K^{p+1}} & \mathbf{0}_{K^{p+1}} & \cdots & V_0^{p+1}\,\mathbf{1}_{K^{p+1}} & \mathbf{0}_{K^{p+1}} \end{bmatrix} \tag{H.14}$$

with $\mathbf{X}_e^\star$ being an $K \times (R+p+1)$ matrix. Since the first row of $\mathbf{X}^j$ is $\mathbf{0}_R^{\mathrm{T}}$, $\mathbf{1}_{K^j} \notin \mathcal{S}_c\,(\mathbf{X}^j)$ for $j = 1, \ldots, p+1$. Thus, $\mathbf{X}_e^\star$ is of full rank. Since $\mathbf{V}\mathbf{\Lambda}^{-1}$ is also full rank, then by invoking Property B.10a, $\mathbf{X}_e$ is full rank.

(b) *impulse additions of single species or mixtures at different time instants:* For p impulse additions, let K^i denote the number of observations before the ith impulse addition ($i = 1, \ldots, p$) and K^{p+1} the number of observations after the pth impulse addition. It can be easily verified that

$$\mathbf{X}_e^\star = \mathcal{H}\mathbf{X}_e = \begin{bmatrix} \mathbf{X}^1 & \mathbf{0}_{K^1} & \mathbf{0}_{K^1} & \cdots & \mathbf{0}_{K^1} & V_0^1\,\mathbf{1}_{K^1} \\ \mathbf{X}^2 & V_\Delta^1\,\mathbf{1}_{K^2} & \mathbf{0}_{K^2} & \cdots & \mathbf{0}_{K^2} & V_0^1\,\mathbf{1}_{K^2} \\ \mathbf{X}^3 & V_\Delta^1\,\mathbf{1}_{K^3} & V_\Delta^2\,\mathbf{1}_{K^3} & \cdots & \mathbf{0}_{K^3} & V_0^1\,\mathbf{1}_{K^3} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{X}^{p+1} & V_\Delta^1\,\mathbf{1}_{K^{p+1}} & V_\Delta^2\,\mathbf{1}_{K^{p+1}} & \cdots & V_\Delta^p\,\mathbf{1}_{K^{p+1}} & V_0^1\,\mathbf{1}_{K^{p+1}} \end{bmatrix} \tag{H.15}$$

with $\mathbf{X}_e^\star$ being an $K \times (R+p+1)$ matrix. Since the first row of $\mathbf{X}^1$ is $\mathbf{0}_R^{\mathrm{T}}$, it follows that $\mathbf{1}_{K^1} \notin \mathcal{S}_c\,(\mathbf{X}^1)$. Thus, the $(R+p+1)$th column of $\mathbf{X}_e^\star$ is independent of its first R columns. Also, assuming the number of rows of $\mathbf{X}^1$ to be greater than $(R+1)$, it is concluded that $[\mathbf{X}^1,\, \mathbf{1}_K]$ is of full rank $(R+1)$. Thus, the vector $\mathbf{0}_K$ of the $(R+1)$th column of $\mathbf{X}_e^\star$ can be expressed only as a trivial combination (all coordinates are zero) of both the first R columns and the $(R+p+1)$th column of $\mathbf{X}_e^\star$. Thus, the $(R+1)$th column is linearly independent of both the first R columns and the $(R+p+1)$th column of $\mathbf{X}_e^\star$. By induction, $\mathbf{X}_e^\star$ is of full rank. Since $\mathbf{V}\mathbf{\Lambda}^{-1}$ is also full rank, then by invoking Property B.10a, $\mathbf{X}_e$ is full rank.

(c) *inlet flowrates with abrupt changes at different time instants:* It is easy to guarantee experimentally that $\mathrm{rank}\,([\mathbf{Z},\, \mathbf{1}_K]) = p+1$. Furthermore, $\mathbf{1}_K$ is trivially independent of $\mathbf{X}$, since the first row of $\mathbf{X}$ is $\mathbf{0}_R^{K^{\mathrm{T}}}$ and, therefore, $\mathrm{rank}\,([\mathbf{X},\, \mathbf{1}_K]) = R+1$. The crucial part is to verify that the columns of $\mathbf{X}$ are independent of the ones of $\mathbf{Z}$. However, this can easily be guaranteed if all inlet flowrates are changed abruptly at least over one time interval. Then, the columns of $\mathbf{X}$ are independent of those of $\mathbf{Z}$ ($\mathrm{rank}\,(\mathbf{Z}) = p$), since even with a discontinuous inlet, reaction rates do not exhibit a discontinuity. □

Proposition 4.6 on page 66:
Let Assumptions A1–7 in Appendix D be verified. Then,

$$\mathrm{rank}\,(\mathbf{D}) = R. \tag{H.16}$$

Proof:
From (4.11) follows that $\text{rank}\,(\mathbf{D}) \leq \min(\text{rank}\,(\mathbf{X}), \text{rank}\,(\mathbf{N}))$. It is obvious that $\text{rank}\,(\mathbf{N}) = \text{rank}\,(\mathbf{X}) = R$ due to the independence of reaction stoichiometries and kinetics. Thus, $\text{rank}\,(\mathbf{D}) \leq R$. The equality in (H.16) comes from Property B.10. □

Proposition 4.7 on page 67:
Let Assumptions A1–7 in Appendix D be verified. Let the number of observed *independent reactions be defined by* $R_o \equiv \text{rank}\,(\mathbf{D}_m)$. *Then,*

$$R_o \leq \text{rank}\,(\mathbf{C}_m) \leq \min(R_o + p + 1, S_m). \tag{H.17}$$

Proof:
$\text{rank}\,(\mathbf{D}_m) = R_o$ implies $\text{rank}\,(\mathbf{X}_o) = \text{rank}\,(\mathbf{N}_o) = R_o$. Since $\mathbf{C}_m = \mathbf{X}_{e,o}\,\mathbf{N}_{e,o}$, $\text{rank}\,(\mathbf{X}_{e,o}) \leq R_o + p + 1, S_m$, and $\text{rank}\,(\mathbf{N}_{e,o}) \leq \min(R_o + p + 1, S_m)$, $\text{rank}\,(\mathbf{C}_m) \leq \min(\text{rank}\,(\mathbf{X}_{e,o}, \mathbf{N}_{e,o})) \leq \min(R_o + p + 1, S_m)$. The lower bound of the rank of $\mathbf{C}_m$ says that in some cases, $\text{rank}\,(\mathbf{C}_m) = \text{rank}\,(\mathbf{D}_m) = R_o$. The rank of $\mathbf{C}_m$ can never be less than that of $\text{rank}\,(\mathbf{D}_m)$, since addition of columns to $\mathbf{X}_o$ to give $\mathbf{X}_{e,o}$ and addition of rows to $\mathbf{N}_o$ to give $\mathbf{N}_{e,o}$ can never decrease the rank. □

Proposition 4.9 on page 67:
Let Assumptions A1–7 in Appendix D be verified. Then, for $S_m \leq S$ *measured species, the nullities of* $\mathbf{C}_m$ *and* $\mathbf{D}_m$ *are given by:*

$$\begin{aligned} \max(S_m - R_o - S_{x,k}, 0) \leq \dim\left(\mathcal{N}\left(\mathbf{C}_m\right)\right) = S_m - \text{rank}\,(\mathbf{C}_m) \leq S_m - R_o, \\ \dim\left(\mathcal{N}\left(\mathbf{D}_m\right)\right) = S_m - \text{rank}\,(\mathbf{D}_m) = S_m - R_o. \end{aligned} \tag{H.18}$$

If $S_m = S$, *then*

$$\begin{aligned} \max(S - R - p - 1, 0) \leq \dim\left(\mathcal{N}\left(\mathbf{C}\right)\right) \leq S - R - 1, \\ \dim\left(\mathcal{N}\left(\mathbf{D}\right)\right) = S - R. \end{aligned} \tag{H.19}$$

Proof:
From (B.1) follows $\dim\left(\mathcal{N}\left(\mathbf{C}_m\right)\right) = S_m - \text{rank}\,(\mathbf{C}_m)$ and $\dim\left(\mathcal{N}\left(\mathbf{D}_m\right)\right) = S_m - \text{rank}\,(\mathbf{D}_m)$. Since $\text{rank}\,(\mathbf{C}_m) \leq \min(R_o + S_{x,k}, S_m)$ (see (H.17)) and $\text{rank}\,(\mathbf{D}_m) = R_o$, equation (H.18) follows. Since, for $S_m = S$, $(R + 1) \leq \text{rank}\,(\mathbf{C}) \leq \min(S, R + p + 1)$ (see 4.13) and $\text{rank}\,(\mathbf{D}) = R$ (see H.16), equation (H.19) follows. □

Proposition 4.11 on page 68:
Let Assumptions A1–7 in Appendix D be verified. Let $S_- = S_+ = S$ *and* $R_- = R_+ = R$, *where the subscripts* $-$ *and* $+$ *denote a quantity before and after variation using any of Variation modes V1–V3 on page 28, respectively. Then, the following properties hold for the nullity change* Δd*:*

(a) *For* $(p_+ - p_-)$ *additional generalized inlets,* $\Delta d \leq 0$.

(b) *If, in addition to (a),* $\text{rank}\,(\mathbf{C}_-) = (R + p_- + 1)$ *and* $\text{rank}\,(\mathbf{C}_+) = (R + p_+ + 1)$, *then* $\Delta d = p_- - p_+ < 0$.

(c) *If* $\text{rank}\,(\mathbf{C}_+) = S$, *then* $d_+ = 0$ *and* $\Delta d = -d_-$.

Proof:
(a) $\mathbf{C}_- = \mathbf{X}_{e,-}\mathbf{N}_{e,-}$ and $\mathbf{C}_+ = \mathbf{X}_{e,+}\mathbf{N}_{e,+}$. Variation modes V1–V3 append columns and/or rows to $\mathbf{X}_{e,-}$, and append columns to $\mathbf{N}_{e,-}$. Such an operation never decreases the ranks of $\mathbf{X}_e$ and $\mathbf{N}_e$ and, thus, $\operatorname{rank}(\mathbf{X}_{e,+}) \geq \operatorname{rank}(\mathbf{X}_{e,-})$ and $\operatorname{rank}(\mathbf{N}_{e,+}) \geq \operatorname{rank}(\mathbf{N}_{e,-})$. Thus, $\operatorname{rank}(\mathbf{C}_+) \geq \operatorname{rank}(\mathbf{C}_-)$. Furthermore,

$$\Delta d = d_+ - d_- = (S - \operatorname{rank}(\mathbf{C}_+)) - (S - \operatorname{rank}(\mathbf{C}_-)) = \operatorname{rank}(\mathbf{C}_-) - \operatorname{rank}(\mathbf{C}_+). \tag{H.20}$$

Thus, $\Delta d \leq 0$.

(b) Proposition 4.11 follows from the substitution of the two expressions $\operatorname{rank}(\mathbf{C}_-) = (R + p_- + 1)$ and $\operatorname{rank}(\mathbf{C}_+) = (R + p_+ + 1)$ into (H.20).

(c) Since $\operatorname{rank}(\mathbf{C}_+) = S$, it follows that $d_+ = \dim(\mathcal{N}(\mathbf{C}_+)) = 0$ and $\Delta d = d_+ - d_- = -d_-$. □

Proposition 4.12 on page 71:
Let Assumptions A1–7 in Appendix D be verified, and the subscripts − and + denote a quantity related to more and fewer measurements, respectively. If $S_- = S$ and $\operatorname{rank}(\mathbf{C}_-) = R + p + 1$, then measuring less species ($S_+ = S_m < S$) leads to the nullity reduction Δd of the corresponding concentration matrices $\mathbf{C}_-$ and $\mathbf{C}_+$:

$$\max(S_m - S + R - R_o, R + p + 1 - S) \leq \Delta d \leq S_m - S + R - R_o + p + 1. \tag{H.21}$$

Proof:
From the assumptions $S_- = S$ and $\operatorname{rank}(\mathbf{C}_-) = R + p + 1$, it follows that $d_- = S - \operatorname{rank}(\mathbf{C}_-) = S - R - p - 1$. From (4.23) follows $\max(S_m - R_o - p - 1, 0) \leq d_+ \leq S_m - R_o$. Thus, from $\delta d = d_+ - d_-$ follows (H.21). □

Proposition 4.13 on page 71:
Let Assumptions A1–7 in Appendix D be verified. Then,

$$\begin{aligned} 1 &\leq \Delta d(\mathbf{C}, \mathbf{D}) \leq \min(p + 1, S - R), \\ 0 &\leq \Delta d(\mathbf{C}_m, \mathbf{D}_m) \leq \min(p + 1, S_m - R_o). \end{aligned} \tag{H.22}$$

Proof:
From $S - \min(R+p+1, S) \leq d_- = \dim(\mathcal{N}(\mathbf{C})) \leq (S - R - 1)$ and $d_+ = \dim(\mathcal{N}(\mathbf{D})) = S - R$ follows that $1 \leq \Delta d(\mathbf{C}, \mathbf{D}) \leq (S-R) - (S - \min(R+p+1, S)) = \min(p+1, S-R)$. Similarly, from $S_m - \min(R_o + p + 1, S_m) \leq d_- = \dim(\mathcal{N}(\mathbf{C}_m)) \leq (S_m - R_o)$ and $d_+ = \dim(\mathcal{N}(\mathbf{D}_m)) = S_m - R_o$ follows that $0 \leq \Delta d(\mathbf{C}_m, \mathbf{D}_m) \leq (S_m - R_o) - (S_m - \min(R_o + p + 1, S_m)) = \min(p + 1, S_m - R_o)$. □

Proposition 4.22 on page 89:
Let Assumptions A1–4 and A6–7 in Appendix D be verified. Furthermore, assume that concentration measurements for $S_m < S$ species and temperature/calorimetric measurements from reaction system (2.40) are available, and the R_o observed independent reactions have a significant heat effect. Then, the following properties hold:

(a) *If* $(-\Delta\mathbf{h}_R) \notin \mathcal{S}_c(\mathbf{N}_m)$, *then*

$$\operatorname{rank}(\mathbf{D}_{T,m}) = \operatorname{rank}(\mathbf{D}_m) + 1 = R_o + 1, \tag{H.23}$$

where $\mathbf{D}_{T,m}$ *is defined as in* (4.48).

(b) *If* $(-\Delta\mathbf{h}_R) \notin \mathcal{S}_c(\mathbf{N}_m)$ *and* $\operatorname{rank}(\mathbf{X}_{e,T}) = R + p_T + 1$, *then*

$$\operatorname{rank}(\mathbf{C}_{T,m}) = \operatorname{rank}(\mathbf{C}_m) + 1 \leq \min(R_o + p + 1, S_m) + 1, \tag{H.24}$$

where $\mathbf{C}_{T,m}$ *is defined as in* (4.49).

Proof:
(a) $\operatorname{rank}(\mathbf{D}_m) = \operatorname{rank}(\mathbf{N}_m) = R_o$ and $\operatorname{rank}(\mathbf{X}) = R$ from the definitions of R_o and R, respectively. From $\operatorname{rank}(\mathbf{X}) = R$, it follows that $\operatorname{rank}(\mathbf{D}_{T,m}) = \operatorname{rank}(\mathbf{N}_{T,m})$ by invoking Property B.10a. Since $(-\Delta\mathbf{h}_R) \notin \mathcal{S}_c(\mathbf{N}_m)$ by assumption, it follows from (4.48) that $\operatorname{rank}(\mathbf{N}_{T,m}) = (R_o + 1)$ and, thus, (H.23).

(b) From $\operatorname{rank}(\mathbf{X}_{e,T}) = R + p_T + 1$, it follows that $\operatorname{rank}(\mathbf{X}_e) = R + p + 1$. Thus, by invoking Property B.10a, $\operatorname{rank}(\mathbf{C}_m) = \operatorname{rank}(\mathbf{N}_{e,m})$ and $\operatorname{rank}(\mathbf{C}_{T,m}) = \operatorname{rank}(\mathbf{N}_{e,T,m})$. Since $\operatorname{rank}(\mathbf{N}_{e,m}) \leq \min(R_o + p + 1, S_m)$ from Proposition 4.7 and $(-\Delta\mathbf{h}_R) \notin \mathcal{S}_c(\mathbf{N}_m)$ by assumption, it follows from (4.49) that $\operatorname{rank}(\mathbf{N}_{e,T,m}) \leq \min(R_o + p + 2, S_m + 1) = (\min(R_o + p + 1, S_m) + 1)$ and, thus, (H.24). □

H.3.3 On-line state reconstruction

Proposition 4.14 on page 72:
Let Assumptions A1–7 in Appendix D be verified. Let the concentrations $\mathbf{c}_m$ *of* $S_m \geq R$ *species be measured, and* $\mathbf{N}_m$ *be the* $R \times S_m$ *submatrix of* $\mathbf{N}$ *corresponding to these* S_m *species. Given* $\mathbf{N}$, $\mathbf{C}_{in}$, $\mathbf{q}_{in}$, q_{out}, $\mathbf{c}_0$, *and* V_0, *if* $\mathbf{N}_m^{\mathrm{T}}$ *has a unique left pseudo-inverse, the reaction variants* $\mathbf{x}$ *can be reconstructed without knowledge of reaction kinetics using*

$$\hat{\mathbf{x}} = \mathbf{N}_m^{+\mathrm{T}} \left(\mathsf{h}(\mathbf{z})\, \mathbf{c}_m - \mathbf{C}_{in,m}\mathbf{z} - V_0 \mathbf{c}_{m,0} \right), \tag{H.25}$$

where the subscript m *denotes a quantity corresponding to the* S_m *measured species. From* (4.29), *the concentrations of the remaining* $(S - S_m)$ *species,* $\mathbf{c}_u$, *can be reconstructed using*

$$\hat{\mathbf{c}}_u = \left(\mathbf{N}_u^{\mathrm{T}} \hat{\mathbf{x}} + \mathbf{C}_{in,u}\mathbf{z} + V_0 \mathbf{c}_{u,0} \right) / \mathsf{h}(\mathbf{z}), \tag{H.26}$$

where the subscript u *denotes a quantity corresponding to the* S_u *unmeasured species.*

Let only an estimate of the initial concentrations of the S_u *species be available,* $\hat{\mathbf{c}}_{u,0}$. *Then the estimation error* $\boldsymbol{\varepsilon}_u \equiv \hat{\mathbf{c}}_u - \mathbf{c}_u$, *with* $\hat{\mathbf{c}}_u$ *being the estimated concentrations, is given by*

$$\varepsilon_u^i(k) = V_0 \varepsilon_{u,0}^i / \mathsf{h}(\mathbf{z}(k)), \qquad i = 1, \ldots, S_u, \tag{H.27}$$

where ε_u^i *is the* i*th element of* $\boldsymbol{\varepsilon}_u$. *If the inlets are present at least intermittently, then the estimation error asymptotically converges to zero.*

Proof:
The measured and unmeasured concentrations induce the following partition in (4.3):

$$\mathsf{h}(\mathbf{z})\,\mathbf{c} = \mathsf{h}(\mathbf{z}) \begin{bmatrix} \mathbf{c}_m \\ \mathbf{c}_u \end{bmatrix} = \begin{bmatrix} \mathbf{N}_m^{\mathrm{T}} \\ \mathbf{N}_u^{\mathrm{T}} \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{C}_{in,m} \\ \mathbf{C}_{in,u} \end{bmatrix} \mathbf{z} + V_0 \begin{bmatrix} \mathbf{c}_{m,0} \\ \mathbf{c}_{u,0} \end{bmatrix}. \tag{H.28}$$

From (4.7), the states $\mathbf{z}$ and λ can be reconstructed from $\mathbf{q}_{in}$, q_{out} and V_0. Owing to the existence of $\mathbf{N}_m^{+\mathrm{T}}$ by assumption, $\mathbf{x}$ can be reconstructed from $\mathbf{c}_m$ leading to (H.25). Having reconstructed $\mathbf{x}$, (H.26) follows.

Since $\hat{\mathsf{h}}(\hat{\mathbf{z}}) = \mathsf{h}(\mathbf{z})$ due to the assumption of perfect knowledge of $\mathbf{q}_{in}$, q_{out} and V_0, (H.27) is obtained from (H.26). Since $\mathsf{h}(\mathbf{z}) = V_0 + \mathbf{1}_p^{\mathrm{T}}\mathbf{z}$, the inlets are present at least intermittently for $k \to \infty$, then $\mathsf{h}(\mathbf{z}) \to \infty$ and, therefore, $1/\mathsf{h}(\mathbf{z}) \to 0$. Thus, $\lim_{k\to\infty} \varepsilon_u^i(k) = V_0 \varepsilon_{u,0}^i \lim_{t\to\infty} 1/\mathsf{h}(\mathbf{z}(k)) = 0$ $(i = 1, \ldots, S_u)$. Thus, $\frac{d}{dt}(\mathsf{h}\,\boldsymbol{\varepsilon}_u) = \dot{\mathsf{h}}\,\boldsymbol{\varepsilon}_u + \mathsf{h}\,\dot{\boldsymbol{\varepsilon}}_u = \mathbf{0}_{S_u}$. Since $\dot{\mathsf{h}} = \mathbf{1}_p^{\mathrm{T}}\mathbf{q}_{in}/\mathsf{h}$, the error dynamics are given by:

$$\dot{\boldsymbol{\varepsilon}}_u = -V_0\,\delta\,\boldsymbol{\varepsilon}_{u,0}, \tag{H.29}$$

where $\delta = (\mathbf{1}_p^{\mathrm{T}}\mathbf{q}_{in})/V$ is the so-called *dilution* term. If at least one of the inlet flowrates is nonzero for $k \to \infty$, then $\delta > 0$. Thus, the error dynamics (H.29) are stable, and ε_u converges asymptotically to $\mathbf{0}_{S_u}$. Consequently, the result proposed by Bastin and Dochain (1990) is obtained:

$$\lim_{t\to\infty} \hat{\mathbf{c}}_u = \mathbf{c}_u, \quad \delta > 0. \tag{H.30}$$

□

H.3.4 Target factor analysis applied to reaction data

Proposition 4.16 on page 77:
Let Assumptions A1–7 in Appendix D be verified. Let $\mathbf{C}_{in}$, V_0, $\mathbf{c}(k)$, $\mathbf{q}_{in}(k)$, *and* $q_{out}(k)$ *of reaction systems described by* (2.38) *be known/measured for all* k. *Then, the following properties hold:*

(a) The number of independent reactions can be determined from $\mathbf{D}$ *in* (4.11) *as*

$$R = \operatorname{rank}(\mathbf{D}). \tag{H.31}$$

(b) $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$ *iff* $\varepsilon_p(\mathbf{n}_{tar}, \mathbf{D}) = 0$, *where* $\varepsilon_p(\mathbf{n}_{tar}, \mathbf{D})$ *is the least-squares error of* $\mathbf{n}_{tar}$ *on the row space of* $\mathbf{D}$ *(see C.15).*

(c) Once $\mathbf{N}$ *has been determined (e.g., after having accepted* R *stoichiometric targets),* $\mathbf{X}$ *and the vector of reaction rates* $\mathbf{r}(k)$ *can be reconstructed.*

Proof:
(a) It follows from (4.11) that $\operatorname{rank}(\mathbf{D}) \leq \min(\operatorname{rank}(\mathbf{X}), \operatorname{rank}(\mathbf{N}))$. It is obvious that $\operatorname{rank}(\mathbf{N}) = \operatorname{rank}(\mathbf{X}) = R$ due to the independence of reaction stoichiometries and kinetics. Thus, $\operatorname{rank}(\mathbf{D}) \leq R$. The equality in (H.31) follows from Property B.10.

(b) Since $R < S$ (from Proposition 2.7a), Property b follows immediately from (C.15) and (4.11):

$$\varepsilon_p(\mathbf{n}_{tar}, \mathbf{D}) = \|\mathbf{n}_{tar}^{\mathrm{T}} \left(\mathbf{I}_S - \mathbf{D}^{+}\mathbf{D}\right)\|_2 = \|\mathbf{n}_{tar}^{\mathrm{T}} \left(\mathbf{I}_S - \mathbf{N}^{+}\mathbf{N}\right)\|_2. \tag{H.32}$$

(c) The reconstruction of $\mathbf{X}$ proceeds according to the TFA procedure described in Subsection 4.3.2.1. Since $V(k)$ and $q_{out}(k)$ (and thus $\lambda(k)$) are assumed to be known, $\mathbf{r}(k)$ can be calculated from (4.7) for all k. □

Theorem 4.17 on page 77:
Let Assumptions A1–7 in Appendix D be verified. Let $\mathbf{C}_{in,m}$, V_0, $\mathbf{c}_m(k)$, $\mathbf{q}_{in}(k)$, *and* $q_{out}(k)$ *of reaction systems described by* (2.38) *be known/measured for all* k. *If* $\operatorname{rank}(\mathbf{\Gamma}_u) \geq S - \max(S_m, R)$, *then* $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$ *iff* $\varepsilon_p(\mathbf{n}_{m,tar}, \mathbf{D}_m) = 0$ *and* $\mathbf{n}_{tar}^{\mathrm{T}} \mathbf{\Gamma} = \mathbf{0}_N^{\mathrm{T}}$.

Proof:
Since $\operatorname{rank}(\mathbf{\Gamma}_u) \leq \min(S_u, N)$, the condition $\operatorname{rank}(\mathbf{\Gamma}_u) = S - \max(S_m, R) = \min(S_u, S - R)$ implies $\min(S_u, S-R) \leq \min(S_u, N)$. In turn, this inequality implies either $N \geq S_u$ or $N \geq S - R$. Thus, the condition $\operatorname{rank}(\mathbf{\Gamma}_u) = \min(S_u, S-R)$ implies that $N \geq \min(S_u, S-R)$ columns of $\mathbf{\Gamma}$ are known. Therefore, the right pseudo-inverse of $\mathbf{\Gamma}_u$ is unique.

With $\mathbf{N}$ satisfying (2.18) and $\operatorname{rank}(\mathbf{\Gamma}_u) = \min(S_u, S-R)$ by assumption:

$$\begin{bmatrix} \mathbf{N}_m & \mathbf{N}_u \end{bmatrix} \begin{bmatrix} \mathbf{\Gamma}_m \\ \mathbf{\Gamma}_u \end{bmatrix} = \mathbf{0}_{R\times N} \quad \text{or} \quad \mathbf{N}_u = -\mathbf{N}_m \mathbf{\Gamma}_m \mathbf{\Gamma}_u^{+}. \tag{H.33}$$

$\Rightarrow$: If $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$, then $\mathbf{n}_{tar}^{\mathrm{T}} \mathbf{\Gamma} = \mathbf{0}_N^{\mathrm{T}}$ from (2.18). Also, $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$ implies $\mathbf{n}_{m,tar} \in \mathcal{S}_r(\mathbf{N}_m)$, which leads to $\varepsilon_p(\mathbf{n}_{m,tar}, \mathbf{D}_m) = 0$.

$\Leftarrow$: $\varepsilon_p(\mathbf{n}_{m,tar}, \mathbf{D}_m) = 0$ implies $\mathbf{n}_{m,tar} \in \mathcal{S}_r(\mathbf{N}_m)$, i.e., $\mathbf{n}_{m,tar}^{\mathrm{T}} = \boldsymbol{\alpha}^{\mathrm{T}} \mathbf{N}_m$. Alternatively, $\mathbf{n}_{tar}^{\mathrm{T}} \mathbf{\Gamma} = \mathbf{0}_N$ leads to $\mathbf{n}_{u,tar}^{\mathrm{T}} = -\mathbf{n}_{m,tar}^{\mathrm{T}} \mathbf{\Gamma}_m \mathbf{\Gamma}_u^{+}$. Combining the two relationships and using (H.26), it follows that $\mathbf{n}_{u,tar}^{\mathrm{T}} = \boldsymbol{\alpha}^{\mathrm{T}} \mathbf{N}_u$. Thus, $\mathbf{n}_{tar}^{\mathrm{T}} = [\mathbf{n}_{m,tar}^{\mathrm{T}}, \mathbf{n}_{u,tar}^{\mathrm{T}}] = \boldsymbol{\alpha}^{\mathrm{T}} \mathbf{N}$ and, therefore, $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$. □

Theorem 4.18 on page 78:
Let Assumptions A1–7 in Appendix D be verified. Let $\mathbf{c}(k)$ *of reaction systems described by* (2.38) *be measured for all* k. *If* $\mathbf{X}_e$ *is full rank, and* $\operatorname{rank}(\mathbf{C}_x \mathbf{\Gamma}) = p + 1$, *then the following properties hold:*

(a) The number of independent reactions can be determined from $\mathbf{C}$ *as*

$$R = \operatorname{rank}(\mathbf{C}) - p - 1. \tag{H.34}$$

(b) $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$ *iff* $\varepsilon_p(\mathbf{n}_{tar}, \mathbf{C}) = 0$ *and* $\mathbf{n}_{tar}^{\mathrm{T}} \mathbf{\Gamma} = \mathbf{0}_N^{\mathrm{T}}$.

(c) Given $\mathbf{C}_{in}$, V_0, *and* $q_{out}(k)$ *for all* k, *once* $\mathbf{N}$ *has been determined (e.g. after having accepted* R *stoichiometric targets), the time profiles* $V(k)$, $\mathbf{z}(k)$ *(and* $\mathbf{q}_{in}(k)$*),* $\mathbf{x}(k)$, *and* $\mathbf{r}(k)$ *can be reconstructed.*

Proof:
(a) Since $\operatorname{rank}(\mathbf{C}_x\mathbf{\Gamma}) = (p+1)$ by assumption, it follows from Proposition 2.7c that $\operatorname{rank}(\mathbf{N}_e) = (R+p+1)$. Also, it is assumed that $\operatorname{rank}(\mathbf{X}_e) = (R+p+1)$. This leads to $\operatorname{rank}(\mathbf{C}) = (R+p+1)$ by invoking Property B.10a, and Theorem 4.18a holds.

(b) If $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$, then $\varepsilon_p(\mathbf{n}_{tar}, \mathbf{C}) = 0$ from (4.4) and $\mathbf{n}_{tar}^{\mathrm{T}}\mathbf{\Gamma} = \mathbf{0}_N^{\mathrm{T}}$ from (2.18) (necessary condition). For the sufficient condition, consider $\varepsilon_p(\mathbf{n}_{tar}, \mathbf{C}) = 0$ and $\mathbf{n}_{tar}^{\mathrm{T}}\mathbf{\Gamma} = \mathbf{0}_N^{\mathrm{T}}$. $\varepsilon_p(\mathbf{n}_{tar}, \mathbf{C}) = 0$ is equivalent to $\mathbf{n}_{tar}$ being in the row space of $\mathbf{N}_e$:

$$\mathbf{n}_{tar}^{\mathrm{T}} = \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{N} + \boldsymbol{\beta}^{\mathrm{T}}\mathbf{C}_x. \tag{H.35}$$

With $\mathbf{N}$ satisfying (2.18), it follows that:

$$\mathbf{n}_{tar}^{\mathrm{T}}\mathbf{\Gamma} = \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{N}\mathbf{\Gamma} + \boldsymbol{\beta}^{\mathrm{T}}\mathbf{C}_x\mathbf{\Gamma} = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{C}_x\mathbf{\Gamma} = \mathbf{0}_N^{\mathrm{T}}. \tag{H.36}$$

Since $\mathbf{C}_x\mathbf{\Gamma}$ is full rank, $\boldsymbol{\beta} = \mathbf{0}_{p+1}$ and $\mathbf{n}_{tar}^{\mathrm{T}} = \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{N}$. Thus, $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$, and Theorem 4.18b follows.

(c) Since $\mathbf{C}_{in}$ and $\mathbf{c}_0$ are assumed to be known, $\mathbf{X}_e$ can be computed once $\mathbf{N}$ has been obtained *via* target testing. From (4.4), $\mathbf{X}_e$ can be partitioned as

$$\mathbf{X}_e = \left[\begin{array}{c|c|c|c} \mathcal{H}^{-1}\mathbf{X} & \mathcal{H}^{-1}\mathbf{Z} & V_0\,\mathcal{H}^{-1}\mathbf{1}_K \end{array}\right]. \tag{H.37}$$

It follows that $\mathfrak{h}^{-1}(k) = \lambda(k)/V(k)$ for all k can be obtained from the last column of $\mathbf{X}_e$, since V_0 is assumed to be known. Thus, knowing $q_{out}(k)$, $\lambda(k)$ can be computed using (4.7). With the knowledge of $\lambda(k)$, the volume $V(k)$ can be reconstructed using the known value of $\lambda(k)/V(k)$.

$\mathbf{X}$ and $\mathbf{Z}$ can be directly obtained from (H.37). Then, using (4.7), the reaction rates $\mathbf{r}(k)$ can be reconstructed. Similarly, $\mathbf{q}_{in}(k)$ can be obtained from (4.7). □

Theorem 4.19 on page 79:
Let Assumptions A1–7 in Appendix D be verified. Let $\mathbf{c}_m(k)$ *of reaction systems described by* (2.38) *be measured for all* k. *If* $\mathbf{X}_e$ *is full rank and* $\operatorname{rank}\left(\left[\begin{smallmatrix}\mathbf{C}_{x,m}\mathbf{\Gamma}_m\\ \mathbf{\Gamma}_u\end{smallmatrix}\right]\right) = S - max(S_m - p - 1, R)$, *then* $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$ *iff* $\varepsilon_p(\mathbf{n}_{m,tar}, \mathbf{C}_m) = 0$ *and* $\mathbf{n}_{tar}^{\mathrm{T}}\mathbf{\Gamma} = \mathbf{0}_N^{\mathrm{T}}$.

Proof:
⇒: If $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$, then $\mathbf{n}_{tar}^{\mathrm{T}}\mathbf{\Gamma} = \mathbf{0}_N^{\mathrm{T}}$ from (2.18). Also, $\mathbf{n}_{tar} \in \mathcal{S}_r(\mathbf{N})$ implies $\mathbf{n}_{m,tar} \in \mathcal{S}_r(\mathbf{N}_{e,m})$, which leads to $\varepsilon_p(\mathbf{n}_{m,tar}, \mathbf{C}_m) = 0$.

⇐: $\varepsilon_p(\mathbf{n}_{m,tar}, \mathbf{C}_m) = 0$ implies $\mathbf{n}_{m,tar} \in \mathcal{S}_r(\mathbf{N}_{e,k})$, i.e., $\mathbf{n}_{m,tar}^{\mathrm{T}} = \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{N}_m + \boldsymbol{\beta}^{\mathrm{T}}\mathbf{C}_{x,k}$. Alternatively, $\mathbf{n}_{tar}^{\mathrm{T}}\mathbf{\Gamma} = \mathbf{0}_N^{\mathrm{T}}$ leads to

$$\mathbf{n}_{m,tar}^{\mathrm{T}}\mathbf{\Gamma}_m + \mathbf{n}_{u,tar}^{\mathrm{T}}\mathbf{\Gamma}_u = \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{N}_m\mathbf{\Gamma}_m + \boldsymbol{\beta}^{\mathrm{T}}\mathbf{C}_{x,k}\mathbf{\Gamma}_m + \mathbf{n}_{u,tar}^{\mathrm{T}}\mathbf{\Gamma}_u = \mathbf{0}_N^{\mathrm{T}}. \tag{H.38}$$

Solving for $\boldsymbol{\beta}$ and $\mathbf{n}_{u,tar}$,

$$\left[\begin{array}{cc}\boldsymbol{\beta}^{\mathrm{T}} & \mathbf{n}_{u,tar}^{\mathrm{T}}\end{array}\right] = -\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{N}_m\mathbf{\Gamma}_m\left[\begin{array}{c}\mathbf{C}_{x,k}\mathbf{\Gamma}_m\\ \mathbf{\Gamma}_u\end{array}\right]^{+}. \tag{H.39}$$

Since by definition $\begin{bmatrix} \mathbf{C}_{x,k}\,\mathbf{\Gamma}_m \\ \mathbf{\Gamma}_u \end{bmatrix}$ is full rank, its pseudo-inverse is unique and, thus, $\boldsymbol{\beta}$ and $\mathbf{n}_{u,tar}$ have a unique solution. It can easily be verified by substitution into (H.38) that the unique solution of (H.39) is $\boldsymbol{\beta} = \mathbf{0}_{(p+1)}$ and $\mathbf{n}_{u,tar}^{\mathrm{T}} = \boldsymbol{\alpha}^{\mathrm{T}}\,\mathbf{N}_u$. Finally, since $\mathbf{n}_{u,tar}^{\mathrm{T}} = \boldsymbol{\alpha}^{\mathrm{T}}\,\mathbf{N}_u$, it follows that $\mathbf{n}_{tar}^{\mathrm{T}} = [\mathbf{n}_{m,tar}^{\mathrm{T}},\, \mathbf{n}_{u,tar}^{\mathrm{T}}] = \boldsymbol{\alpha}^{\mathrm{T}}\mathbf{N}$ and, thus, $\mathbf{n}_{tar} \in \mathcal{S}_r\,(\mathbf{N})$. □

H.3.5 Special cases

Proposition 4.23 on page 90:
Let Assumptions A1–7 in Appendix D be verified. Then, the rank of $\mathbf{C}$ *of batch reaction system is determined by:*

$$\operatorname{rank}(\mathbf{C}) = R + 1. \tag{H.40}$$

Proof:
Since $\mathbf{1}_K \notin \mathcal{S}_c\,(\mathbf{X})$, $\operatorname{rank}([\mathbf{X},\, \mathbf{1}_K]) = (R+1)$. By invoking Corollary 4.4a, $\operatorname{rank}(\mathbf{C}) = \operatorname{rank}(\mathbf{N}_e)$. Since according to Proposition 2.7d, $\mathbf{c}_0 \notin \mathcal{S}_r\,(\mathbf{N})$, it follows that $\operatorname{rank}(\mathbf{N}_e) = R+1$ and, thus, the proposition. □

Proposition 4.24 on page 90:
Let Assumptions A1–7 in Appendix D be verified. For single or multiple runs of constant-density batch and semibatch reaction systems, if $\operatorname{rank}(\mathbf{C}) = R + p + 1$*, then* $\mathbf{1}_K \in \mathcal{S}_c\,(\mathbf{X}_e)$*. Proof:*
Single runs: First, the proof for column-mean centering is given. The condition $\operatorname{rank}(\mathbf{C}) = (R + p + 1)$ implies $\operatorname{rank}(\mathbf{N}_e) = \operatorname{rank}(\mathbf{X}_e) = R + p + 1$. For constant-density batch reaction systems, $V(k) = V_0$ and $\lambda(k) = 1$ for all k. Since $\mathbf{V} = V_0\,\mathbf{I}_K$ and $\mathbf{\Lambda} = \mathbf{I}_K$, it can be seen from $\mathbf{X}_e = \mathfrak{H}^{-1}\,[\mathbf{X},\, \mathbf{Z},\, V_0\mathbf{1}_K]$ that the last column of $\mathbf{X}_e$ contains the $\mathbf{1}_K$-vector.

For constant-density semibatch reaction systems, $\dot{V}(k) = \mathbf{1}_p^{\mathrm{T}}\,\mathbf{q}_{in}(k)$ and $\lambda(k) = 1$ for all k. Thus, $\dot{V} = \mathbf{1}_p^{\mathrm{T}}\,\mathbf{q}_{in} = \mathbf{1}_p^{\mathrm{T}}\,\dot{\mathbf{z}}$ or, for K measurements:

$$\mathbf{Z}\,\mathbf{1}_p = (\mathbf{V} - V_0\,\mathbf{I}_K)\;\mathbf{1}_K. \tag{H.41}$$

With (H.41), it can be seen that $\mathcal{S}_c\,(\mathbf{X}_e)$ contains the $\mathbf{1}_K$-vector:

$$\mathbf{X}_e \begin{bmatrix} \mathbf{0}_R \\ \mathbf{1}_p \\ 1 \end{bmatrix} = \mathbf{V}^{-1}\;(\mathbf{V}\,\mathbf{1}_K - V_0\,\mathbf{1}_K + V_0\,\mathbf{1}_K) = \mathbf{1}_K.$$

Multiple runs: The proof is given for column-mean centering. The condition $\operatorname{rank}(\mathbf{C}) = (R+p+1)$ implies $\operatorname{rank}(\mathbf{N}_e) = \operatorname{rank}(\mathbf{X}_e) = R+p+1$. From Property B.6, it follows that it is sufficient to proof that $\mathcal{S}_c\,(\mathbf{X}_e)$ contains a $\mathbf{1}_K$-vector. First, the rank of spectral

data from a constant-density batch reaction system is studied. Note that $\mathbf{V} = \mathbf{\Lambda} = \mathbf{I}_K$. Then, from (H.14) and (H.15), $\mathbf{X}_e$ becomes

$$\mathbf{X}_{e,I3} = \left[\begin{array}{c|ccccc} \mathbf{X}^1 & \mathbf{0}_{K^1} & \mathbf{0}_{K^1} & \cdots & \mathbf{0}_{K^1} & V_0^1\,\mathbf{1}_{K^1} \\ \mathbf{X}^2 & V_0^2\,\mathbf{1}_{K^2} & \mathbf{0}_{K^2} & \cdots & \mathbf{0}_{K^2} & \mathbf{0}_{K^2} \\ \mathbf{X}^3 & \mathbf{0}_{K^3} & V_0^3\,\mathbf{1}_{K^3} & \cdots & \mathbf{0}_{K^3} & \mathbf{0}_{K^3} \\ \hline \mathbf{X}^{p+1} & \mathbf{0}_{K^{p+1}} & \mathbf{0}_{K^{p+1}} & \cdots & V_0^{p+1}\,\mathbf{1}_{K^{p+1}} & \mathbf{0}_{K^{p+1}} \end{array}\right],$$

$$\mathbf{X}_{e,I2} = \left[\begin{array}{c|ccccc} \mathbf{X}^1 & \mathbf{0}_{K^1} & \mathbf{0}_{K^1} & \cdots & \mathbf{0}_{K^1} & V_0^1\,\mathbf{1}_{K^1} \\ \mathbf{X}^2 & V_\Delta^1\,\mathbf{1}_{K^2} & \mathbf{0}_{K^2} & \cdots & \mathbf{0}_{K^2} & V_0^1\,\mathbf{1}_{K^2} \\ \mathbf{X}^3 & V_\Delta^1\,\mathbf{1}_{K^3} & V_\Delta^2\,\mathbf{1}_{K^3} & \cdots & \mathbf{0}_{K^3} & V_0^1\,\mathbf{1}_{K^3} \\ \hline \mathbf{X}^{p+1} & V_\Delta^1\,\mathbf{1}_{K^{p+1}} & V_\Delta^2\,\mathbf{1}_{K^{p+1}} & \cdots & V_\Delta^p\,\mathbf{1}_{K^{p+1}} & V_0^1\,\mathbf{1}_{K^{p+1}} \end{array}\right].$$

Summing up the last $(p+1)$ columns of $\mathbf{X}_{e,I3}$ leads to the $\mathbf{1}_K$ vector and, hence, the $\mathbf{1}_K$ vector lies in the column space of $\mathbf{X}_{e,I3}$. Similarly, the $\mathbf{1}_K$ vector lies in the column space of $\mathbf{X}_{e,I2}$, since its last column is already the $\mathbf{1}_K$ vector.

The rank of spectral data from a constant-density semibatch reaction system is studied now. Note that $\mathbf{\Lambda} = \mathbf{I}_K$. Let $\mathbf{V} := \begin{bmatrix} \mathbf{V}^1 & \cdots & \mathbf{0}_{K^1\times K^1} \\ \vdots & \ddots & \vdots \\ \mathbf{0}_{K^{p+1}\times K^{p+1}} & \cdots & \mathbf{V}^{p+1} \end{bmatrix}$. For simplifying the notation, the dimensions for $\mathbf{1}$ and $\mathbf{0}$ are suppressed in the following two equations. Then,

$$\mathbf{V}\,\mathbf{X}_{e,I3} = \left[\begin{array}{c|ccccc|ccc} \mathbf{X}^1 & \mathbf{Z}^1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & V_0^1\,\mathbf{1} \\ \mathbf{X}^2 & \mathbf{1}\,(\acute{\mathbf{z}}^1)^{\mathrm{T}} & \mathbf{Z}^2 & \cdots & \mathbf{0} & V_\Delta^1\,\mathbf{1} & \cdots & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{X}^{B+1} & \mathbf{1}\,(\acute{\mathbf{z}}^1)^{\mathrm{T}} & \mathbf{1}\,(\acute{\mathbf{z}}^2)^{\mathrm{T}} & \cdots & \mathbf{Z}^{B+1} & \mathbf{0} & \cdots & V_0^{B+1}\,\mathbf{1} & \mathbf{0} \end{array}\right]$$

$$\mathbf{V}\,\mathbf{X}_{e,I2} = \left[\begin{array}{c|ccccc|ccc} \mathbf{X}^1 & \mathbf{Z}^1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & V_0^1\,\mathbf{1} \\ \mathbf{X}^2 & \mathbf{1}\,(\acute{\mathbf{z}}^1)^{\mathrm{T}} & \mathbf{Z}^2 & \cdots & \mathbf{0} & V_\Delta^1\,\mathbf{1} & \cdots & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{X}^{B+1} & \mathbf{1}\,(\acute{\mathbf{z}}^1)^{\mathrm{T}} & \mathbf{1}\,(\acute{\mathbf{z}}^2)^{\mathrm{T}} & \cdots & \mathbf{Z}^{B+1} & V_\Delta^1\,\mathbf{1} & \cdots & V_\Delta^B\,\mathbf{1} & \mathbf{0} \end{array}\right]$$

Let $p = \sum_{j=1}^{B+1} p_j$. In case of excitation by B additional process runs, it is known from (H.3.5) that the $(B+1)$ process runs contain the $\mathbf{1}_K$ vector for, e.g., the first process

run (columns $(R+1)$ to $(R+p_1)$, and $(R+p+B+1)$) and, for the $(B+1)$th process run (columns $(R+\sum_{j=1}^{B} p_j+1)$ to $(R+p)$, and $(R+p+B)$). Hence, summing up the last $(p+B+1)$ columns of $\mathbf{X}_{e,I3}$ leads to the $\mathbf{1}_K$ vector and, hence, the $\mathbf{1}_K$ vector lies in the column space of $\mathbf{X}_{e,I3}$. A similar conclusion can be drawn for $\mathbf{X}_{e,I2}$. □

H.4 Reaction and flow variants/invariants in the factorization of spectral data

H.4.1 Rank analysis of spectral matrices

Proposition 5.5 on page 98:
Let Assumptions A8–11 in Appendix D be verified. Then,

$$\begin{aligned} \operatorname{rank}(\mathbf{A}) &= \operatorname{rank}(\mathbf{C}), \qquad \operatorname{rank}(\mathbf{H}) = \operatorname{rank}(\mathbf{D}) = R \\ \operatorname{rank}\left(\bar{\mathbf{A}}_c\right) &= \operatorname{rank}\left(\bar{\mathbf{C}}_c\right), \qquad \operatorname{rank}\left(\bar{\mathbf{H}}_c\right) = \operatorname{rank}\left(\bar{\mathbf{D}}_c\right) \end{aligned} \tag{H.42}$$

Proof:
Since $\operatorname{rank}(\mathbf{E}) = S$ by assumption, the proposition follows by invoking Property B.11a.□

Corollary 5.7 on page 100:
Let Assumptions A1–7 in Appendix D be verified. If $\operatorname{rank}(\mathbf{X}_e) = (R+p+1)$*, then* $\operatorname{rank}(\mathbf{X}_e^{aug}) = (R+p+1+S_p)$ *is trivially satisfied.*

Proof:
From the definition of $\mathbf{X}_e^{aug}$, it is obvious that $\operatorname{rank}(\mathbf{X}_e^{aug}) = \operatorname{rank}(\mathbf{X}_e) + S_p$. Since by assumption $\operatorname{rank}(\mathbf{X}_e) = (R+p+1)$, the corollary follows. □

H.4.2 Calibration

Proposition 5.9 on page 103:
Let $\operatorname{rank}(\mathbf{E}) = S$*. Then, the concentrations of the* S *species are predicted correctly from a new spectrum,* $\mathbf{a}_n$*, using forward calibration:*

$$\hat{\mathbf{c}}_n^{\mathrm{T}} = \mathbf{a}_n^{\mathrm{T}} \hat{\mathbf{B}}, \quad \textit{where} \quad \hat{\mathbf{B}} = \mathbf{A}^{+}\,\mathbf{C}. \tag{H.43}$$

iff $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$.

Proof:
⇒: The spectral matrix $\mathbf{A}$ can be rank deficient, $A = \operatorname{rank}(\mathbf{A}) < S$. Since $\operatorname{rank}(\mathbf{E}) = S$ by assumption, (5.3) and (4.4) can be formulated in the framework of a bilinear model (C.9)–(C.10) with the spectral data $\mathbf{A}$ being the inputs and the concentration data $\mathbf{C}$ the outputs:

$$\begin{aligned} \mathbf{A} &= \mathbf{T}\,\mathbf{P}^{\mathrm{T}}, \qquad \mathbf{P}^{\mathrm{T}} = \mathbf{Q}^{\mathrm{T}}\,\mathbf{E}, \\ \mathbf{C} &= \mathbf{T}\,\mathbf{Q}^{\mathrm{T}}, \end{aligned} \tag{H.44}$$

where $\mathbf{T}$ is the $K \times A$ scores matrix, and $\mathbf{P}$ and $\mathbf{Q}$ the loading matrices of dimension $L \times A$ and $S \times A$, respectively.

From (H.44), $\mathbf{A}^+ = \mathbf{P}^{+\mathrm{T}}\mathbf{T}^+$. Thus,

$$\hat{\mathbf{B}} = \mathbf{A}^+ \mathbf{C}_k = \mathbf{P}^{+\mathrm{T}}(\mathbf{T}^+ \mathbf{T})\,\mathbf{Q}^{\mathrm{T}} = \mathbf{P}^{+\mathrm{T}}\mathbf{Q}^{\mathrm{T}}. \tag{H.45}$$

In the equation above, $\mathbf{T}^+\mathbf{T} = \mathbf{I}_A$ was used. By hypothesis, the new data obey $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$, i.e., $\mathbf{a}_n^{\mathrm{T}} = \mathbf{t}_n^{\mathrm{T}}\mathbf{A}$. Since, in addition, $\mathrm{rank}(\mathbf{E}) = S$, it follows that $\mathbf{c}_n^{\mathrm{T}} = \mathbf{t}_n^{\mathrm{T}}\mathbf{C}$. The prediction is then given by:

$$\begin{aligned} \hat{\mathbf{c}}_n^{\mathrm{T}} &= \mathbf{a}_n^{\mathrm{T}}\hat{\mathbf{B}} = \mathbf{t}_n^{\mathrm{T}}\mathbf{T}\,(\mathbf{P}^{\mathrm{T}}\,\mathbf{P}^{+\mathrm{T}})\,\mathbf{Q}^{\mathrm{T}} = \mathbf{t}_n^{\mathrm{T}}\mathbf{T}\,\mathbf{Q}^{\mathrm{T}} = \mathbf{t}_n^{\mathrm{T}}\mathbf{C} \\ &= \mathbf{c}_n^{\mathrm{T}}. \end{aligned} \tag{H.46}$$

In the equation above, $\mathbf{P}^{\mathrm{T}}\mathbf{P}^{+\mathrm{T}} = \mathbf{I}_A$ was used.

$\Leftarrow$: Since $\mathrm{rank}(\mathbf{E}) = S$ by assumption, $\mathbf{a}_n \notin \mathcal{S}_r(\mathbf{A})$ implies the corresponding (complete) concentration vector $\mathbf{c}_n \notin \mathcal{S}_r(\mathbf{C})$. The proof proceeds by contradiction. Let $\hat{\mathbf{c}}_n^{\mathrm{T}} \neq \mathbf{c}_n^{\mathrm{T}}$. For $\mathbf{c}_n \in \mathcal{S}_r(\mathbf{C})$, it follows that $\mathbf{c}_n^{\mathrm{T}} = \mathbf{t}_n^{\mathrm{T}}\mathbf{C}$. It can be easily verified that $\mathbf{Q}^{+\mathrm{T}}\mathbf{Q}^{\mathrm{T}}$ and $\mathbf{E}\,\mathbf{P}^{+\mathrm{T}}\mathbf{Q}$ span the same row space. Then,

$$\hat{\mathbf{c}}_n^{\mathrm{T}} = \mathbf{a}_n^{\mathrm{T}}\hat{\mathbf{B}} = \mathbf{c}_n^{\mathrm{T}}\mathbf{E}\,\hat{\mathbf{B}} = \mathbf{t}_n^{\mathrm{T}}\mathbf{C}\,\mathbf{E}\,\mathbf{P}^{+\mathrm{T}}\mathbf{Q} = \mathbf{t}_n^{\mathrm{T}}\mathbf{C} = \mathbf{c}_n^{\mathrm{T}} \tag{H.47}$$

This is in contradiction to the assumption above. Thus, the correct prediction $\hat{\mathbf{c}}_n = \mathbf{c}_n$ implies $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$. □

Proposition 5.11 on page 105:
If $\mathrm{rank}(\mathbf{A}) = S$, *then* $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$.

Proof:
Since $\mathrm{rank}(\mathbf{A}) = S$ by assumption, then from Property B.10 $\mathrm{rank}(\mathbf{C}) = \mathrm{rank}(\mathbf{E}) = S$. Thus, the rows of $\mathbf{C}$ span the entire S-dimensional space and, therefore, $\mathbf{c}_n^{\mathrm{T}} = \mathbf{t}_n^{\mathrm{T}}\mathbf{C}$ and $\mathbf{a}_n^{\mathrm{T}} = \mathbf{t}_n^{\mathrm{T}}\mathbf{A}$. It follows that $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$. □

Proposition 5.12 on page 105:
Let Assumptions A1–12 in Appendix D be verified, $\mathrm{rank}(\mathbf{A}) = (R + p + 1)$, *and* $\mathbf{R}$ *an* $(R_n + p_n + 1) \times (R + p + 1)$ *matrix of full rank. If*

$$\mathbf{N}_{e,n} = \mathbf{R}\,\mathbf{N}_e, \tag{H.48}$$

then $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$.

Proof: Since $\mathrm{rank}(\mathbf{A}) = (R + p + 1)$ and $\mathrm{rank}(\mathbf{E}) = S$ by assumption, $\mathrm{rank}(\mathbf{C}) = (R + p + 1)$ follows from Proposition 5.5. Thus, from Corollary 4.4b, $\mathrm{rank}(\mathbf{X}_e) = \mathrm{rank}(\mathbf{N}_e) = R + p + 1$. From $\mathrm{rank}(\mathbf{X}_e) = (R + p + 1)$, it follows that $\mathbf{X}_e^+\,\mathbf{X}_e = \mathbf{I}_{R+p+1}$. Then, with $\mathbf{c}_n^{\mathrm{T}} = \mathbf{x}_{e,n}^{\mathrm{T}}\mathbf{N}_{e,n}$, $\mathbf{N}_{e,n} = \mathbf{R}\,\mathbf{N}_e$, and $\mathbf{A} = \mathbf{X}_e\,\mathbf{N}_e\,\mathbf{E}$, it follows that

$$\mathbf{a}_n^{\mathrm{T}} = \mathbf{x}_{e,n}^{\mathrm{T}}\,\mathbf{N}_{e,n}\,\mathbf{E} = \mathbf{x}_{e,n}^{\mathrm{T}}\,\mathbf{R}\,\mathbf{N}_e\,\mathbf{E} = \mathbf{x}_{e,n}^{\mathrm{T}}\,\mathbf{R}\,\mathbf{X}_e^+\,(\mathbf{X}_e\,\mathbf{N}_e\,\mathbf{E}) = \mathbf{t}_n^{\mathrm{T}}\,\mathbf{A}, \tag{H.49}$$

where $\mathbf{t}_n^{\mathrm{T}} = \mathbf{x}_{e,n}^{\mathrm{T}}\,\mathbf{R}\,\mathbf{X}_e^+$. Thus, $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$. □

Corollary 5.13 on page 105:
Let Assumptions A1–12 in Appendix D be verified, $\mathrm{rank}(\mathbf{A}) = (R+p+1)$, $\mathbf{N}_n = \mathbf{R}^\dagger\,\mathbf{N}$,

and $\mathbf{R}^{\dagger}$ *and* $\mathbf{R}^{\ddagger}$ *matrices of full rank of dimension* $R_n \times R$ *and* $(p_n+1)\times(p+1)$*, respectively. If*

$$\mathbf{C}_{x,n} = \mathbf{R}^{\ddagger}\,\mathbf{C}_x, \tag{H.50}$$

then $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$.

Proof: If $\mathbf{N}_n = \mathbf{R}^{\dagger}\,\mathbf{N}$ and $\mathbf{C}_{x,n} = \mathbf{R}^{\ddagger}\,\mathbf{C}_x$, then

$$\mathbf{N}_{e,n} = \left[\begin{array}{c}\mathbf{N}_n\\\hline \mathbf{C}_{x,n}\end{array}\right] = \left[\begin{array}{c}\mathbf{R}^{\dagger}\,\mathbf{N}\\\hline \mathbf{R}^{\ddagger}\,\mathbf{C}_x\end{array}\right] = \left[\begin{array}{c|c}\mathbf{R}^{\dagger} & \mathbf{0}_{R_n\times p+1}\\\hline \mathbf{0}_{p_n+1\times R} & \mathbf{R}^{\ddagger}\end{array}\right]\left[\begin{array}{c}\mathbf{N}\\\hline \mathbf{C}_x\end{array}\right] = \mathbf{R}\,\mathbf{N}_e. \tag{H.51}$$

Since $\mathbf{N}_{e,n} = \mathbf{R}\,\mathbf{N}_e$, from Proposition 5.12, it follows that $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$. □

Corollary 5.14 on page 106:
Let Assumptions A1–12 in Appendix D be verified, $\mathbf{N}_n = \mathbf{R}^{\dagger}\,\mathbf{N}$*, and* $\operatorname{rank}(\mathbf{A}) = R + p + 1$*. If* $\operatorname{rank}(\mathbf{C}_x) = S_x$*, then* $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A})$.

Proof: If $\operatorname{rank}(\mathbf{C}_x) = S_x$, by the definition of S_x, $\mathbf{C}_{x,n} \in \mathcal{S}_r(\mathbf{C}_x)$. The result follows from Corollary 5.13. □

H.4.3 Factor analysis

Proposition 5.21 on page 121:
Let Assumptions A1–12 in Appendix D be verified and $\mathbf{R}$ *an* $(R+p+1)\times K_r$ *matrix. If*

$$\mathbf{N}_e = \mathbf{R}\,\mathbf{C}_r, \tag{H.52}$$

then $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A}_r)$.

Proof:
Since $\mathbf{N}_e = \mathbf{R}\,\mathbf{C}_r$ by assumption, it follows that $\mathbf{C} = \mathbf{X}_e\mathbf{N}_e = \mathbf{R}^{\dagger}\mathbf{C}_r$, where $\mathbf{R}^{\dagger}$ is a $K \times K_r$ matrix with $\mathbf{R}^{\dagger} \equiv \mathbf{X}_e\mathbf{R}$. Thus, $\mathbf{A} = \mathbf{R}^{\dagger}\mathbf{A}_r$ and, thus, $\mathbf{a}_n \in \mathcal{S}_r(\mathbf{A}_r)$. □

References

Aguilar, R., Gonzalez, J., Alvarez-Ramirez, J. and Barron, M. A. (1997), 'Temperature regulation of a class of continuous chemical reactor based on a nonlinear Luenberger-like observer', *Journal of Chemical Technology and Biotechnology* **70**(3), 209–216.

Amrhein, M. (1998), Titration in analytical chemistry, Technical Report 1998.05, Institut d'Automatique, EPFL.

Amrhein, M., Husmann, E., Srinivasan, B., Bonvin, D. and Schumacher, M. M. (1997), Control and online optimization of a laboratory-scale semi-batch reactor using near-infrared spectroscopic measurements, *in* '6th International Conference on Automation, Robotics and Artificial Intelligence Applied to Analytical Chemistry and Laboratory Medicine (ICAR)', Montreux, Switzerland.

Amrhein, M., Srinivasan, B., Bonvin, D. and Schumacher, M. M. (1996), 'On the rank deficiency and rank augmentation of the spectral measurement matrix', *Chemometrics Intellig. Lab. Sys.* **33**(1), 17–33. first presented on the first Internet Conference on Chemometrics in October 1994.

Amrhein, M., Srinivasan, B., Bonvin, D. and Schumacher, M. M. (1998), 'Calibration of spectral reaction data: Importance of rank augmentation and data pre-treatment', accepted by *Chemometrics Intellig. Lab. Sys.*

Bak, J. and Larsen, A. (1995), 'Quantitative gas analysis with FT-IR: a method for co calibration using partial least-squares with linearized data', *Applied Spectroscopy* **49**(4), 437–443.

Bastin, G. and Dochain, D. (1990), *On-line Estimation and Adaptive Control of Bioreactors*, Elsevier, Amsterdam.

Bastin, G. and Lévine, J. (1993), 'On state accessibility in reaction systems', *IEEE Trans. Automatic Contr.* **38**(5), 733–742.

Bastin, G., Chen, L. and Chotteau, V. (1992), Can we identify biotechnological processes?, *in* 'ICCAF-5/IFAC-BIO-2 Symposium on Modelling and Control of Biological Processes', Colorado, USA, pp. 83–88.

Birth, G. S. and Hecht, H. G. (1987), The physics of near-infrared reflectance, *in* P. C. Williams and K. Norris, eds, 'Near-infrared Technology in Agricultural and Food Industries', Am. Assoc. Cereal Chem., pp. 1–15.

Bonvin, D. and Rippin, D. W. T. (1990), 'Target factor analysis for the identification of stoichiometric models', *Chem. Eng. Sci.* **49**, 3417–3426.

Box, G. E. P., Hunter, W. G. and Hunter, J. S. (1978), *Statistics for Experimenters. An Introduction to Design, Data Analysis and Model Building*, John Wiley & Sons, New York.

Brereton, R. G. and Elbergali, A. K. (1994), 'Use of double windowing, variable selection, variable ranking and resolvability indices in window factor analysis', *J. Chemometrics* **8**, 423–437.

Brown, P. J. (1992), 'Wavelength selection in multicomponent near-infrared calibration', *J. Chemometrics* **6**, 151–161.

Brown, P. J. (1993), *Measurement, Regression, and Calibration*, Clarendon Press, Oxford.

Bugnon, P., Chottard, J.-. C., Jestin, J.-. L., Jung, B., Laurency, G., Maeder, M., Merbach, A. E. and Zuberbühler, A. D. (1994), 'Second-order globalisation for the determination of activation parameters in kinetics', *Analytica Chimica Acta* **298**, 193–201.

Burns, D. A. and Ciurzak, E. W. (1992), *Handbook of Near-Infrared Analysis*, Vol. 13 of *Practical Spectroscopy*, Marcel Dekker, Inc., New York.

Cannizzaro, C. (1998), Reactir 1000 evaluation report, Technical report, Institute of Chemical Engineering, EPFL, Lausanne.

Cattell, R. B. (1966), 'The Scree test for the number of factors', *Multivar. Behav. Res.* **1**, 245–276.

Chang, K., Brown, C. W. and Force, R. K. (1993), 'Laser-induced fluorescence multicomponent analysis in full spectra and their fourier transform by principal component regression', *Applied Spectroscopy* **47**(11), 1801–1807.

Chen, C.-T. (1984), *Linear System Theory and Design*, Holt Rinehart and Winston, New York.

Chen, L., Zhang, J., Morris, A. J., Montague, G. A., Kent, C. A. and Norton, J. P. (1995), Combining neural networks with physical knowledge in modelling and state estimation of bioprocesses, *in* the proceedings of the European Control Conference (ECC'95), Rome, Italy.

Colthup, N. B., Daly, L. H. and Wiberley, S. E. (1990), *Introduction to Infrared and Raman Spectroscopy*, 3rd edn, Academic Press Ltd., Londom.

Crocombe, R. A., Olson, M. L. and Hill, S. L. (1984), Quantitative fourier transform infrared methods for real complex samples, *in* G. L. McClure, ed., 'Computerized Quantitative Infrared Analysis', ASTM, pp. 95–130.

Crouch, S. R. (1993), 'Trends in kinetic methods of analysis', *Analytica Chimica Acta* **283**, 453–470.

de Jong, S. (1991), 'Chemometrical applications in an industrial food research laboratory', *Mikrochimica Acta* **II**, 93–101.

de Jong, S. (1993), 'Simpls: an alternative approach to partial least squares regression', *Chemometrics Intellig. Lab. Sys.* **18**, 251–263.

de Vallière, P. R. L. (1989), State and Parameter Estimation in Batch Reactors for the Purpose of Inferring On-line the Rate of Heat Production, Doctoral thesis No. 8847, ETH Zürich, Switzerland.

Denham, M. C. (1997), 'Prediction intervals in partial least squares', *J. Chemometrics* **11**(1), 39–52.

DiFoggio, R. (1995), 'Examination of some misconceptions about near-infrared analysis', *Applied Spectroscopy* **49**(1), 67–75.

Dochain, D. and Chen, L. (1992), 'Local observability and controllability of stirred tank reactors', *J. Process Control* **2**(3), 139–144.

Duboc, P. (1997), Transient Growth of *Saccharomyces cerevisiae*, a Quantitative Approach, Doctoral thesis No. 1648, EPF Lausanne, Switzerland.

Dyson, R. M., Kaderli, S., Lawrance, G. A., Maeder, M. and Zuberbühler, A. D. (1997), 'Second-order global analysis: The evaluation of series of spectrophotometric titrations for improved determination of equilibrium constants', *Analytica Chimica Acta* **353**, 381–393.

Efron, B. and Tibshirani, R. (1986), 'Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy', *Statistical Science* **1**(1), 54–77.

Faber, K. and Kowalski, B. R. (1997), 'Propagation of measurement errors for the validation of predictions obtained by principal component regression and partial least squares', *J. Chemometrics* **11**(3), 181–238.

Faber, N. M., Buydens, L. M. C. and Kateman, G. (1994), 'Aspects of pseudorank estimation methods based on an estimate of the size of the measurement error', *Analytica Chimica Acta* **296**, 1–20.

Fjeld, M., Asbjørnsen, O. A. and Åström, K. J. (1974), 'Reaction invariants and their importance in the analysis of eigenvectors, state observability and controllability of the continuous stirred tank reactor', *Chem. Eng. Sci.* **29**, 1917–1926.

Frans, S. D., McConnell, M. L. and Harris, J. M. (1985), 'Multiwavelength detection and reiterative least squares resolution of overlapped liquid chromatographic peaks', *Analytical Chemistry* **57**, 1552–1559.

Gampp, H., Maeder, M. and Zuberbühler, A. D. (1988), 'Computerized data acquisition and data reduction in spectrophotometric analysis. Part 2: Numerical analysis with and without an underlying chemical model', *Trends in Analytical Chemistry* **7**(4), 147–150.

Gampp, H., Maeder, M., Meyer, C. J. and Zuberbühler, A. D. (1985), 'Calculation of equilibrium constants from multiwavelength spectroscopic data — III. Model-free analysis of spectrophotometric and ESR titrations', *Talanta* **32**, 1133–1139.

Garland, M., Visser, E., Terwiesch, P. and Rippin, D. W. T. (1997), 'On the number of observable species, observable reactions and observable fluxes in chemometric studies and the role of multichannel integration', *Analytica Chimic Acta* **351**, 337–358.

Geladi, P. and Wold, S. (1987), 'Local principal component models, rank maps and contextuality for curve resolution and mulit-way calibration inference', *Chemometrics Intellig. Lab. Sys.* **2**, 273–281.

Gemperline, P. J. (1989), 'Mixture analysis using factor analysis. I. Calibration and quantitation', *J. Chemometrics* **3**, 549–568.

Gemperline, P. J., Long, J. R. and Gregoriou, V. G. (1991), 'Nonlinear multivariate calibration using principal components regression and artificial neural networks', *Analytical Chemistry* **63**, 2313–2323.

Golub, G. H. and Van Loan, C. F. (1983), *Matrix Computations*, North Oxford Academic, Oxford.

Griffiths, P. R. and de Haseth, J. A. (1986), *Fourier Transform Infrared Spectroscopy*, Vol. 83 of *Chemical Analysis: A series of Monographs on Analytical Chemistry and its Applications*, John Wiley & Sons, New York.

Haaland, D. M. (1992), Multivariate calibration methods applied to the quantitative analysis of infrared spectra, *in* P. C. Jurs, ed., 'Computer-Enhanced Analytical Spectroscopy', Vol. 3 of *Modern Analytical Chemistry*, Plenum Press, pp. 1–30.

Haaland, D. M. and Thomas, E. V. (1988), 'Partial least-squares methods for spectral analyses. I. Relation to other quantitative calibration methods and the extraction of qualitative information', *Analytical Chemistry* **60**, 1193–1202.

Hamer, J. W. (1989), 'Stoichiometric interpretation of multireaction data: an application to fed-batch fermentation data', *Chem. Eng. Sci.* **44**(10), 2363–2374.

Hamilton, J. C. and Gemperline, P. J. (1990), 'Mixture analysis using factor analysis. II: Self-modeling curve resolution', *J. Chemometrics* **4**, 1–13.

Harmon, J. L., Duboc, P. and Bonvin, D. (1995), 'Factor analytical modeling of biochemical data', *Comp. Chem. Eng.* **19**(12), 1287–1300.

Hill, C. G. (1977), *An Introduction to Chemical Engineering Kinetics & Reactor Design*, John Wiley & Sons.

Himmelblau, D. M., Jones, C. R. and Bischoff, K. B. (1967), 'Determination of rate constants for complex kinetics models', *Ind. Eng. Chem. Fund.* **6**(4), 539–543.

Hopke, P. H. (1989), 'Target transformation factor analysis', *Chemometrics Intellig. Lab. Sys.* **6**, 7–189.

Huber, P. J. (1981), *Robust Statistics*, John Wiley & Sons, New York.

Isidori, A. (1989), *Nonlinear Control Systems*, Springer-Verlag, Berlin.

Izquierdo-Ridorsa, A., Saurina, J., Hernández-Cassou, S. and Tauler, R. (1997), 'Second-order multivariate curve resolution applied to rank-deficient data obtained from acid–base spectrophotometric titrations of mixtures of nucleic bases', *Chemometrics Intellig. Lab. Sys.* **38**, 183–196.

Jackson, J. E. (1991), *A User's Guide to Principal Components*, John Wiley & Sons.

Karjalainen, E. J. (1989), 'The spectrum reconstruction problem: Use of alternating regression for unexpected spectral components in two-dimensional spectroscopies', *Chemometrics Intellig. Lab. Sys.* **7**, 31–38. P C 3554.

Keener, J. P. (1988), *Principles of Applied Mathematics: Transformation and Approximation*, Addison–Wesley, New York.

Keller, H. R. and Massart, D. L. (1991), 'Peak purity control in liquid chromatography with photodiode-array detection by a fixed size moving window evolving factor analysis', *Analytica Chimica Acta* **246**(5), 379–390.

Keller, H. R., Massart, D. L., Liang, Y. Z. and Kvalheim, O. M. (1992), 'Evolving factor analysis in the presence of heteroscedastic noise', *Analytica Chimica Acta* **263**, 29–36.

Knorr, F. J., Thorsheim, H. R. and Harris, J. M. (1981), 'Multichannel detection and numerical resolution of overlapping chromatographic peaks', *Analytical Chemistry* **53**(6), 821–825.

Koch, E. (1977), *Non-isothermal Reaction Analysis*, Academic Press.

Kormos, D. W. and Waugh, J. S. (1983), 'Abstract factor analysis of solid-state nuclear magnetic resonance spectra', *Analytical Chemistry* **55**, 633–638.

Kortum, G. (1969), *Reflectance Spectroscopy: Principles, Methods, Applications*, Springer Verlag, New York.

Krutschkoff, R. G. (1967), 'Classical and inverse regression methods of calibration', *Technometrics* **9**, 425–439.

Kvalheim, O. M. and Liang, Y.-Z. (1992), 'Heuristic evolving latent projections: Resolving two-way multicomponent data. 1. selectivity, latent-projective graph, datascope, local rank, and unique resolution', *Analytical Chemistry* **64**, 936–946.

Lamberti, C., Bordiga, S., Cerrato, G., Morterra, C., Scarano, D., Spoto, G. and Zecchina, A. (1993), 'Minuit subroutine for spectra deconvolution', *Computer Physics Communication* **74**, 119–141.

Lay, D. C. (1994), *Linear Algebra and Its Applications*, Addison–Wesley, New York.

Lucasius, C. B., Beckers, M. L. M. and Kateman, G. (1994), 'Genetic algorithms in wavelength selection: a comparitive study', *Analytica Chimica Acta* **286**, 135–153.

Maeder, M. (1987), 'Evolving factor analysis for the resolution of overlapping chromatographic peaks', *Analytical chemistry* **59**(1), 527–539.

Maeder, M. and Zuberbühler, A. D. (1990), 'Nonlinear least-squares fitting of multivariate absorption data', *Analytical Chemistry* **62**, 2220–2224.

Magnus, J. R., , and Neudecker, H. (1988), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley & Sons.

Malinowski, E. R. (1978), 'Theory of error for target factor analysis with applications to mass spectrometry and nuclear magnetic resonance spectrometry', *Analytica Chimica Acta* **103**, 339–354.

Malinowski, E. R. (1988), 'Statistical F-tests for abstract factor analysis and target testing', *J. Chemometrics* **3**, 49–60.

Malinowski, E. R. (1991), *Factor Analysis in Chemistry*, John Wiley & Sons.

Malinowski, E. R. (1992), 'Window factor analysis: theoretical derivation and application to flow injection analysis data', *J. Chemometrics* **6**(1), 29–40.

Malinowski, E. R. and McCue, M. (1977), 'Qualitative and quantitative determination of suspected components in mixtures by target transformation factor analysis of their mass spectra', *Analytical Chemistry* **49**(2), 284–287.

Manne, R. (1995), 'On the resolution problem in hyphenated chromatography', *Chemometrics Intellig. Lab. Sys.* **27**, 89–94.

Martens, H. and Naes, T. (1989), *Multivariate Calibration*, John Wiley & Sons, New York.

McMullen, D. W., Jaskunas, S. R. and Tinoco, I. (1967), 'Application of matrix rank analysis of the optical rotatory dispersion of tmv rna', *Biopolymers* **5**, 589–613.

Mottola, H. A. (1993), 'Kinetic determinations of reactants utilizing uncatalyzed reactions', *Analytica Chimica Acta* **280**, 279–287.

Nijmeijer, H. and van der Schaft, A. J. (1990), *Nonlinear Dynamical Control Systems*, Springer-Verlag, New York.

Nørgaard, L. and Ridder, C. (1994), 'Rank annihilation factor analysis applied to flow injection analysis with photodiode-array detection', *Chemometrics Intellig. Lab. Sys.* **23**, 107–114.

Oehme, F. and Richter, W. (1987), *Instrumental Titration Techniques: Principles and Applications of the Volumetric Analysis with Emphasis on Aspects of the Measuring and Instrumental Procedures*, ABC of Measuring and Analytical Procedures, Dr. Alfred Heuthig Verlag, Heidelberg, Germany.

Osten, D. W. (1988), 'Slection of optimal regression models via cross-validation', *J. Chemometrics* **2**, 39–48.

Paatero, P. and Tapper, U. (1993), 'Analysis of different modes of factor analysis as least squares fit problems', *Chemometrics Intellig. Lab. Sys.* **18**, 183–194.

Perkampus, H.-. H. and Kaufmann, R. (1991), *Kinetische Analyse mit Hilfe der UV/VIS-Spektrometrie*, Datenverarbeitung in den Naturwissenschaften, VCH, Weinheim, Germany.

Perry, R. H. and Green, D. W. (1984), *Perry's Chemical Engineers' Handbook*, Chemical Engineering Series, 6th edn, McGraw–Hill, Singapore.

Perspektive (1997), Revolutionary development in process engineering, Semesterly publication, Systag Ltd., Rüschlikon, Switzerland.

Phatak, A. and de Jong, S. (1997), 'The geometry of partial least squares', *J. Chemometrics* **11**(4), 311–338.

Phatak, A., Reilly, P. M. and Penlidis, A. (1992), 'The geometry of 2-block partial least squares regression', *Communications in statistics* **21**(6), 1517–1553.

Phatak, A., Reilly, P. M. and Penlidis, A. (1993), 'An approach to interval estimation in partial least squares regression', *Analytica chimica acta* **277**(2), 495–501.

Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W. T. (1994), *Numerical Recipes in C. The Art of Scientific Computing*, 2nd edn, Cambridge University Press.

Pretsch, E., Clerc, T. and Seibl, J. (1981), *Structural Analysis of Organic Compounds by Combined Application of Spectroscopic Methods*, Vol. 1 of *Studies in Analytical Chemistry*, Elsevier, Amsterdam.

Prinz, O. (1992), Chemometric Methods for Investigating Chemical Reaction Systems, Doctoral Thesis No. 9708, ETH Zürich, Switzerland.

ReactIR (1994), Trademark of Applied Systems Inc., Annapolis, MD, USA.

Rothfuss, R., Rudolph, J. and Zeitz, M. (1996), 'Flatness based control of a nonlinear chemical reactor model', *Automatica* **32**(10), 1433–1439.

Sánchez, E. and Kowalski, B. R. (1988*a*), 'Tensorial calibration: I. First order calibration', *J. Chemometrics* **2**, 247–263.

Sánchez, E. and Kowalski, B. R. (1988*b*), 'Tensorial calibration: II. Second order calibration', *J. Chemometrics* **2**, 265–283.

Saurina, J., Hernández-Cassou, S., Tauler, R. and Izquierdo-Ridorsa, A. (1998), 'Multivariate resolution of rank-deficient spectrophotometric data from first-order kinetic decomposition reactions', *J. Chemometrics* **12**, 183–203.

Sheiner, L. B. (1984), 'Analysis of pharmacokinetic data using parametric models — I: Regression models', *Journal of Pharmacokinetics and Biopharmaceutics* **12**(1), 93–117.

Sheiner, L. B. (1985), 'Analysis of pharmacokinetic data using parametric models — II: Point estimates of an individual's parameters', *Journal of Pharmacokinetics and Biopharmaceutics* **13**, 515–541.

Sheiner, L. B. (1986), 'Analysis of pharmacokinetic data using parametric models — III: Hypothesis tests and confidence intervals', *Journal of Pharmacokinetics and Biopharmaceutics* **14**(5), 539–555.

Shrager, R. I. and Hendler, R. W. (1982), 'Titration of individual components in a mixture with resolution of difference spectra, pKs, and redox transitions', *Analytical Chemistry* **54**(7), 1147–1152.

American Chemical Society, American Institute of Chemical Engineers, Chemical Manufacturers Association, Council for Chemical Research, Synthetic Organic Chemical Manufacturers Association. (1996), Technology Vision 2020: Report of the U.S. Chemical Industry. ftp://acs.org.

Soroush, M. (1997), 'Nonlinear state-observer design with application to reactors', *Chem. Eng. Sci.* **52**(3), 387–404.

Srinivasan, B., Amrhein, M. and Bonvin, D. (1998), 'Reaction and flow variants/ invariants in chemical reaction systems with inlet and outlet streams', *AIChE J.* **44**(8), 1858–1867.

Steinhauser, U. (1996), Einflussreiche Beobachtungen in der Explorativen Faktorenanalyse: Identifikation und Einflussbeschränkung, Doctoral Thesis, University of Freiburg, Switzerland.

Stone, M. and Brooks, R. J. (1990), 'Continuum regression: cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal components regression', *Journal of the Royal Statistical Society B* **52**(2), 237–269.

Sylvestre, E. A. and Maggio, M. S. (1974), 'Curve resolution using a postulated reaction', *Technometrics* **16**(3), 353–368.

Tauler, R., Smilde, A. and Kowalski, B. (1995), 'Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution', *J. Chemometrics* **9**, 31–58.

Titration (1954), *Titration Methods: Acid–base, Precipitation, and Complex-formation Reactions*, Vol. 2 of *Volumetric Analysis*, Interscience Publishers, New York.

Titration (1957), *Titration Methods: Oxidation-Reduction Reactions*, Vol. 3 of *Volumetric Analysis*, Interscience Publishers, New York.

Toft, J. (1995), 'Evolutionary rank analysis applied to multidetectional chromatographic structures', *Chemometrics Intellig. Lab. Sys.* **29**, 189–212.

Valluri, S., Soroush, M. and Nikravesh, M. (1998), 'Shortest-prediction-horizon nonlinear model-predictive control', *Chem. Eng. Sci.* **53**(2), 273–292.

Van Stokkum, I. H. M. (1997), 'Parameter precision in global analysis of time-resolved spectra', *IEEE transactions on Instrumentation and Measurement* **46**(4), 764–768.

Vandeginste, B. G. M., Leyten, F., Gerritsen, M., Noor, J. W. and Kateman, G. (1987), 'Evaluation of curve resolution and iterative target transformation factor analysis in quantitative analysis by liquid chromatography', *J. Chemometrics* **1**, 57–71.

Vetterli, M. and Kovačević, J. (1995), *Wavelets and Subband Coding*, Signal Processing, Prentice Hall, Englewood Cliffs, NJ.

Vidyasagar, M. (1993), *Nonlinear Systems Analysis*, Prentice Hall, Englewood Cliffs, NJ.

Vogels, J. T. W. E., Tas, A. C., van den Berg, F. and van der Greef, J. (1993), 'A new method for classification of wines based on proton and carbon-13 NMR spectroscopy in combination with pattern recognition techniques', *Chemometrics Intellig. Lab. Sys.* **21**, 249–258.

Walczak, B. and Massart, D. L. (1995), 'Robust principal components regression as a detection tool for outliers', *Chemometrics Intellig. Lab. Sys.* **27**, 41–54.

Waller, K. V. and Mäkilä, P. M. (1981), 'Chemical reaction invariants and variants and their use in reactor modeling, simulation, and control', *Ind. Eng. Chem. Process Des. Dev.* **20**(1), 1–11.

Weiss, S. (1991), Faktorenanalyse in der Kinetik, PhD thesis, University of Tübingen, Germany.

Windig, W. (1990), 'Interactive self-modeling multivariate analysis', *Chemometrics Intellig. Lab. Sys.* **9**, 7–30.

Wise, B. M. (1991), Adapting Multivariate Analysis for Monitoring and Modeling of Dynamic Systems, PhD thesis, University of Washington, USA.

Wold, H. (1966), Estimation of principal components and related models by iterative least squares, *in* P. R. Krishnaiah, ed., 'Multivariate Analysis', Academic Press, pp. 391–420.

Wold, H. and Lyttkens, E. (1970), Nonlinear iterative partial least squares (nipals) estimation procedures, *in* 'Proceedings of the 37th Session of the International Statistical Institute', The Hague, The Netherlands, pp. 29–47.