Improved Economic Performance
of
Municipal Solid Waste Combustion Plants
by
Model Based Combustion Control

Improved Economic Performance
of
Municipal Solid Waste Combustion Plants
by
Model Based Combustion Control

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op

maandag 18 maart 2013 om 15:00 uur

door

Martijn LESKENS

werktuigkundig ingenieur
geboren te ’s-Gravenhage

Dit proefschrift is goedgekeurd door de promotoren:
Prof. dr. ir. P.M.J. Van den Hof
Prof. ir. O.H. Bosgra

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. P.M.J. Van den Hof,	Technische Universiteit Delft, promotor
Prof. ir. O.H. Bosgra,	Technische Universiteit Delft, promotor
Prof. dr. D. Dochain,	Université Catholique de Louvain
Prof. dr. ir. G. Brem,	Universiteit Twente
Prof. dr. S. Weiland,	Technische Universiteit Eindhoven
Prof. dr. ir. A.H.M. Verkooijen,	Technische Universiteit Delft
Dr. ir. L.B.M. van Kessel,	DNV KEMA
Prof. dr. ir. J. Hellendoorn,	Technische Universiteit Delft, reservelid

Dr. ir. L.B.M. van Kessel heeft als begeleider in belangrijke mate aan de totstandkoming van het proefschrift bijgedragen.

Het werk in dit proefschrift is financieel deels ondersteund door TNO.

MIX
Paper from responsible sources
FSC® C014810
FSC www.fsc.org

ISBN: 978-94-6191-664-8

Keywords: model based control, waste combustion

Printed in the Netherlands.

Acknowledgements

This thesis is the result of fourteen years of PhD research at both the Delft University of Technology and TNO. This long period has required a lot of patience and support from people in my surroundings, in particular from my supervisors and closest family. I particularly take this opportunity here to thank these people.

First and foremost of all I would like to thank Paul Van den Hof and Okko Bosgra from the Delft Center of Systems and Control (DCSC) and Robert van Kessel from DNV KEMA for their motivating supervisory work, support, (incredible) patience and many pleasant and fruitful discussions on both issues related to model based control of waste combustion plants and other subjects not related to the PhD research. A special thanks goes to Paul for the supportive and leading role in particularly the final stages of the research.

I also would like to thank TNO, in particular my former colleagues Robert and Gerrit from the (also) former department of Thermal Conversion at the TNO Environment, Energy and Process Innovation institute in Apeldoorn, for initiating and financially supporting this PhD research and providing me with the opportunity to do it.

Furthermore, I would like to thank many other people I have met during my period as a PhD researcher, in particular at the Delft University of Technology and TNO, for providing a pleasant atmosphere to work in. As I have worked for several employers and at multiple departments during the fourteen years of my PhD research, the list of these people is too long to mention here. As an exception, I would like to mention here that I will certainly miss the conversations at DCSC with Arjan about football and those with Xavier about anything but football.

Finally, I would like to thank my family, in particular my wife Angela, for the (again incredible) patience and support they have shown in especially the final years of the PhD project.

Martijn Leskens
Rosmalen, February 2013

viii

Contents

Chapter 1

Introduction

The main objective of this thesis is to explore the opportunities of model based combustion control in improving the economic performance of municipal solid waste combustion (MSWC) plants. In this introductory chapter, this research objective is motivated and formally stated. Also, a specific solution strategy is proposed to tackle this objective and the contents and main contributions of this thesis are outlined.

1.1 State-of-the-art and challenges in the operation of MSWC plants

1.1.1 The aims of municipal solid waste combustion

The combustion of municipal solid waste (MSW) is used for its inertisation, reduction of its volume and the conversion of its energy content into heat and/or electricity[1]. By inertising the waste the health risks associated with waste disposal are reduced. Inertisation was, in fact, the motivation for building the very first MSWC plants (see *e.g.* [111]), with the earliest one constructed in 1870 in Paddington, a borough of London, [32, 111]. Volume reduction is an important motivation for applying MSWC in places where landfill, the main alternative to combustion, is not a feasible option, such as islands and highly densely populated areas, in particular cities. The heat and/or electrical power resulting from the conversion of the chemical energy content of the waste is fed to the surroundings *c.q.* delivered to a neighbouring city, public buildings (*e.g.* a hospital) or industrial plants. Although the earliest MSWC plants were not equipped with this conversion capability, plants that were built shortly after *were* equipped with it, *e.g.* the first Danish MSWC plant built in 1903 in Copenhagen [111], which is depicted in figure 1.1. Presently, the importance of the function of converting the MSW energy content into heat and/or electricity is growing due to the interest in durable and renew-

[1] In the literature one sometimes speaks of energy *production* (see *e.g.* [67]). However, this implies a violation of the first law of thermodynamics *c.q.* the law of conservation of energy and, therefore, it has been chosen here to refer to energy *conversion*

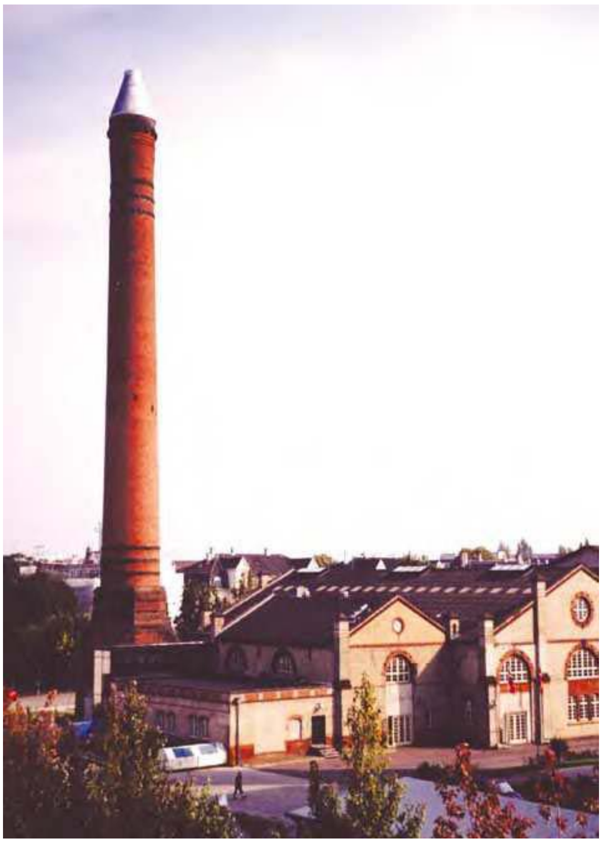

Figure 1.1: One of the earliest MSWC plants: the MSWC plant built in 1903 *at Frederiksberg, Copenhagen, Denmark (printed with permission of Ramboll).*

able energy[2] and the recognition of MSW as partially being a source of this type of energy. The latter view on MSW is a consequence of the fact that MSW typically consists to a large extent of *biomass* [107] (as a result of which the energy that is derived *c.q.* converted from MSW is also denoted as *bioenergy*).

At present, MSWC plays a major role within the whole waste disposal scheme of many countries, with the alternatives being landfill and recycling (incl. composting). In some countries, *e.g.* in Japan, MSWC is even the dominant method of waste disposal. In the Netherlands, MSWC is the second most important waste disposal method ($34\ \%$ of all produced MSW in 2004 [111]), with recycling dominating ($64\ \%$) and landfill playing only a minor role ($2\ \%$).

[2] Renewable sources of energy, in a traditional sense, are (i) those that nature can regrow, such as wood, crops, or other plants (biomass), and (ii) those that are available through the earths unique physical set-up, such as wind, water and solar radiation [69].

1.1.2 The process of municipal solid waste combustion

Typical modern large scale MSWC plants, also denoted as *Waste-To-Energy* (WTE) plants, are depicted in figure 1.2 and, schematically, in figure 1.3. Waste is collected

Figure 1.2: MSWC plant AVR at Rozenburg (printed with permission).

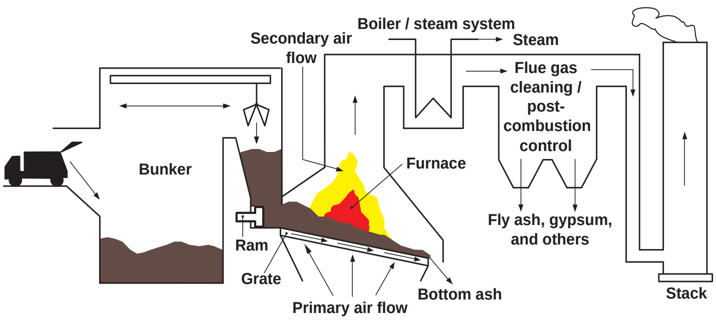

Figure 1.3: Schematic view of a typical modern large-scale MSWC / WTE plant.

from households and transported to the combustion plant by *e.g.* truck or ship. There it is stored in a large bunker from which it is transported by cranes into a large chute. At the bottom of the chute the waste is pushed into the furnace by a ram. While the waste is travelling on this grate, pushed forward by rolls or bars in this grate and possibly moved forward by gravity due to this grate being inclined, it is combusted using O_2 from the so called primary air flow, which is fed to the waste layer from below through holes in the grate. Another air flow denoted as secondary air flow is fed, through the furnace side wall above the waste layer, to the flue gas coming from this layer. Its aim is to cause a second (*post-*) combustion, not only by providing O_2 but also by mixing up the gas, to ultimately reduce certain emissions such as *e.g.* the amount of CO. Further on in the furnace, the flue gas passes a boiler delivering heat which is transformed into steam. This steam is subsequently converted into energy in the form of heat and/or electricity which then is delivered to *e.g.* a surrounding city, public building or industrial plant. After passing the boiler, flue gas cleaning equipment removes as much as possible the combustion residues in the flue gas that are not allowed to be released to the surroundings through the stack.

1.1.3 Operational objectives and constraints for MSWC plants

The operation of modern large scale MSWC plants is determined by economic and environmental objectives and constraints[3]. More specific, the main *economic* objectives and constraints for operators of such plants are determined by the aims of *revenue maximization* and *cost reduction* and are pre-dominantly given as:

- maximization of the waste throughput and energy conversion *c.q.* steam production, being the main sources of revenue for such plants with, notably, the waste throughput accounting for roughly $80\ \%$ of these revenues.
- minimization of variations in process variables (*e.g.* in standard deviation or variance sense) that cause increases in operational and maintenance (broken component replacement and downtime) costs. An example here is the variation in furnace temperature as this causes a reduced lifetime of furnace parts like *e.g.* refractory furnace materials ([110]) or grate bars. Variations in furnace temperature and composition also can cause corrosion [110] and, thereby, an increase in maintenance costs. Also, variations in flue gas related variables may cause a significant decrease in efficiency in flue gas cleaning equipment and, thereby, the costs of operating this equipment. For example, large temperature variations may significantly decrease the efficiency of selective non-catalytic reduction (SNCR) equipment used for NO_x removal as the temperature range in which this equipment works well is relatively small (850 - 1000 $[°C]$).
- fulfillment of constraints to ensure maximal component lifetime and, thereby, minimal maintenance costs, *e.g.* fulfillment of an upper limit on the furnace temperature to prevent too high thermal stresses and associated component breakdown and also, for example, chlorine associated high temperature corrosion of the boiler pipes (flue gas side; see *e.g.* [35, 73]).

Environmental requirements consist of constraints imposed on flue gas emissions in order to reduce health and environmental risks to a minimum. Particularly important such constraints are (see *e.g.* [110]) (i) acid gases like SO_x, HCl, HF and NO_x, (ii) heavy metals like Cd and Hg and (iii) polychlorinated dibenzo*dioxins* and dibenzo*furans*. A very common environmental constraint is also a lower bound (typically of 6 $[Vol.\ \%\ dry]$) on the O_2-concentration in the flue gas to guarantee excess

[3] The terms 'objective' and 'constraint' are used here with an interpretation similar to that used in mathematical optimization theory, *i.e.* with 'objective' being a quantity to be minimized or maximized, similar to an objective function in an optimization problem, and 'constraint' denoting a restriction not to be violated.

air and, thereby, to prevent the formation of CO due to imcomplete combustion.

The order of importance of the objectives and constraints above for a waste incineration plant operator depends on the market situation, in particular on whether the waste market is one of *under*- or *overcapacity*, *i.e.* on whether the supply of waste is more resp. less than can be processed *c.q.* combusted with the current total combustion capacity (*i.e.* with the available waste incineration plants). The main difference between these two market situations is that in the undercapacity one the main focus for an MSWC plant operator lies on maximization of the revenues whereas in the overcapacity situation the focus lies on cost reduction. More specific, in an undercapacity market situation, the waste price is high, which renders a shift of the economically optimal operating point (revenues - costs) to one with maximum revenues while accepting possible larger costs. On the other hand, in an overcapacity market situation, the waste price is low (with waste incineration plant operators competing for the relatively small supply of waste) and the economically optimal operating point shifts to one where cost reduction dominates.

The focus on revenue maximization in an undercapacity market situation translates to a focus on maximization of the waste throughput and steam production while the focus on cost reduction in an overcapacity market situation translates to a focus on minimization of process variations. Likewise, maintenance related constraints might be of lower importance in an undercapacity market situation than in an overcapacity market situation.

Note that the focus on constrained maximization of the waste throughput and steam production in an undercapacity market situation causes the optimal operating point for an MSWC plant in this market situation to be as close as possible to one of the constraints. More specific, maximization of the waste throughput and steam production pushes all plant variables up or down until this so called *dominating* constraint is met, which provides the name of *constraint pushing* operating behavior to this type of operation. Exact fulfillment of this constraint is, however, not possible as a certain distance, also denoted as *back-off*, has to be maintained on average to prevent too frequent violations due to the occurrence of disturbances.

Until recently, the waste market in the Netherlands was one of undercapacity while the last couple of years it has changed into one of overcapacity. In this thesis, both situations are considered. *i.e.* economic improvement is sought for both considered market situations.

1.1.4 The need for improving the economic performance of MSWC plants

Even though fulfillment of the ever becoming tighter environmental constraints is and will remain an important goal and challenge for MSWC plant operators and managers, typically the main goal and challenge for them currently is that of optimizing the economic performance (within the operating envelope determined by the environmental constraints). More specific, they are under an increasing pressure to operate economically more optimally due to the increasing business character of the environment they have to operate in, with market forces and competition increasingly dictating the plant operation. The latter is a consequence of developments like privatisation (towards share-holder based ownership) and (subsequent) public to private ownership changes, with new owners having a stronger focus on profitability. See *e.g.* [75] for a discussion of these developments in the Netherlands.

Motivated by the industrial need for optimizing the economic MSWC plant operation, the aim set in this thesis is to optimize this operation through model based combustion control. In the next section, this choice is motivated via a discussion of the state-of-the-art and challenges in MSWC plant combustion control.

1.2 State-of-the-art and challenges in combustion control of MSWC plants

1.2.1 MSWC plant control systems

Modern large scale MSWC plants employ two types of control systems:

- control systems related to the *flue gas cleaning equipment* (also referred to as *post-combustion* or *air pollution control systems*)
- a (here) so called *combustion* control system

The purpose of flue gas cleaning equipment is, as mentioned in section 1.1, to remove as much as possible the combustion residues in the flue gas that are not allowed to be released to the surroundings through the stack, *i.e.* to reduce the amount of toxic components that enter the environment via the flue gas to below the limits required by law. Flue gas cleaning equipment consists of several independently operating control systems. Examples of this kind of control systems are [44, 48, 79, 86, 92] *electrostatic precipitators*, which are used to remove fly ash/particulate emissions, *fabric (bag) filters*, which are also used to remove fly ash, *scrubbers*, which are used to control the levels of acid gases (SO_x, HCl and HF), and *selective (non-) catalytic reduction control systems*, which are used to remove NO_x.

The combustion control system controls, as the name suggests, the actual waste combustion process *c.q.* the furnace and boiler part of the MSWC plant. Typically, the combustion control system is of the type depicted in figure 1.4[4], *i.e.* its aim is to minimize the deviations between, on the one hand, measured steam production and flue gas O_2-concentration and, on the other hand, corresponding setpoints by manipulating

- the ram frequency, which manipulates the waste inlet flow
- the frequency/speed of the grate bars or rolls, whichever is used at the MSWC plant at hand

[4]Note the nomenclature used in this figure: manipulated variables (MVs) for plant variables that are manipulated by the controller and controlled variables (CVs) for plant variables that are actually to be controlled by this controller. This is according to the nomenclature commonly used in the process control literature.

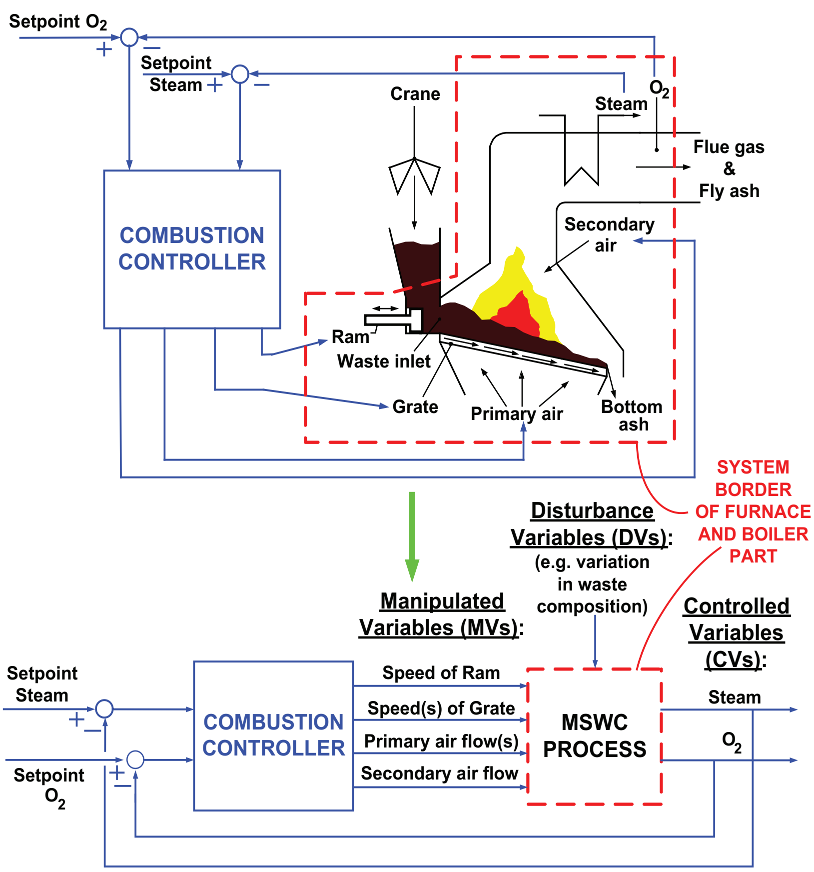

Figure 1.4: Combustion control of MSWC plants.

- the primary air flow(s)
- the secondary air flow

The waste inlet flow is typically manipulated by setting the frequency of one or several rams that push the waste out of the chute onto the grate. The grate most of the times consists of either a large number of partially overlapping bars, each one moving back and forth and thereby pushing the waste from left to right over the grate, or a set of rolls (see *e.g.* [31] for a display of some of the commonly used grate types). The frequencies/speeds of the grate elements, bars or rolls, are manipulated to send the waste through the burning zone of the waste on the grate with a certain residence time. Generally, the grate is divided in several zones (along its length), usually four to five, with the frequency/speed over each zone being constant and independently manipulable, rendering a potential of four to five MVs to move the waste over the grate. A common situation is to use only one MV to steer the zone frequencies/speeds simultaneously with a certain fixed frequency/speed distribution over the zones. This situation is also considered in this thesis. The main function of the primary air flow is to deliver the O_2 needed for combustion. The primary air flow is obtained from the ambient air, spread over the grate zones and then sent through holes in the grate to enter the waste layer. Typically, both the total primary air flow and the air flow for each zone individually are manipulable, although under the constraint that the sum of the zone primary air flows must equal the total primary air flow. As with the grate speeds, a common situation is that the total primary air flow is manipulated whereas this flow is distributed with a fixed distribution over the grate zones. Again, this situation is also considered in this thesis. The secondary air flow, which is also obtained from the ambient air, is fed to the gas phase above the waste layer on the grate. Its main function is, as mentioned in section 1.1, (post-) combustion of unburnt components of the flue gas, in particular CO, by adding extra O_2 to and mixing of the flue gas.

The aims to be fulfilled by the combustion control system are similar to those for the overall operation of the MSWC plant, *i.e.*

- to maximize waste throughput and steam production in order to maximize the revenues

- to minimize the variations in various process variables in order to minimize operational and maintenance costs

- to fulfill constraints imposed out of environmental and maintenance considerations

with a particular focus *c.q.* priority in these objectives and constraints for the MSWC plant operator being determined by the market situation, in particular by whether this market situation is one of under- or overcapacity (see section 1.1.3). The only difference with the overall MSWC plant operation objectives is mainly the smaller number of environmental constraints to be dealt with by the combustion control system as many of these are to be fulfilled by the flue gas cleaning equipment. This all implies that

- in an undercapacity market situation, where the focus of an MSWC plant operator is on revenue maximization, the settings of the combustion control system will be such that its focus is on maximizing steam production and waste throughput within the space defined by the constraints, where fulfillment of maintenance related constraints might have a lower priority, whereas

- in an overcapacity market situation, where the focus of an MSWC plant operator is on cost reduction, these settings will be such that the focus of the combustion control system will be on minimizing process (variable) variations, typically on minimizing those present in steam production and O_2-concentration with respect to the corresponding setpoints, and on remaining within the space defined by the environmental and maintenance constraints, where the latter are certainly of importance.

Note also that the focus of currently employed combustion control systems, see the discussion just above, is typically on minimizing process variations present in steam production and O_2-concentration (with respect to the corresponding setpoints): no combustion control system has been implemented yet that fulfills the aims of (i) maximizing the waste throughput and steam production and (ii) fulfilling the imposed constraints in a direct, explicit manner (*c.q* in a constraint pushing / constrained maximization manner), despite the presence of an undercapacity market situation recently (in the Netherlands).

The process variations in MSWC plants are mainly due to the variation in waste composition, which represents the largest source of unmeasured disturbances on an MSWC plant and can make the fulfillment of the combustion control objectives and constraints very difficult. It is a challenge to find an MSWC plant combustion control strategy that reduces the negative effect of these large unmeasured disturbances on the fulfillment of these objectives and constraints to a minimum.

1.2.2 Improved economic performance of MSWC plants by means of improved combustion control

As may be evident from the discussion above, the combustion control system is the main instrument in fulfilling the overall *economic* objectives of an MSWC plant, while the flue gas cleaning equipment plays a minor role in fulfilling these objectives. This is a first main motivation for pursuing, in this thesis, an approach of improving the combustion control performance in order to improve the overall economic performance of MSWC plants.

Another main motivation for pursuing the approach of improving the combustion control performance in order to improve the overall MSWC plant economic performance is that improving the combustion control system is a relatively low cost solution compared to making fundamental changes to the plant layout/design, *e.g.* by replacing the grate for a water-cooled one (see *e.g.* [45]) to reduce thermal stresses and, thereby, the maintenance costs.

A third main motivation for pursuing the improvement of the MSWC plant combustion control performance for overall economic performance improvement is that there are clear opportunities for improving the performance of currently employed MSWC plant combustion controllers. In the remainder of this section, these opportunities are identified and a specific way of exploiting these opportunities is proposed and motivated, which (indeed) is *model based* combustion control.

1.2.3 Opportunities and challenges for improved MSWC plant combustion control

Currently, (only) two types of combustion control strategies are employed at MSWC plants on a significant scale: PID and fuzzy type. Typically, current MSWC plant combustion control systems are of the (multivariable) proportional-integral-derivative (PID) type (see *e.g.* [22]), *i.e.* consisting of proportional, integral and/or derivative operations on errors between setpoints and corresponding CVs. With the aim to obtain a more stable plant behavior and, thereby, an improved overall economic and environmental performance, *fuzzy control* based combustion controllers have also been implemented on MSWC plants, sometimes in combination with a neural network model and genetic programming techniques [13, 14, 22, 29, 46, 74, 95], though the number of applications is still small. With fuzzy control (see *e.g.* [78]), a set of *fuzzy c.q. if-then* rules are used to control the plant, together with a *fuzzification* interface to transform the measured CV signals into information that can be handled by these rules and a *defuzzification* interface to transform the actions determined by the rules to implementable MV signals. A fuzzy control system is generally designed in a trial-and-error way with the fuzzy rules derived from expert knowledge of the plant dynamics, where the experts are plant operators and/or the control system designer. An important note here is that, although implemented with the aim to improve the performance of PID type of combustion control strategies, the reported applications of fuzzy combustion control of MSWC plants have delivered no such improvement yet [22].

The performance of the currently employed PID and fuzzy type of MSWC plant combustion control strategies is not optimal due to the fact that these control strategies do not optimally handle the main characteristics of the MSWC plant dynamics and combustion control problem, which are in particular

- the multivariable, interacting nature of the MSWC plant dynamics
- (at least in case of a PID type of MSWC plant combustion control strategy:) the nonlinear nature of the MSWC plant dynamics
- the constrained nature of the MSWC plant combustion control problem

Major causes for this non-optimal handling are

- the *inherent* inability of the employed combustion control strategies of optimally handling the mentioned characteristics
- the fact that PID and fuzzy type of combustion control strategies are not designed on the basis of rigorous plant knowledge in the form of a mathematical model but, rather, on the basis of some 'mental' model that operators and/or combustion control system vendors have of the MSWC plant dynamics.

One of the main challenges and opportunities for improved MSWC plant combustion control and operation is the removal of these causes by means of a newly chosen control strategy, resulting in an improved overall economic MSWC plant performance. A control strategy that allows for this removal, and thereby represents one such opportunity and answer to this challenge, is *model based control* and in particular *model predictive control*, as is explained now in more detail.

1.2.4 Improved combustion control through model based control

Model based control is a control strategy whose design is based, implicitly or explicitly, on a set of equations that rigorously describes the dynamics of the plant to be controlled. This rules out *e.g.* control strategies where control design is based on an imprecise, 'mental' plant model and on heuristic tuning rules, as is typically the case with MSWC plant combustion control strategies of the PID type. Also, this rules out control strategies where models are employed consisting of a set of if-then rules obtained in a heuristic (trial-and-error) manner, as is the case with fuzzy control.

The usage of a rigorous model for control design allows for a systematic and, thereby, optimal handling of multivariable, interacting processes, processes with time delays and/or with inverse responses (non-minimum phase behaviour). Also, unstable process can be handled. Additionally, model based control allows, in the form of *model predictive control* (MPC), for a systematic handling of multiple, conflicting objectives and constraints.

With MPC the control problem to be solved, *i.e.* constraints and objectives, is formulated explicitly and translated to a constrained optimization problem that is to be solved online. Given the state of the plant to be controlled at a certain time instant, predictions of the dynamic plant behavior over some a priori chosen future time horizon are made for given trajectories of the MVs and, by solving the optimization problem, that trajectory for the MVs is selected that is optimal in the sense of the *a priori* defined control problem. See figure 1.5. The optimal control *c.q.* constrained

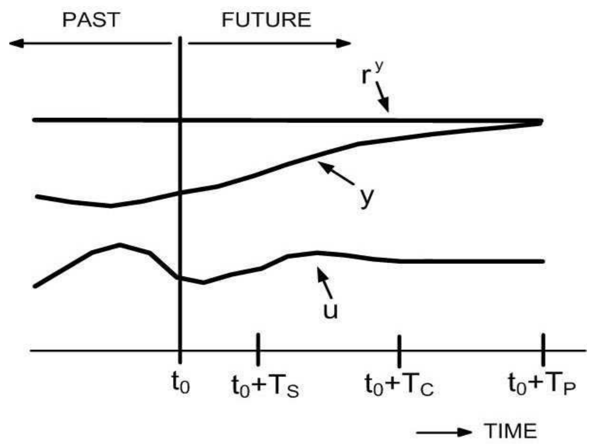

Figure 1.5: Model Predictive Control (see e.g. [25]); r^y = desired CV trajectory; y = real (past) and predicted (future) CV trajectory; u = real (past) and predicted (future) MV trajectory; T_s = sampling time; T_c = control horizon, i.e. over which MVs are optimized; T_p = CV prediction horizon.

optimization problem is re-solved each (sample) time the state of the plant to be controlled is determined, which is done through measurements and/or estimation. As a result of this strategy, MPC is also referred to as *moving* or *receding horizon control*.

Note that it is the optimization based nature of MPC that allows for a systematic handling of multiple, possibly conflicting control objectives and constraints.

It may be clear from the discussion here that, due to the usage of a rigorous plant model and (in case of MPC) due to its optimization based nature, model based control represents a very suitable control strategy for improving the combustion control performance of MSWC plants and, thereby, their overall economic performance.

Actually, model based control has already been considered for obtaining this improvement but that has not lead to an application yet, in contrast to PID and fuzzy based combustion strategies. In the next section, the results of these considerations *c.q.* the state-of-the-art and challenges in model based combustion control of MSWC plants are discussed with the main aim of identifying suitable research directions when exploring the opportunities of model based combustion control for improving the overall economic performance of MSWC plants.

1.3 State-of-the-art and challenges in model based combustion control of MSWC plants

A significant amount of work has been done on model based combustion control of MSWC plants, in particular with respect to the derivation of the model needed for that purpose. However, no full MSWC plant combustion control application has been studied yet, neither in a simulation or real-life setting, although closely related control problems have been studied in a model based control perspective. More specific, particularly motivated by obtaining an

- optimal, rather than sub-optimal, handling of the variability of the waste composition
- optimal handling of the severity of the combustion conditions
- optimal handling of operational constraints
- optimal economic operation
- a systematic rather than an intuitive and experience based operation

of the MSWC plant, Rovaglio, Manca and co-workers [90, 89, 67, 16] have pursued a model based approach to the MSWC plant combustion control problem and closely related control problems. Their approach uses relatively low order first-principles models based on material, energy and momentum balances. Such models have been derived and validated for several types of MSWC plants, including of the type considered here (see [16]), and for several types of MSWC plant control problems. Model based control solutions have been proposed and evaluated for (i) inverse response compensation [67], with the inverse response present in the transfer function from the ram frequence/waste inlet flow to steam production, and (ii) the reduction of NO_x in the MSWC plant flue gases [89]. With the first application, the aim is to optimally reduce the oscillations introduced by the combination of an inverse response and a too tightly tuned controller. For that purpose, amongst others, a model is used to replace measurements used by the controller for *predictions*. With the second application, a combined conventional/MPC control system is proposed and evaluated (although the MPC part is not referred to as such but as an *on-line optimization* strategy). Although not considering the full MSWC plant combustion control problem the corresponding reference is of much interest here because the considered control problems are close in nature. Notably, the control strategy proposed in [89] is claimed to succesfully have been applied on a real-life MSWC plant, although the corresponding results are not presented. Finally, in [16] a model based solution to the full MSWC plant combustion control problem is pursued but only a solution to the modeling problem is proposed.

Apart from the work of Rovaglio, Manca and co-workers, other work of particular interest here is the first-principles modeling work in [41], which provides dynamic MSWC plant models very well suited for model based control due to their low complexity *c.q.* order.

Specific observations from the literature on model based combustion control of MSWC plants are:

- Only nonlinear first-principles models are considered whereas linear empirical models are favorite in other (non-MSWC plant combustion control related) model based control applications, in particular in MPC applications: see *e.g.* [83]. Linear empirical MSWC plant models may be less costly to obtain than nonlinear first-principles models. Also, linear model based combustion control strategies may lead to a similar improvement in combustion control performance as nonlinear ones. In other words, both the optimal way of MSWC plant modeling (first-principles based versus empirical) and the choice of model structure and resulting nature of the control strategy (linear versus nonlinear) are still unresolved issues.

- The derivation of first-principles model equations for model based MSWC plant combustion control applications is a mature area.

- However, there is no conclusive evidence available yet, through validation on real-life MSWC plant data, that the resulting first-principles models are capable of sufficiently well representing the MSWC plant dynamics in the operating range excited in the final (model based) control application. More specific, the MSWC plant model validation exercises encountered in the literature (see, again, [16] and [90]) have been performed in open-loop, rather than in closed-loop. Also, start-up data not containing the faster MSWC plant dynamics have been used for validation [90], *i.e.* (not containing) dynamics that are expected to be highly relevant for control.

- All implementational issues with respect to model based MSWC plant combustion control other than those related to modeling still need to be sorted out. Particular challenges here are *e.g.*

  - The choice of the type of model based control. For instance: should MPC be chosen or does it suffice to use a control strategy that is more easy to implement, *e.g.* a model based improvement of the conventional PID type of MSWC plant combustion control strategies? Also, does a linear model based control strategy suffice or is it necessary to use a nonlinear one to obtain a significant improvement in control performance?

  - Minimization of the computational complexity. This is particularly relevant for nonlinear model based control strategies, where the usage of a nonlinear model easily leads to a large computation time.
  - The implementation of measures that guarantee a stable closed-loop (MSWC) plant.
  - The choice of state estimator, *i.e.* a tool that provides optimal model based inference of un- or inaccurately measured plant variables from (other) measured ones. This is particularly relevant for model predictive control applications where typically a state estimator is required to account for absence of measurements or the presence of noise.
  - Providing robustness against model error, *i.e.* obtaining a sufficient control performance even in case of model errors.

  These and other issues need to be properly addressed if one is to obtain a well performing model based MSWC plant combustion control strategy.

- All *performance* issues with respect to model based MSWC plant combustion control other than those related to modeling still need to be sorted out. More specific, even though model based control has been put forward as a technique to significantly improve conventional (PID and fuzzy type) MSWC plant combustion control strategies, this improvement has not been demonstrated yet, even not in a simulation setting for an *a priori* assessment (which is helpful in the often difficult task of convincing MSWC plant operators and managers to use model based control at their plants). More specific, it has not been demonstrated yet that model based MSWC plant combustion control strategies handle much better the main characteristics of the MSWC plant dynamics and combustion control problem such as nonlinearity, interaction and constraints to, thereby, lead to an improvement of the overall economic MSWC plant performance.

With these issues and the earlier provided motivation for investigating model based MSWC plant combustion control in mind, the main problem addressed in this thesis can now be formally stated and a specific solution strategy can be given to tackle this problem. This is done in the next section. After that, the contents and main contributions of this thesis are outlined.

1.4 Problem statement and solution strategy

1.4.1 Main thesis research objective

Motivated by the

- the industrial need to improve the economic performance of MSWC plants
- the fact that this performance improvement can be obtained in a more cost-effective way through controller improvement than through design/layout changes

- the major influence of the combustion control system of an MSWC plant on its economic performance
- the potential of a large improvement in MSWC plant combustion control performance by replacing the conventional such control systems for model based ones

the following main research objective is addressed in this thesis:

Explore the opportunities of model based combustion control in improving the economic performance of MSWC plants

In order to fulfill this objective in a tractable manner, a specific solution strategy has been chosen in this thesis that takes into account the issues and challenges in model based MSWC plant combustion control as presented in the previous section. This solution strategy is outlined now.

1.4.2 Solution strategy

In order to allow a tractable fulfillment of the main research objective, it is decomposed into research objectives related to modeling issues and one related to the remaining issues.

With respect to the modeling issues, the following research objectives are addressed, which all follow from the modeling related issues and challenges of model based MSWC plant combustion control presented in section 1.3:

- *Explore the opportunities of first-principles modeling for obtaining a model suitable for model based MSWC plant combustion control*
- *Explore the opportunities of linear empirical modeling for obtaining a model suitable for model based MSWC plant combustion control*

As discussed in section 1.3, the first objective has been largely but not fully addressed in the literature. More specific, the derivation of the equations of first-principles MSWC plant combustion models is a mature area but the resulting models have not been fully validated yet within the operating range excited in the final (model based control) application. This research objective is added here with the aim to find more conclusive evidence that this type of model is able to accurately simulate the corresponding dynamics. Tackling the second objective has been made possible by the fortunate circumstance that experiments could be performed at MSWC plants for the purpose of linear empirical modeling. To limit the scope, the more difficult nonlinear empirical modeling techniques [64] are not considered here. Notably, the added value of linear empirical modeling is also determined by the added value of this model structure for model based MSWC plant combustion control. This issue is addressed below in another non-modeling related research objective.

With respect to the non-modeling related issues of model based MSWC plant combustion control, the following research objective is addressed in this thesis, which again follows from the discussion on the state-of-the-art and challenges in model based MSWC plant combustion control presented in section 1.3:

- *Given a suitable model of the MSWC plant combustion process, explore the opportunities of model based control for improving the combustion control performance of MSWC plants, in particular such that the overall economic performance of such plants is improved.*

To allow for a comprehensive yet also limited assessment of the improvement that can be obtained with model based MSWC plant combustion control, three types of such control strategies are compared:

- linear MPC (LMPC), which refers to an MPC strategy that employs a linear model and a linear or quadratic optimal control problem formulation
- nonlinear MPC (NMPC), which refers to an MPC strategy employing a nonlinear model and/or a nonlinear and non-quadratic optimal control problem formulation
- PID type of combustion control which design is, in contrast to common practice, based on a model describing the MSWC plant combustion dynamics

The motivation for comparing these combustion control strategies is that this allows for establishing the role of handling the nonlinear and the multivariable, interacting nature of the MSWC plant dynamics in obtaining a significantly improved combustion control performance (which assessment may have a practical consequence like *e.g.* that one can resort to the possibly more cheaper to perform and/or implement linear modeling and model based control techniques to already obtain a significant improvement in combustion control performance, rather than their nonlinear counterparts). The objective of determining the improvement that can be obtained with a model based, rather than a non-model based, PID type of combustion control strategy has been added here also to assess whether one can resort to this much cheaper to implement type of combustion control strategy, with the corresponding control soft- and hardware and architecture generally already in place, rather than its MPC counterpart.

The research objective above also includes a search for optimal choices with respect to the implementational issues related to model based MSWC plant combustion control. In particular, NMPC related implementational issues are considered here as these are the most challenging to resolve.

1.5 Outline and main contributions of the thesis

The contents of the thesis follow the decomposition of the main research objective into a modeling part and a part on the remaining, performance and implementation related, model based MSWC plant combustion control issues. More specific, the first next three chapters are on modeling while the subsequent three chapters are on the remaining issues.

In chapter 2, first results are presented of the exploration of the opportunities of first-principles modeling for obtaining a model suitable for model based MSWC plant combustion control. These results concern the actual modeling part whereas model quality *c.q.* validation aspects are discussed in chapter 4. More specific, after a brief review of the first-principles modeling work related to model based MSWC plant combustion control, a specific first-principles model is proposed to be used for that purpose. This model is a simple extension of an existing model and represents the first main contribution of this thesis. The motivation for the proposed extension is to allow the model to be used in combination with a model based on-line waste composition and calorific value estimator recently proposed in the literature [106] and thereby (through providing an on-line estimate for unmeasured disturbances) to allow for an improved simulation, validation and model based MSWC plant combustion control performance.

After that, in chapter 3, first results are presented of the exploration of the opportunities of linear empirical modeling *c.q.* system identification for obtaining a model suitable for model based MSWC plant combustion control. More specific, issues *c.q.* potential obstacles for arriving at such a model are identified together with solutions to overcome these obstacles. Subsequently, from the identified solutions a methodology is derived that is claimed to be suitable for obtaining a model suitable for model based MSWC plant combustion control. This methodology is the main contribution of this chapter. Another contribution is the solution to the so called partial closed-loop identification problem, which handling here is motivated by the fact that it may be encountered at MSWC plants when aiming to derive a model through system identification.

Subsequently, in chapter 4, the exploration of the opportunities of both first-principles modeling and linear system identification for obtaining a model suitable for model based MSWC plant combustion control is finalized. More specific, through an application of these modeling approaches, in particular those proposed in chapters 2 and 3, on data experimentally obtained from a large scale Dutch MSWC plant a final assessment is made with respect to their ability of delivering such a model. This is the main contribution of this chapter. Another contribution is a new - system identification based - way of validating first-principles MSWC plant models.

In chapter 5 the question is addressed whether currently employed PID-type of MSWC plant combustion control strategies can already be improved via a model based design approach, *i.e.* without having to resort to advanced model based control strategies like *e.g.* MPC. More specific, starting point here is a PID type of combustion control strategy that is applied in practice. This control strategy is compared to a new PID combustion control strategy proposed in this chapter and which is derived from a closer investigation of the MSWC plant dynamics, in particular the dynamics exhibited by models that have been obtained through the modeling approaches of chapters 2 and 3. The results from this comparison are used to assess the margin for improvement present for PID based MSWC plant combustion control. The main contributions of this chapter are the results of this assessment and the new PID-type of combustion control strategy.

In chapter 6 implementational issues related to model based MSWC plant combustion control are addressed, more specific those related to LMPC and NMPC based MSWC plant combustion control. In particular, both an LMPC and an NMPC based MSWC plant strategy are presented that have been found to be suitable for MSWC plant combustion control and which contain solutions proposed to resolve the identified implementational issues. These control strategies and solutions are the main contributions of this chapter. The focus of this chapter is largely on NMPC related implementational issues as these are the most challenging to resolve.

In chapter 7, then, the opportunities of MPC are explored for improving the combustion control and overall economic performance of MSWC plants. This is done through assumed close-to-realistic simulations involving the linear and nonlinear MPC strategies presented in chapter 6. The main question addressed in chapter 7 is how and to what extent the favorable properties of MPC (see *e.g.* sections 1.2.3 - 1.2.4) can actually be exploited to improve the mentioned MSWC plant performances. The answer to this question is the main contribution of this chapter.

Finally, in chapter 8, the main conclusions of this thesis and recommendations for future work are given.

It is noted that part of this thesis is contained in a number of published journal and conference papers. More specific, the MSWC plant system identification methodology presented in chapter 3 and its application discussed in chapter 4 have been published in [56]. Some of the system identification and first-principles modeling results have also been discussed in [51], which also includes a discussion on and application of the calorific value sensor. The partial closed-loop identification problem and its solution presented in chapter 4 have been discussed in [58]. The novel PID MSWC plant combustion control strategy presented in chapter 5 has also been discussed in [59]. Some of the NMPC results presented in chapters 6 and 7 have been published in [61]. The moving horizon estimation based NMPC approach discussed in appendix E can also be found in [57]. In addition, other results by the author related to the subject of this thesis but which are not discussed here can be found in [55], which includes a first assessment of LMPC for improved MSWC plant combustion control, and in reports written as part of a number of European Union projects (incl. the ECOTHERM and NextGenBioWaste projects) into which the author of this thesis was involved: see *e.g.* [54] and [60]. Finally, the work on model based control of MSWC plants has also resulted in a patent: see [43].

Chapter 2

First-principles modeling for model based MSWC plant combustion control

2.1 Introduction

In this chapter, results are presented of the exploration of the opportunities of first-principles modeling for obtaining a model suitable for model based MSWC plant combustion control. These results concern the actual modeling part whereas model quality aspects are discussed in a later chapter. More specific, it is shown that the literature provides ample opportunities for first-principles MSWC plant models suitable for model based combustion control. One of these models is chosen and, being the main contribution of this chapter, extended to arrive at a new model that allows for an improved simulation, validation and model based combustion control performance compared to existing models. More specific, this extension is due to the incorporation of the equations underlying the so called calorific value sensor (CVS) [106], which is an on-line estimator of the MSWC plant waste composition and calorific value. This incorporation leads to a more detailed description of the waste composition in the model which, combined with the ability to estimate the main parameter of this description from large scale MSWC plant plant data through the CVS, allows for the mentioned improvements. In other words, these improvements are due to the ability to incorporate more information on the main source of MSWC plant disturbances in the new model.

The contents of this chapter are as follows. First, in section 2.2, the literature on MSWC plant modeling for the purpose of model based combustion control is reviewed. Here, it is also motivated why the specific model that forms the basis for the new model to be proposed here has been chosen for that purpose. After that, in section 2.3, this new model is outlined. In this section, first, the existing model underlying this new model is outlined. After that, the CVS equations are provided and the integration of these equations in the existing model are discussed, to form the new model proposed here. Section 2.3 also includes a system theoretic investigation of the new model where it, amongst others, is shown that this model can be reduced in complexity when used for simulating normal combustion conditions. Finally, in section 2.4, the main conclusions of this chapter are given.

2.2 State-of-the-art of first-principles modeling for model based MSWC plant combustion control

The literature provides ample opportunities for the derivation of first-principles models suitable for model based MSWC plant combustion control. Here, the corresponding literature is reviewed and these opportunities are identified. This section also includes the choice for one of the identified models as the basis for the new model to be presented in the next section, together with a motivation for this choice.

For completeness, the literature review starts with a more global historic overview of the developments with respect to first-principles modeling of MSWC plants.

Historic overview of first-principles modeling of MSWC plants

Most of the work on modeling of MSWC plants has been based on first-principles. The earliest attempts to model the MSWC process date back to the early seventies [119]. The incentive for this modeling work were the CO and dioxin problems that MSWC plants were facing back then. The first modeling efforts focused on the gas phase with computational fluid dynamics (CFD) being the main modeling tool. Five to ten years ago the focus shifted to the modeling of the waste layer on the grate, as no satisfactory model of it was available until then [31]. Accurate modeling of the underlying thermal and chemical processes proved, and still proves, to be a challenging problem to scientists due the wide variations in waste composition and the many different pollutants they may be generating [119]. The waste layer models that can be found in the literature are used to obtain a better understanding of the influence on the combustion process of process parameters such as *e.g.* particle mixing [116], channel formation [118], moisture level and devolatilisation [117]. Often, the obtained waste layer models are coupled with already available CFD models for the gas phase. An overview of first-principles modeling work on MSW combustion can be found in [119]. Much work in this direction has particulary been performed by Swithenbank *et al.* [30, 31, 96, 98, 116, 117, 118, 119] at the Sheffield University Waste Incineration Centre. Part of the first-principles MSWC plant modeling work is oriented on model based control, in particular on model based combustion control and closely related control applications. This work is discussed below in more detail.

State-of-the-art of first-principles modeling for model based MSWC plant combustion control

A significant amount of work has been done on first-principles modeling for model based MSWC plant combustion control and closely related control applications. In particular, Rovaglio, Manca *et al.* [89, 90, 67, 16] have pursued a first-principles modeling approach for such control applications. Their models describe the dynamics of part or all of the furnace and boiler part of the MSWC plant through mass, energy and momentum balances. These models are of a relatively low complexity (order) while still being rich in detail from a control point of view. Also, many of these models have been validated on large scale plant data with positive outcome. All these characteristics indicate that low order first-principles models are suitable candidates for model based control of MSWC plants.

Within the work of Rovaglio, Manca *et al.*, the most complete, recent and interesting model is discussed in [16], where a full dynamic model of the furnace and boiler part of the MSWC plant is provided and validated on large scale plant data. Also, this model is particularly aimed at model based combustion control. A similar but less extended version of this model can be found in (the appendix of) [67], where the aim is inverse response compensation though improved control [67] with the inverse response present in the transfer function from the ram frequence/waste inlet flow to steam production. Other relevant models can be found in [90] and [89]. These models, however, describe the dynamics of another type of MSWC plant (*i.e.* one with a rotary kiln) than considered here, although these dynamics are similar. Moreover, in [89] the focus of the model is on describing the NO_x content in the flue gas.

Apart from the work of Rovaglio, Manca *et al.*, the only other work of sufficient interest here is the first-principles modeling work by Van Kessel *et al.* [32, 41, 42]. Of particular interest here is the work presented in [41], where similar but (even) more simple models are presented as in [16], *i.e.* models fully and only aimed at the dynamics relevant for the MSWC plant combustion control problem.

In this chapter, one of the models in [41] is used as the basis for the new model proposed here to be used for model based MSWC plant combustion control. A first reason for using this model as a starting point and not *e.g.* the model used in [16] is its lower complexity, which typically is advantageous from a model based control point of view. Another reason is that this model has been validated (too) on large scale MSWC plant with successfull outcome, which is discussed in more detail later on in this thesis.

2.3 A new first-principles model for model based MSWC plant combustion control applications

2.3.1 Introduction

In this section, a new first-principles model for model based combustion control of MSWC plants is outlined. This model is a simple extension of a model that is already available in the literature [41]. This extension is largely due to the integration of the equations underlying the so called calorific value sensor (CVS) [106], which is an on-line estimator of the MSWC plant waste composition and calorific value, and is motivated by the fact that it allows for an improved simulation, validation and model based combustion control performance compared to existing models. More specific, the incorporation of the CVS equations leads to a more detailed description of the waste composition in the model which, combined with the ability to estimate the main parameter of this description from large scale MSWC plant plant data through the CVS, allows for the mentioned improvements. In other words, these improvements are due to the ability to incorporate more information on the main source of MSWC plant disturbances in the new model. These improvements are not quantified in this thesis but are subject for future work.

Apart from being extended with the CVS equations, the literature model is also extended with variables representing air leakage and recirculation flow. These are typically, though not necessarily, occurring MSWC plant variables. The inclusion of these two flows is performed to further enhance the applicability of the new first-principles MSWC plant model to be presented here, in particular also for model based combustion control.

To clearly demonstrate the contributions made by incorporating the CVS equations, the equations of the new model are provided here in an indirect manner as adaptations to the chosen literature model. More specific, first this model is outlined. After that, the CVS equations are discussed. Subsequently, the integration of these equations into the literature model is discussed, thereby providing the equations of the new model proposed here. Following that, system theoretic properties of the model are discussed and, via a study of one of these properties, it is shown that the computational complexity of the new model can be reduced. This is beneficial for model based control design as the computational complexitiy of such a control strategy is generally dependent of the model computational complexity.

2.3.2 Review of an existing model

Introduction

In his thesis [41], Van Kessel presents two slightly different first-principles based MSWC plant models suitable for model based control design, with the first one being a slightly more complex version of the second one due to a more complex assumption on the waste composition. More specific, the more simple model assumes the waste layer to consist of one single mass containing all material, *i.e.* combustibles, inert and water, whereas the more complex one assumes this layer to consist of two separate masses, one for the inert and combustibles, and one for the water part of the waste. The advantage of the latter model is that it allows for a better incorporation of the control relevant MSWC plant dynamics due to its ability to capture the inverse response present in the transfer function from waste inlet flow to steam production (meaning that the steam production will first decrease below its original operating point before ending up above this operating point when a step is applied to the waste inlet flow), whereas the more simple model does not have this ability. Capturing this inverse response well in a model to be used for MSWC plant combustion control design is important as failure to do so may lead to a significantly degraded control performance. See *e.g.* [67]. Because of that and because the more complex model allows for a better incorporation of the control relevant MSWC plant dynamics, this model has been chosen here as the starting point for the derivation of the new first-principles model proposed here for model based MSWC plant combustion control. The equations and main underlying assumptions of this model are outlined now in detail. This also provides the major part of the equations and main underlying assumptions of the new model.

Global overview and main underlying assumptions

The more complex model of [41] describes the dynamics of the furnace and boiler part of the MSWC plant, more specific the dynamics of that part of the MSWC plant within the system border depicted in figure 1.4. The main inputs of the model are

- the waste flow entering the waste layer in the furnace ($\phi_{w,in}$)
- the primary air flow (ϕ_{prim})
- the secondary air flow (ϕ_{sec})
- the (so called) interfacial area (a)

as these are related to the MVs of relevance for the MSWC plant combustion control problem, where it is noted that the waste inlet flow is a function of the ram speed MV and the interfacial area is a function of the grate speed MV. The main outputs of the model are

- the steam production *c.q.* flow (ϕ_{st}) produced by the MSWC plant
- the oxygen concentration in the flue gas ($Y_{O_2,fg}$)

as these are the main CVs of the MSWC plant combustion control problem. The main equations describing these dynamics are mass and energy balances for the solid waste layer on the grate and for the gas phase above it and an energy balance for the steam system. The model has been derived by means of a high degree of lumping and under the CSTR assumption (continuously stirred tank reactor; see *e.g.* [114]), which are the main causes for the resulting model to contain only a very low number of ordinary differential equations (eventually four). Additionally, the gas phase above the waste layer is assumed to be stationary, with no accumulation of mass ($d.../dt = 0$) assumed. Also, the solid waste layer on the grate is assumed to consist of only a moisture part and a combined combustible and inert part. Finally, a simplified description of the dynamics of the steam system is used that has been obtained from [4].

The equations of the literature model and further, minor simplifications and assumptions are now discussed in detail, starting with the equations describing the dynamics of the solid waste layer on the grate. It is noted that not all equations are explicitly stated in [41]. The implicitly given ones, however, are readily derived from the equations present in this thesis. Also, a small modification of the model equations is suggested with respect to the combustion kinetics.

Dynamics of the solid waste layer on the grate

The literature model assumes the waste layer to consist of a combustible and inert mass part M_{comb} $[kg]$ and a moisture mass part M_{mois} $[kg]$. The balances for these masses are given as

$$\frac{dM_{comb}}{dt} = X_{comb}\phi_{w,in} - \phi_{out} - RaM_{comb} \tag{2.1}$$

and

$$\frac{dM_{mois}}{dt} = X_{mois}\phi_{w,in} - k_{evap}aM_{mois} \tag{2.2}$$

Here, $\phi_{w,in}$ $[kg/s]$ represents the waste flow entering the waste layer through the rams, X_{mois} $[kg/kg]$ the moisture mass fraction of this flow and $X_{comb} = 1-X_{mois}$ $[kg/kg]$ its combustible and inert mass fraction. The mass flow terms $k_{evap}aM_{mois}$ $[kg/s]$ and RaM_{comb} $[kg/s]$ represent the amount of waste that leaves the solid waste layer via the gas phase through evaporation resp. combustion, with a $[m^2/kg]$ representing the interfacial area and with the parameters k_{evap} $[kg/(m^2\ s)]$ and R $[kg/(m^2\ s)]$ determining the rate of evaporation of the moisture part of the waste layer resp. the rate of combustion of its combustible part. The mass flow ϕ_{out} $[kg/s]$ represents the unconverted part of the waste layer that leaves the grate as bottom and fly ash, which is modeled as

$$\phi_{out} = X_{inert}\phi_{w,in} \tag{2.3}$$

where X_{inert} $(\leq X_{comb})$ $[kg/kg]$ represents the inert mass fraction of $\phi_{w,in}$.

The *energy* balance for the solid waste layer on the grate is modeled as

$$\begin{aligned}(C_{p,comb}M_{comb} + C_{p,mois}M_{mois})\frac{dT_s}{dt} = {} & (C_{p,comb}X_{comb} + C_{p,mois}X_{mois}) \times \\ & \phi_{w,in}(T_{w,in} - T_s) + \\ & C_{p,fg}\phi_{prim}(T_{prim} - T_s) + \\ & X_{\Delta H_{comb}}RaM_{comb}\Delta H_{comb} + \\ & k_{evap}aM_{mois}\Delta H_{evap} + \\ & \sigma A_s(\varepsilon_{gt}T_g^4 - \alpha_{gt}T_s^4)\end{aligned} \tag{2.4}$$

In this energy balance, the accumulation of heat in the waste layer is the result of convective energy flows entering through the waste inlet flow and primary air flow (the first two terms on the right side of the equation), heat resulting from the thermal *c.q.* combustion and evaporation processes taking place in the waste layer (third and fourth term on the right side) and heat resulting from radiation from the gas phase to the waste layer (final term on the right side). $C_{p,comb}$, $C_{p,mois}$ and $C_{p,fg}$ $[J/(kg\ K)]$ represent the specific heat capacities of M_{comb}, M_{mois} resp. the flue gas, where it is assumed that the specific heat enthalpy of the primary air flow ϕ_{prim} $[kg/s]$ is equal to that of the flue gas. The error that the latter approximation introduces is less than $1\ \%$ [41]. The primary *c.q.* ambient air heat capacity $C_{p,air}$ may, however, be better known and be used as an approximation to $C_{p,fg}$, rather than the other way round. Note that the specific heat capacity of the waste inlet flow is computed as the mass fraction weighted average $(C_{p,comb}X_{comb} + C_{p,mois}X_{mois})$. T_s $[K]$ represents the temperature of the waste layer, which is assumed to be equal all over this layer. Furthermore, $T_{w,in}$ and

T_{prim} $[K]$ represent the temperatures of the waste inlet flow resp. primary air flow, which both may be replaced for the ambient air temperature T_{air} without introducing a significant error; ΔH_{comb} $[J/kg]$ represents the combustion enthalpy of the combustible and inert part of the waste and ΔH_{evap} $[J/kg]$ the evaporation enthalpy of the moisture part of the waste. $X_{\Delta H_{comb}}$ $[-]$ represents the fraction of the total energy resulting from combustion $RaM_{comb}\Delta H_{comb}$ that is fed to the solid waste layer, while the remaining fraction $(1 - X_{\Delta H_{comb}})$ is assumed to be fed to the gas phase above it (which may be *e.g.* due to phenomena such as *channeling*). Note that no such split up is assumed for the evaporation heat transfer. $\sigma = 5.67051 * 10^{-8}$ $[W/(m^2\ K^4)]$ is the constant of Stefan-Boltzmann; A_s $[m^2]$ represents the area over which the transfer of radiation energy takes place ($\approx$ area of combustion, which is of the same order of magnitude as the area of the grate); T_g is the temperature of the gas phase, which is assumed to be equal all over this phase; and ε_{gt} $[-]$ and α_{gt} $[-]$ (typically $\varepsilon_{gt} = \alpha_{gt}$ [41]) represent constants related to the emissivity and absorptivity of the flue gas. See appendix A of [41] for a derivation of this radiation term and the assumptions made in obtaining this term. In fact, a linearized version of this radiation energy term is proposed in [41] with the motivation that it introduces an error of at most 5 % under normal MSWC operating conditions. Here, however, it is proposed to use the original form of this radiation energy term as its usage introduces no (numerical) problems whatsoever (see the discussion on the gas phase energy balance later on in this section) while it additionally is a more accurate description of the radiation heat transfer process.

Combustion kinetics

At first, the combustion reaction rate R is described in [41] with the formula for a first order surface reaction rate with externally limited mass transfer:

$$R = \left(\frac{1}{\frac{1}{k_d} + \frac{1}{k_0 e^{\frac{-E_a}{R_g T_s}}}} \right) \left(\frac{Y_{O_2,in} M_{O_2}}{\nu_{O_2}} \right) \tag{2.5}$$

Here, $Y_{O_2,in}$ is the oxygen concentration $[mol/m^3]$ of the primary air flow c.q. the ambient air: $Y_{O_2,in} = Y_{O_2,air}$. ν_{O_2} represents the amount of oxygen that is consumed during combustion per amount of waste in $[(kg\ O_2)/(kg\ waste)]$. The remaining parameters are: the mole mass of oxygen $M_{O_2} = 2*0.016 = 0.032$ $[kg/mol]$, the mass transfer coefficient k_d $[m/s]$, the pre-exponential constant k_0 $[m/s]$, the activation energy E_a $[J/mol]$ and the universal gas constant $R_g = 8.3145$ $[J/(molK)]$. Model validation results, however, suggest that for MSWC plants the reaction rate may very well be modeled as

$$R = k_d \left(\frac{Y_{O_2,in} M_{O_2}}{\nu_{O_2}} \right) \tag{2.6}$$

i.e. that the reaction rate for MSWC plants is mass transfer dominated, in which case $k_d << k_0 e^{\frac{-E_a}{R_g T_s}}$. For that reason it is recommended here to use expression (2.6) for modeling R.

In [41] the mass transfer coefficient k_d is determined through a Sherwood relation for mass transfer:

$$k_d \quad = \quad Sh\left(\frac{D}{d_p}\right) = C_{Sh}\ Re^{0.5}\ Sc^{0.33}\left(\frac{D}{d_p}\right) \tag{2.7}$$

where Sh $[-]$ represents the Sherwood number, d_p $[m]$ is a representation for the diameter of the waste particles, D $[m^2/s]$ is the diffusion coefficient of oxygen in air, Re $[-]$ is the Reynolds number, Sc $[-]$ is the Schmidt number and C_{Sh} is some constant. However, due to the presence of many unknown parameters and due to the absence of any other model variable in (2.7), this expression for k_d is not in a really tractable form. It has therefore been proposed [1] to replace eqn. (2.7) for the more tractable one

$$k_d \quad = \quad C_{k_d}\sqrt{\phi_{prim}} \tag{2.8}$$

with C_{k_d} some constant. Note that this expression for k_d can be derived from (2.7) by replacing the velocity in the Reynolds number by the primary air flow and lumping the remaining parameters into one single scalar C_{k_d}. Note also the independence of the expression from the waste layer temperature T_s.

'Dynamics' of the gas phase

The gas phase is assumed to be stationary, which is valid due to the gas phase dynamics generally being much faster than the waste layer dynamics. See [41] for a motivation for this assumption in terms of mathematical expressions of the relevant time constants. Under the stationarity assumption, the energy balance for the gas phase is given as

$$\begin{aligned} 0 \quad = \quad & C_{p,fg}\phi_{prim}(T_s - T_g) + C_{p,fg}\phi_{sec}(T_{sec} - T_g) + \\ & (1 - X_{\Delta H_{comb}})RaM_{comb}\Delta H_{comb} + \sigma A_s(\alpha_{gt}T_s^4 - \varepsilon_{gt}T_g^4) + \\ & C_{p,fg}(RaM_{comb})(T_s - T_g) + C_{p,fg}(k_{evap}aM_{mois})(T_s - T_g) \end{aligned} \tag{2.9}$$

with ϕ_{sec} $[kg/s]$ the secondary air flow and T_{sec} $[T]$ the temperature of this flow. The stationary mass balance for the gas phase is used to compute the total flue gas flow ϕ_{fg} $[kg/s]$:

$$\phi_{fg} \quad = \quad \phi_{prim} + \phi_{sec} + RaM_{comb} + k_{evap}aM_{mois} \tag{2.10}$$

The energy balance (2.9) contains similar terms as those in the energy balance for the waste layer (2.4). Note that the convective energy flow term due to $\phi_{w,in}$ is replaced by one corresponding to the secondary air flow ϕ_{sec} (which is assumed to enter only the gas phase but not the waste layer) and by one corresponding to the combusted and evaporated waste flow RaM_{comb} resp. $k_{evap}aM_{mois}$. The temperature T_{sec} may again, *i.e.* as was the case with $T_{w,in}$ and T_{prim} in eqn. (2.4), be replaced for the ambient air temperature T_{air} without introducing a significant error.

The energy balance (2.9) is used to calculate T_g. When using the linearized version for the radiation term as given in [41], an explicit expression for T_g can be obtained

[1] By Van Kessel, through personal communication.

from this energy balance. When using its non-linearized counterpart, this temperature is implicitly given and needs to be solved iteratively via some method that computes the zeros of a generic nonlinear function. It has been found that Newton's method (see *e.g.* [71]) effectively and efficiently solves the gas phase energy balance for the gas temperature T_g, with T_s as a suitable initial guess. Because of this and because of leading to a more accurate computation of T_g due to the usage of the full nonlinear version of the radiation term, it is proposed here to use the latter way of computing T_g.

Flue gas oxygen concentration

The oxygen concentration in the flue gas $Y_{O_2,fg}$, expressed in $[mol/m^3]$, is derived from a stationary oxygen mole balance over the furnace as

$$Y_{O_2,fg} = \left(\frac{\rho_{fg}}{\phi_{fg}}\right)\left(\left(\frac{\phi_{prim}+\phi_{sec}}{\rho_{air}}\right)Y_{O_2,in} - \left(\frac{\nu_{O_2} RaM_{comb}}{M_{O_2}}\right)\right) \quad (2.11)$$

with ρ_{fg} $[kg/m^3]$ the flue gas density and ρ_{air} $[kg/m^3]$ the ambient air density.

Dynamics of the steam system

The steam flow ϕ_{st} $[kg/s]$ is assumed to be produced by a natural convection boiler that includes an economiser and a superheater. This boiler is essentially a heat exchanger and transforms heat present in the flue gas into steam. Motivated by the work of [4], the boiler dynamics are described as

$$A_b \frac{d\phi_{st}}{dt} = Q + \phi_{fd}(H_{fd} - H_w) - \phi_{st} H_c \quad (2.12)$$

with A_b $[(J\ \ s)/kg]$ a constant that depends on, amongst others, boiler construction data, ϕ_{fd} $[kg/s]$ the feed water to the boiler, H_{fd} $[J/kg]$ its enthalpy, H_w $[J/kg]$ the enthalpy of the water in the boiler and H_c $[J/kg]$ its condensation enthalpy. Q $[J/s]$ represents the heat flux delivered by the flue gas to the boiler and is described as

$$Q = \eta \phi_{fg} C_{p,fg}(T_g - T_{g,out}) \quad (2.13)$$

with η $[-]$ representing the efficiency of the heat transfer from the flue gas to the water on the boiler side and $T_{g,out}$ $[K]$ the flue gas temperature at the outlet of the boiler. Typically, a measurement of the latter temperature is available at an MSWC plant. In case it is not, one can resort to a simple expression for this temperature such as *e.g.*

$$T_{g,out} = \alpha_{T_g} T_g \quad (2.14)$$

with α_{T_g} $[-]$ some scalar (< 1). It it noted that $H_{fd} \approx H_w$ and one, hence, can approximate the boiler dynamics by

$$A_b \frac{d\phi_{st}}{dt} = Q - \phi_{st} H_c \quad (2.15)$$

if ϕ_{fd} is not too large.

2.3.3 The calorific value sensor

Introduction

The basic idea behind the calorific value sensor can be stated as to, first, model the composition of the burning waste on the grate and, secondly, to compute the parameters that make up this composition model from given stationary component material balances over the gas phase for which all other parameters are known, either by assumption or via measurements. From the computed burning waste composition parameters, then, the corresponding calorific value is straightforwardly calculated using a well-known formula.

Several slightly different versions of the CVS exist. Here, a version is discussed which is equivalent to the original version of the CVS, as described in *e.g.* [41], but which explanation slightly differs. The gas phase mass balances required for this CVS are mass balances for O_2, CO_2 and N_2. Important measurements that are required to close these balances are flue gas component measurements for O_2, H_2O and CO_2[2].

The basic idea behind the CVS will now be explained in detail. For that purpose, first the required gas phase component mass balances are presented. Then it is discussed how the desired burning waste composition parameters are computed from these mass balances and, subsequently, how the calorific value for this waste is computed from these composition parameters. Finally, it is briefly shown how to compute the remaining mass balance parameters (*i.e.* those assumed known) from given measurements.

Gas phase component mass balances

For the derivation of the CVS it is again assumed that the gas phase of the MSWC plant is stationary, *i.e.* as was the case for the model discussed in section 2.3.2, and that the mass flows entering and leaving this gas phase are as depicted in figure 2.1. Here, X $[kg/kg]$ represents (again) mass fraction and ϕ $[kg/s]$ mass flow. More specific, for the derivation of the gas phase component mass balances and the resulting CVS it is assumed that

- four mass flows $\phi_{O_2,fg,in}$, $\phi_{CO_2,fg,in}$, $\phi_{H_2O,fg,in}$ and $\phi_{N_2,fg,in}$ enter the gas phase from the waste layer on the grate, which are the result of the combustion and evaporation process taking place in this layer.

- a secondary air flow ϕ_{sec} and air leakage flow ϕ_{leak} enter the gas phase while originating from the ambient air. These flows are assumed to consist of O_2, H_2O and N_2 (the amount of CO_2 in the ambient air is assumed to be negligibly small), as expressed by the mass fractions $X_{O_2,air}$, $X_{H_2O,air}$ and $X_{N_2,air}$.

- the flue gas ϕ_{fg} leaving the gas phase is assumed to contain O_2, CO_2, H_2O and N_2, as expressed by the mass fractions $X_{O_2,fg}$, $X_{CO_2,fg}$, $X_{H_2O,fg}$ and $X_{N_2,fg}$. This implies, amongst others, that the amount of CO in the flue gas is negligibly small *c.q.* that the combustion is assumed to be complete.

[2] It is noted that H_2O and CO_2 in the flue gas are not always measured at MSWC plants.

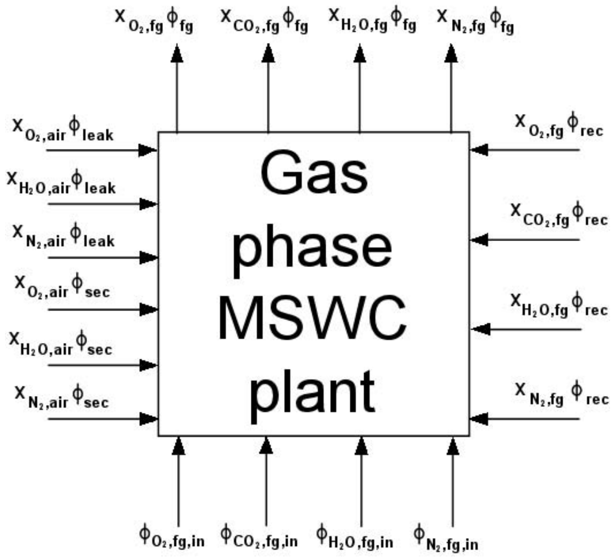

Figure 2.1: The mass flows entering and leaving the gas phase of an MSWC plant.

- a portion of ϕ_{fg} is assumed to be fed back to the gas phase as the recirculation flow ϕ_{rec} and this flow, therefore, is assumed to have the same composition as ϕ_{fg}. The point where ϕ_{rec} is tapped from ϕ_{fg} is assumed to lie outside the considered system boundary.

Under these assumptions, four component mass balances are readily derived for the gas phase, *i.e.* those for O_2, CO_2, H_2O and N_2. For the derivation of the CVS, however, only those for O_2, CO_2 and N_2 are needed:

$$X_{O_2,fg}(\phi_{fg} - \phi_{rec}) = \phi_{O_2,fg,in} + X_{O2,air}(\phi_{sec} + \phi_{leak}) \tag{2.16}$$

$$X_{CO_2,fg}(\phi_{fg} - \phi_{rec}) = \phi_{CO_2,fg,in} \tag{2.17}$$

$$X_{N_2,fg}(\phi_{fg} - \phi_{rec}) = \phi_{N_2,fg,in} + X_{N2,air}(\phi_{sec} + \phi_{leak}) \tag{2.18}$$

These balances form the basis for the CVS together with certain expressions for $\phi_{O_2,fg,in}$, $\phi_{CO_2,fg,in}$ and $\phi_{N_2,fg,in}$. These expressions are obtained under the following assumptions:

- the waste layer consists of three fractions: a moisture fraction, an inert fraction and a combustible fraction (Hence, in comparison to the literature model of section 2.3.2, the combined combustible and inert part has been split up). In addition, the combustible component is assumed to consist of CH_yO_z.
- the combustion reaction is given as

$$CH_yO_z + \tfrac{1}{2}(2\eta_C + \tfrac{1}{2}y - z)O_2 \longrightarrow (1-\eta_C)C + \eta_C CO_2 + \tfrac{1}{2}yH_2O \tag{2.19}$$

These assumptions allow the derivation of the following expressions for the mass flows $\phi_{O_2,fg,in}$, $\phi_{CO_2,fg,in}$ and $\phi_{N_2,fg,in}$ as a function of the waste flow $\phi_{CH_yO_z}$ that leaves the combustible part of the waste layer on the grate through combustion:

$$\phi_{O_2,fg,in} = X_{O_2,air}\phi_{prim} - \frac{1}{2}(2\eta_C + \frac{1}{2}y - z)(\frac{MM_{O_2}}{MM_{CH_yO_z}})\phi_{CH_yO_z} \quad (2.20)$$

$$\phi_{CO_2,fg,in} = \eta_C(\frac{MM_{CO_2}}{MM_{CH_yO_z}})\phi_{CH_yO_z} \quad (2.21)$$

$$\phi_{N_2,fg,in} = X_{N_2,air}\phi_{prim} \quad (2.22)$$

Here, MM_{O_2} represents the mole mass of O_2 ($[kg/mole]$), $MM_{CH_yO_z}$ that of CH_yO_z, etc. (the values of these mole masses can readily be obtained from the literature; also MM is used here to denote mole mass to distinguish it from the notation used here for mass, which is M).

The assumptions stated just above also imply the gas phase component mass balance for H_2O to be given as

$$X_{H_2O,fg}(\phi_{fg} - \phi_{rec}) = \phi_{H_2O,fg,in} + X_{H2O,air}(\phi_{sec} + \phi_{leak}) \quad (2.23)$$

with

$$\phi_{H_2O,fg,in} = X_{H_2O,air}\phi_{prim} + \phi_{evap} + \frac{1}{2}y(\frac{MM_{H_2O}}{MM_{CH_yO_z}})\phi_{CH_yO_z} \quad (2.24)$$

where ϕ_{evap} is the amount of H_2O flowing from the waste layer due to evaporation. Also, the assumptions imply the following expression for the char flow ϕ_C:

$$\phi_C = (1 - \eta_C)(\frac{MM_C}{MM_{CH_yO_z}})\phi_{CH_yO_z} \quad (2.25)$$

This char flow leaves the grate in solid form as bottom and fly ash. This H_2O mass balance and this expression for ϕ_C are of no use for the derivation of the CVS. These are of relevance, though, for the new model to be proposed here, which is the reason for stating these here.

The calorific value sensor

The aim of the CVS is to, first, compute the composition parameter z of the burning CH_yO_z part of the waste on the grate via the mass balances (2.16) - (2.22) and, then, to compute the calorific value of this waste, denoted as $\Delta H_{CH_yO_z}$, from this composition parameter using a well-known specific formula.

To be short, the estimator for z is derived from eqns. (2.16) - (2.18) and (2.20) - (2.22) as

$$z = 2\eta_C + \frac{1}{2}y - 2\eta_C((\frac{X_{N_2,fg}}{X_{CO_2,fg}})(\frac{X_{O_2,air}}{X_{N_2,air}}) - (\frac{X_{O_2,fg}}{X_{CO_2,fg}}))(\frac{MM_{CO_2}}{MM_{O_2}}) \qquad (2.26)$$

and the estimate for $\Delta H_{CH_yO_z}$ $[J/kg]$ is computed from this estimate for z via the so called formula of Michel [41] as

$$\Delta H_{CH_yO_z} = \frac{(408.4 + 102.4y - 156.8z)}{MM_{CH_yO_z}} 10^{-3} \qquad (2.27)$$

Note that the expression for z does not depend on any mass flow, in particular not on $\phi_{CH_yO_z}$, ϕ_{evap}, ϕ_{rec} and ϕ_{leak}. Also, note that the estimates for z and $\Delta H_{CH_yO_z}$ require values for

- the flue gas mass fractions $X_{O_2,fg}$, $X_{CO_2,fg}$ and $X_{N_2,fg}$
- the ambient air mass fractions $X_{O_2,air}$ and $X_{N_2,air}$
- η_C and y

The values for $X_{O_2,fg}$, $X_{CO_2,fg}$ and $X_{N_2,fg}$ can be obtained from measurements of corresponding molar volumes $Y_{O_2,fg}$, $Y_{CO_2,fg}$ and $Y_{H_2O,fg}$, corresponding mole fractions or, most common, corresponding volume fractions, and standard equations. The values for $X_{O_2,air}$ and $X_{N_2,air}$ are obtained from measurements of the ambient air temperature, T_{air} $[K]$, and ambient air relative humidity, RH_{air} $[\%]$, and also standard equations. The values for η_C and y, however, are not measured and have to be specified *a priori* by assumption. The value for η_C typically lies close to 1. The value for y is sometimes estimated via an off-line analysis of the waste composition. Such analyses for different types of solid fuels have shown the value for y to lie approximately within the range of 1.6 to 1.8, with 1.72 being a good estimate. Fortunately, the CVS has been shown to be relatively insensitive to errors in y and η_C [41].

There is more to the CVS than presented here. For example, in [41, 106] more quantities are estimated than only z and $\Delta H_{CH_yO_z}$. In particular, the CVS typically also takes the water content of the burning waste into account to deliver a water compensated estimate for the calorific value.

The references [41, 106] also discuss applications of the CVS to data from large scale MSWC plants. From these applications it was concluded that the CVS is capable of continuously monitoring the calorific value of the waste being converted.

2.3.4 The new model

In this section, the equations are specified of the new first-principles model proposed here for model based MSWC plant combustion control applications. This model is obtained through integrating the set of CVS equations

- the gas phase component mass balances given by eqns. (2.16) - (2.18) and (2.20) - (2.24)
- the expression for the char flow ϕ_C given by eqn. (2.25)
- the calorific value $\Delta H_{CH_yO_z}$ and its expression as a function of the waste composition parameters y and z as given by Michels formula (2.27)

and, also, air leakage flow ϕ_{leak} and recirculation flow ϕ_{rec} into the literature model of section 2.3.2. A main characteristic of the resulting model is that it contains z as a disturbance input. Together with estimates for this variable obtained with the CVS, this allows for improved simulation capabilities of the model, including more accurate simulation of important control relevant variables such as steam production and oxygen concentration, and, thereby, also for improved model validation results and an improved model based MSWC plant combustion control performance. No quantification of the improvements that can be made by including z as a disturbance input is made here. This is subject of future research.

Integration of the mentioned CVS equations into the literature model of section 2.3.2 is rather straightforward, though requires a number of modifications to be made to this model. A first such modification is the division of the mass balance (2.1) for the combustible and inert waste layer part M_{comb} into a separate mass balance for the combustible part $M_{CH_yO_z}$ $[kg]$

$$\frac{dM_{CH_yO_z}}{dt} = X_{CH_yO_z}\phi_{w,in} - RaM_{CH_yO_z} \tag{2.28}$$

and a separate mass balance for the inert part M_{inert} $[kg]$

$$\frac{dM_{inert}}{dt} = X_{inert}\phi_{w,in} - \phi_{out} \tag{2.29}$$

Note that the way to model ϕ_{out} proposed in [41] as

$$\phi_{out} = X_{inert}\phi_{w,in} \tag{2.30}$$

leads to a model where the inert part of the waste M_{inert} does not change with time, *i.e.* $\frac{M_{inert}}{dt} = 0$. Hence, when applying this substitution, one may as well remove the inert mass balance (2.29) from the model and replace M_{inert} in the energy balance (2.32) by a constant parameter of the same name. Another substitution for ϕ_{out} that is suggested here, one that removes the restrictive assumption of a constant M_{inert}, is to express ϕ_{out} as

$$\phi_{out} = C_{\phi_{out}}\ a\ M_{inert} \tag{2.31}$$

with $C_{\phi_{out}}$ a positive constant. The first motivation for using this expression is that it reflects the effect that a larger (smaller) M_{inert} renders a larger (smaller) ϕ_{out}. A second motivation is that it reflects the effect that a larger (smaller) grate speed renders a larger (smaller) ϕ_{out}, where it is assumed that a is a monotone increasing function of the grate speed MV. As will be shown later on in this thesis, this assumption is a valid one.

The moisture mass balance (2.2) of the literature model of section 2.3.2 also holds for the new model proposed here. Also, the reaction rate R and corresponding parameter are computed as in this model (see eqns. (2.6) and (2.8)).

With the introduction of $M_{CH_yO_z}$ and M_{inert} as separate waste layer masses, the waste layer energy balance (2.4) of the literature model of section 2.3.2 is modified accordingly to become

$$\begin{aligned}(C_{p,mois}M_{mois} + C_{p,CH_yO_z}M_{CH_yO_z} + C_{p,inert}M_{inert})\tfrac{dT_s}{dt} = \\ (C_{p,mois}X_{mois} + C_{p,CH_yO_z}X_{CH_yO_z} + C_{p,inert}X_{inert})\phi_{w,in}(T_{w,in} - T_s) + \\ C_{p,fg}\phi_{prim}(T_{prim} - T_s) + X_{\Delta H_{CH_yO_z}}RaM_{CH_yO_z}\Delta H_{CH_yO_z} + \\ k_{evap}aM_{mois}\Delta H_{evap} + \sigma A_s(\varepsilon_{gt}T_g^4 - \alpha_{gt}T_s^4)\end{aligned} \tag{2.32}$$

The mass flows entering the gas phase from the waste layer, *i.e.* $\phi_{O_2,fg,in}$, $\phi_{CO_2,fg,in}$, $\phi_{H_2O,fg,in}$ and $\phi_{N_2,fg,in}$, are computed according to eqns. (2.20) - (2.22) and (2.24), with $\phi_{CH_yO_z}$ and ϕ_{evap} computed as

$$\begin{aligned}\phi_{CH_yO_z} &= RaM_{CH_yO_z} \\ \phi_{evap} &= k_{evap}aM_{mois}\end{aligned} \tag{2.33}$$

The gas phase is assumed stationary again with its mass balance given as

$$\phi_{fg} = \phi_{fg,in} + \phi_{sec} + \phi_{rec} + \phi_{leak} \tag{2.34}$$

where

$$\phi_{fg,in} = \phi_{O_2,fg,in} + \phi_{CO_2,fg,in} + \phi_{H_2O,fg,in} + \phi_{N_2,fg,in} \tag{2.35}$$

The recirculation flow ϕ_{rec} in (2.34) may be measured or may have to be modeled. The latter may result *e.g.* in an expression

$$\phi_{rec} = \alpha_{rec}\phi_{fg} \tag{2.36}$$

with α_{rec} representing a (mass) fraction that is a function of an MV to control this flow. In that case,

$$\phi_{fg} = (\frac{1}{1-\alpha_{rec}})(\phi_{fg,in} + \phi_{sec} + \phi_{leak}) \tag{2.37}$$

with α_{rec} computed from the corresponding MV value.

The flue gas flow *after* the point where ϕ_{rec} is drawn from this flow (note that ϕ_{fg} represents the flue gas flow *before* this point) is given as

$$\phi'_{fg} = \phi_{fg,in} + \phi_{sec} + \phi_{leak} \tag{2.38}$$

As ϕ_{fg} and ϕ'_{fg} have different values for $\alpha_{rec} > 0$ it is, in order to properly compute the steam flow ϕ_{st}, important to determine which of these two flows actually passes the boiler. Here, it is assumed that ϕ_{fg} represents the flue gas flow passing the boiler.

With $\phi_{O_2,fg,in}$, $\phi_{CO_2,fg,in}$, $\phi_{H_2O,fg,in}$, $\phi_{N_2,fg,in}$ and ϕ_{fg} known quantities, the mass fractions in the flue gas are readily computed from the gas phase component mass balances (2.16) - (2.18) and (2.23) as:

$$X_{O_2,fg} = \frac{\phi_{O_2,fg,in} + X_{O2,air}(\phi_{sec} + \phi_{leak})}{\phi_{fg} - \phi_{rec}} \tag{2.39}$$

$$X_{CO_2,fg} = \frac{\phi_{CO_2,fg,in}}{\phi_{fg} - \phi_{rec}} \tag{2.40}$$

$$X_{H_2O,fg} = \frac{\phi_{H_2O,fg,in} + X_{H2O,air}(\phi_{sec} + \phi_{leak})}{\phi_{fg} - \phi_{rec}} \tag{2.41}$$

$$X_{N_2,fg} = \frac{\phi_{N_2,fg,in} + X_{N2,air}(\phi_{sec} + \phi_{leak})}{\phi_{fg} - \phi_{rec}} \tag{2.42}$$

These mass fractions are easily transformed into quantities that are actually measured at MSWC plants such as *e.g.* molar volumes $Y_{O_2,fg}$, $Y_{CO_2,fg}$, $Y_{H_2O,fg}$ and $Y_{N_2,fg}$, or molar or, most common, volume fractions. Note that now not only O_2 is a model output but also CO_2, H_2O and N_2.

By integrating eqn. (2.25) of the CVS equations into the new model here, this model also allows for the computation of the char flow ϕ_C.

Using the assumption that $C_{p,fg,in} = C_{p,fg}$, the gas phase energy balance for the new MSWC model is derived as

$$\begin{aligned} 0 \quad = \quad & C_{p,fg}\phi_{fg,in}(T_s - T_g) + C_{p,fg}\phi_{sec}(T_{sec} - T_g) + \\ & C_{p,fg}\phi_{rec}(T_{rec} - T_g) + C_{p,fg}\phi_{leak}(T_{leak} - T_g) + \\ & (1 - X_{\Delta H_{CH_yO_z}})RaM_{CH_yO_z}\Delta H_{CH_yO_z} + \\ & \sigma A_s(\alpha_{gt}T_s^4 - \varepsilon_{gt}T_g^4) \end{aligned} \tag{2.43}$$

which can, again, effectively and efficiently be solved for T_g by means of Newton's method. Note that in the energy balances (2.32) and (2.43), as in the literature model of section 2.3.2, the temperatures $T_{w,in}$, T_{prim} and T_{sec} may be replaced for the ambient air temperature T_{air} without introducing a significant error. This also can be done for T_{leak}. Depending on the location along the furnace where the recirculation flow is tapped from, T_{rec} may be replaced for *e.g.* $T_{g,out}$ or T_g. In the latter case, the corresponding convective energy term, the third one in (2.43), is removed.

Both in the gas phase energy balance (2.43) and in the waste layer energy balance (2.32), the heat capacity of the air $C_{p,air}$ is approximated, as in the literature model of section 2.3.2, by that of the flue gas $C_{p,fg}$. Vice versa, $C_{p,air}$ may be used as an approximation to $C_{p,fg}$. In contrast to the model of section 2.3.2, however, $C_{p,fg}$ may now also be approximated as a weighted sum of the gas phase component heat capacities, thereby exploiting the accurate knowledge available in the new MSWC model on the flue gas composition and the fact that the heat capacity of a gas mixture can be approximated as a mole fraction weighted average of its compononent heat capacities when the latter are expressed in $[J/mol\ K]$ (rather than $[J/kg\ K]$) [1]. Additionally, the gas component heat capacities can be obtained as polynomial functions of the gas phase temperature T_g, with these functions derived from literature (*e.g.* [1]). In order to reduce the computational complexity here, though at the cost of accuracy, one may choose to change this dependency on T_g into one on T_s.

The equations describing the steam system (2.12) - (2.14) require no modification and, therefore, are directly used to describe this system in the new MSWC plant model proposed here.

2.3.5 A system theoretic investigation of the new model

Introduction

The new model proposed above is investigated here from a system theoretic point of view. In particular, the structure and two important control relevant system theoretic properties of the model are investigated: controllability and observability. The main reasons for that are (i) to disclose fundamental control relevant properties of the model, (ii) to find out whether the computational complexity of the model can be reduced and (iii), by introducing a specific model format, to ease the explanation of control theoretic concepts later on in this thesis. The reduction of the computational complexity of the model is beneficial for model based control design *e.g.* because this typically results in a lower computational and implementational complexity of the resulting control system.

Model structure

The new model proposed above is to be used for (model based) control and, thereby, must be able to simulate control relevant variables *c.q.* controlled variables (CVs) as a function of the variables that can be manipulated *c.q.* the manipulated variables (MVs). In that respect, it is important to note that equations describing the actuator and sensor dynamics still may have to be added to the new model proposed here. In particular, if the model is to be truly useful for control design, equations have to be added that describe

- the relation between the ram speed MV and $\phi_{w,in}$
- the relation between the grate speed MV and a

It is noted that the second, non-obvious, one of these relations, has been discovered during model validation exercises to be discussed later in this thesis. There, a linear relationship between the grate speed MV and a has been found to suffice.

The form of the new model considered in this thesis is one where the two actuator relationships discussed above have been added while no other sensor and actuator dynamics have been added as these are assumed to be negligible. The resulting model can be shown to be of the differential-algebraic equation(s) (DAE) format

$$\frac{dx_1(t)}{dt} = f(x_1(t), x_2(t), u(t), d_m(t), d_{nm}(t), \theta)$$

$$
\begin{aligned}
0 &= g(x_1(t), x_2(t), u(t), d_m(t), d_{nm}(t), \theta) \\
y(t) &= h(x_1(t), x_2(t), u(t), d_m(t), d_{nm}(t), \theta)
\end{aligned}
\tag{2.44}
$$

with

- $u(t)$ a column vector containing the manipulated variables (MVs) of the MSWC plant: ram speed MV, grate speed MV, ϕ_{prim}, ϕ_{sec} and ϕ_{rec}
- $y(t)$ a column vector containing the controlled variables (CVs) of the MSWC plant: (in particular) ϕ_{steam} and O_2 expressed in volume fraction.
- $x_1(t)$ a column vector containing the differential states: $M_{CH_yO_z}$, M_{inert}, M_{mois}, T_s, ϕ_{st} and, possibly, differential states related to the added actuator dynamics.
- $x_2(t)$ a column vector containing the algebraic states, which here is only T_g
- $d_m(t)$ a column vector containing the measured disturbances: *e.g.* T_{air}
- $d_{nm}(t)$ a column vector containing the nonmeasured disturbances: *e.g.* $X_{CH_yO_z}$, X_{inert} and X_{mois}.
- θ a column vector with the remaining, constant model parameters such as *e.g.* ΔH_{evap}.

Although actually being of the format (2.44), the new model proposed here may effectively also be viewed as being of the nonlinear state-space format

$$
\begin{aligned}
\frac{dx(t)}{dt} &= f(x(t), u(t), d_m(t), d_{nm}(t), \theta) \\
y(t) &= h(x(t), u(t), d_m(t), d_{nm}(t), \theta)
\end{aligned}
\tag{2.45}
$$

where no algebraic states x_2 are present and only differential states $x_1(t) \to x(t)$ are present. This view is allowed when the model (2.44) is integrated over time through an implicit model structure of the form of (2.45). This is the case when the algebraic equations $g(x_1, x_2, ...) = 0$ are first solved, in an inner loop, for x_2 and subsequently the resulting value for x_2 is substituted in the remaining equations $f(.)$ and $h(.)$ which then, in an outer loop, are integrated over time as being of the form of (2.45). In this thesis, the new MSWC plant model is assumed to be solved in the latter manner, with $x_2 = T_g$ being effectively and efficiently solved in the inner loop by means of Newton's method (and the resulting outer loop model equations being integrated with a fourth-order Runge-Kutta method). Because of that, the model format (2.45) is used in the remainder of this thesis when discussing control theoretic concepts.

Controllability

Roughly stated, controllability is a system property that quantifies the extent with which this system can be moved around. (See *e.g.* [93] for a formal definition). Typically, an investigation into the controllability of a system results in a yes/no answer to the question whether, with the given set of inputs u, some *a priori* chosen set of (differential) states x (state controllability) or outputs y (output or input-output controllability) can be moved from any initial set of values to any final set of values within any given finite time. Many ways of determining the controllability of a system are available. Also, controllability determination techniques are available for both linear and nonlinear systems.

Investigating the controllability of a plant (model) is used for *e.g.* actuator location and model reduction. In the first case, a controllability test is used *e.g.* to determine whether the given set of MVs provides enough controllability, in which case new MVs may be added if possible. Likewise, redundant actuators may be identified and removed. In the second case, (differential) states or combinations of states may be identified that are uncontrollable and removed from the model because of not providing any contribution to the model part of interest for the application while unncessesarily increasing the complexity of the model.

Here, (only) the question is addressed whether the new MSWC plant model proposed here is controllable for the set of inputs $u = \{\ \phi_{w,in},\ a,\ \phi_{prim},\ \phi_{sec}$ and $\phi_{rec}\ \}$ and the two main CVs $y = \{\ \phi_{st},\ O_2\ \}$, providing an answer to the practically relevant question whether the latter two CVs can actually be moved from any initial set of values to any final set of values withing any given finite time. To answer this question, the notion of output controllability is used. More specific, a measure of output controllability for linear systems is applied to linearizations of the new model proposed above at many operating points, hence using a linear controllability measure for investigating the controllability of a nonlinear system. The output controllability measure used here is a simple extension of a commonly used measure for linear state controllability and consists of testing whether the following rank condition is fulfilled:

$$rank(\mathcal{C}_y) \quad = \quad m \tag{2.46}$$

with $m = 2$ the number of outputs y and

$$\mathcal{C}_y \quad = \quad [CB \quad CAB \quad CA^2B \quad ... \quad CA^{n-1}B \quad D] \tag{2.47}$$

where $n = 5$ is the number of (differential) states x and where A, B, C and D are the Jacobians resulting from linearizing the new MSWC plant model with respect to x and u: $A = \frac{\partial f(.)}{\partial x}$, $B = \frac{\partial f(.)}{\partial u}$, $C = \frac{\partial h(.)}{\partial x}$, $D = \frac{\partial h(.)}{\partial u}$. From this testing it followed that the new MSWC plant model is controllable for the given set of outputs, thereby indicating the absence of any controllability problems in the most common practical situation where ϕ_{st} and O_2 are the only two CVs.

Observability

Roughly stated, observability is a system property that quantifies the extent with which information on the state values x is present in the outputs y. (See *e.g.* [93] for a formal definition). Typically, an investigation into the observability of a system results in a yes/no answer to the question whether state values x at some time instant can be inferred from finite time records available for y (and u) after this time instant. Several ways of determining the observability of a system are available. Also, observability determination techniques are available for both linear and nonlinear systems.

Investigating the observability of a plant (model) is used for *e.g.* sensor location and model reduction. In the first case, observability tests are used to *e.g.* add sensors *c.q.* outputs y to allow for inference of values of the states of interest which otherwise could not be inferred. In the second case, these tests are used to identify states or combinations of states that are unobservable and to remove these from the model because of not providing any contribution to the model part of interest for the application while unncessesarily increasing the complexity of the model.

Here, a specific interesting model reduction related result is discussed that followed from analyzing the observability of the new MSWC plant model proposed here. More specific, observability tests on this model for different sets of measurements y revealed that under normal combustion conditions M_{inert} is virtually unobservable for *any* practically typically available such set of measurements, more specific for any set of measurements not containing ϕ_{out} (see eqn. (2.31)).

The observability tests were performed through a standard linear observability test on linearized versions of the new MSWC plant model proposed here obtained at many distinct operating points. The standard linear observability test refered to here is to check whether (see *e.g.* [93])

$$rank(\mathcal{O}) \quad = \quad n \tag{2.48}$$

where $n = 5$ is the number of states x, with

$$\mathcal{O} \quad = \quad \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \tag{2.49}$$

and where A and C are the Jacobians resulting from linearizing the new MSWC plant model with respect to x: $A = \frac{\partial f(.)}{\partial x}$ and $C = \frac{\partial h(.)}{\partial x}$.

The unobservability of M_{inert} for any practically typically available set of measurements under normal combustion conditions can be explained by taking a closer look at the equations of the new model proposed here. More specific, such a closer look reveals that (i) M_{inert} affects all model variables other than ϕ_{out}, if computed according to eqn. (2.31), through T_s, *i.e.* through the waste layer energy balance (2.32), and (ii) the influence of M_{inert} on T_s is negligibly small under normal combustion conditions. This second fact can be illustrated by means of a simulation where the inert mass balance (2.29) is excited by an additional random perturbation and all remaining external variables (*e.g.* $\phi_{w,in}$) are held constant. See figure 2.2. Note from this figure that T_s shows no (visible) change at all while M_{inert} exhibits a large variation. Only under extreme conditions, M_{inert} has a noticable effect on T_s. From closer investigation of the waste layer energy balance (2.32) these extreme conditions can be shown to be those where $M_{inert} >> M_{mois}$ and $M_{inert} >> M_{CH_yO_z}$ and, simultaneously, $M_{inert} \downarrow 0$ (which represents the situation where only a tiny inert mass is present in the furnace). In order to see that, note that the waste layer energy balance (2.32) is of the form

$$\frac{dT_s}{dt} = \frac{f(M_{mois}, M_{CH_yO_z})}{(C_{p,mois}M_{mois} + C_{p,CH_yO_z}M_{CH_yO_z} + C_{p,inert}M_{inert})} \tag{2.50}$$

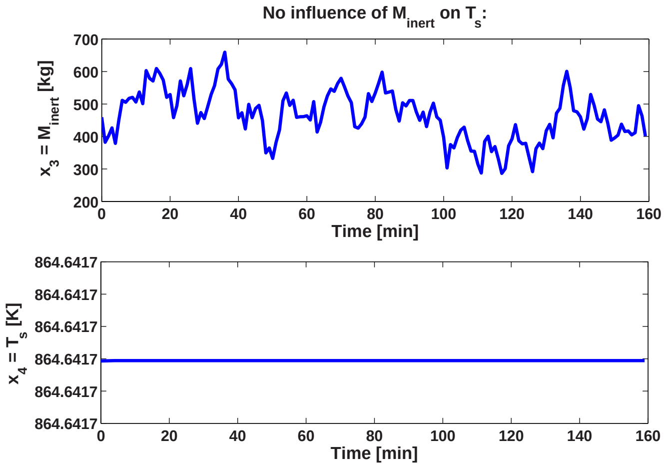

Figure 2.2: M_{inert} and T_s for a simulation with the new MSWC plant model with a random perturbation on the inert mass balance (2.29).

Hence, under the mentioned extreme conditions, the M_{inert}-term in the denominator of (2.50) dominates over the other terms in this denominator and causes a large amplification when $M_{inert} \downarrow 0$.

Summarizing, under normal combustion conditions M_{inert} has no significant effect on any other model variable except for ϕ_{out}. As the latter flow generally is not measured in practice, this implies unobservability of M_{inert} for any practically typically available set of measurements under normal combustion conditions.

The facts that (i) M_{inert} and ϕ_{out} are typically not of interest for MSWC plant combustion control and that (ii) M_{inert} is (almost) unobservable under normal combustion conditions imply that these variables can be removed from the new MSWC plant model proposed here when used under such conditions. This has the advantage of a computationally less complex model and thereby of a computationally and implementationally less complex model based controller. The removal of M_{inert} and ϕ_{out} can be done in two ways: either through the complete removal of the model equations describing M_{inert} and ϕ_{out} or the replacement of the inert mass balance (2.29) by a computationally more cheap but less accurate approximation. In this thesis, the second option has been chosen with the inert mass balance (2.29) replaced by its steady-state version

$$M_{inert} = \frac{X_{inert}\phi_{w,in}}{C_{\phi_{out}}\ a} \tag{2.51}$$

(which is called *residualization* [93]). This choice was arbitrary because no significant difference in gain in computation time has been observed between the two mentioned ways of model reduction. The only advantage of the chosen way of reduction is that it provides a rough estimate of M_{inert}.

By applying the replacement of the inert mass balance (2.29) by its steady-state solution, a reduction in simulation time of approximately 2 to 3 $\%$ has been observed, which is not spectacular but still may lead to significant improvements for *e.g.* an NMPC control strategy (which performance may depend on the model simulation time). A simple simulation with the resulting reduced order version of the new model shows that only the computation of M_{inert} and ϕ_{out} changes significantly through this replacement but not the computation of the remaining model variables. See *e.g.* figure 2.3, where one can also see that replacing (2.29) for (2.51) has no noticeable influence on the prediction of T_s, and hence the remaining MSWC plant variables, for a step change on $\phi_{w,in}$.

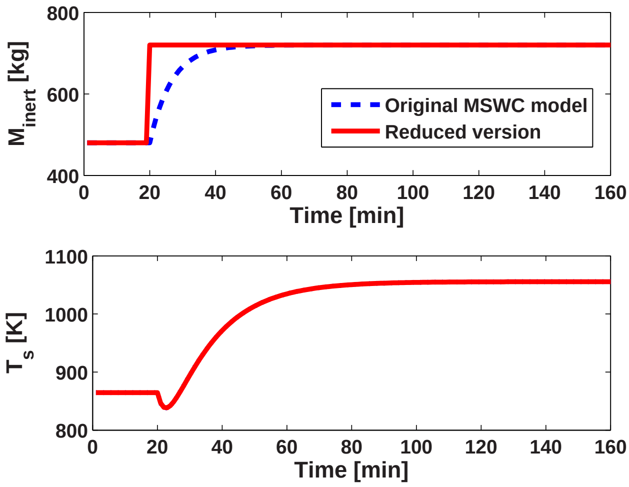

Figure 2.3: Simulation with full order version (blue, dashed line) and reduced order version (red, solid line) of the new model proposed here for a step change on $\phi_{w,in}$.

2.4 Conclusions

In this chapter, opportunities of first-principles modeling have been explored for obtaining a model suitable for model based MSWC plant combustion control. In particular, opportunities concerning the actual modeling part have been explored whereas model quality aspects are discussed in a later chapter. It has been found that the literature provides ample opportunities for first-principles MSWC plant models suitable for model based combustion control. Being the main contribution of this chapter, one of the available literature models has been selected because of its simplicitly and accuracy (to be discussed in more detail later on in this thesis) and extended to allow for an improved simulation, validation and model based combustion control performance compared to existing models. This extension involves the incorporation of the equations underlying the so called calorific value sensor, which is an on-line estimator of the MSWC plant waste composition and calorific value. This leads to a more detailed description of the waste composition in the model which, combined with the ability to estimate the main parameter of this description from large scale MSWC plant plant data through the CVS, allows for the mentioned improvements. The determination of the actual improvements is not carried out in this thesis but subject of future work. Also, system theoretic properties of the new model have been investigated with the main conclusion being that the new model can be reduced in model complexity when used to simulate normal MSWC plant combustion conditions. This reduction is advantageous from a model based control point of view as a computationally lower complex model typically results in a lower computational and implementational complexity for the resulting control system.

Chapter 3

A system identification methodology for MSWC plants

3.1 Introduction

In this chapter, results are presented of the exploration of the opportunities of linear system identification for obtaining a model suitable for model based MSWC plant combustion control, *i.e.* a sufficiently accurate low order linear black-box model of the furnace and boiler part of an MSWC plant. More specific, issues *c.q.* potential obstacles for arriving at such a model are identified together with solutions to overcome these obstacles. Subsequently, from the identified solutions a methodology is derived that is claimed to be suitable for obtaining a model suitable for model based MSWC plant combustion control. This methodology is the main contribution of this chapter.

Most of the issues to be resolved for proper system identification of MSWC plants (for reasons of space, from now on the adjective *linear* is left out) are also encountered at the identification of other industrial processes. As a consequence *c.q.* advantage, solutions to these issues are readily available in the corresponding literature. This fact is exploited here in the sense that most of the solutions in the MSWC plant system identification methodology proposed here are obtained from this *process identification* literature. Some issues relevant for proper system identification of MSWC plants are, however, not explicitly dealt with in this literature and are resolved here. In particular, one may encounter the so called *partial closed-loop* experimental setting at MSWC plants. The issue of how to properly handle data experimentally obtained in such a setting has not been resolved explicitly in the literature before and is resolved here. This is another major contribution of this chapter.

It is noted that this chapter does not contain yet a full assessment of linear system identification as a tool for obtaining a model suitable for model based MSWC plant combustion control. This is defered to the next chapter.

The contents of this chapter are as follows. First, in section 3.2, system identification preliminaries are discussed that are relevant for the remainder of this chapter. This includes a general introduction into system identification and a more detailed elaboration of the so called prediction error method, which is one of the main system identification methods. Then, in section 3.3, elements of process identification are discussed that are relevant for the MSWC plant system identification methodology proposed here. Here, issues are identified from the available literature that need to be resolved for proper identification of industrial processes together with solutions for these issues. The contents of this section also represent a brief overview of the state-of-the-art of process identification. Subsequently, in sections 3.4 and 3.5, issues are discussed that are particularly relevant for MSWC plant system identfication but which are not explicitly dealt with in the process identification literature, together with solutions proposed to overcome these issues. This includes the partial closed-loop identification issue refered to above. After that, in section 3.6, the MSWC plant system identification methodology proposed here is outlined. Finally, in section 3.7, the conclusions of this chapter are given.

3.2 System identification preliminaries

3.2.1 Introduction

In this section, elements of system identification are discussed that are relevant for the discussion in the remainder of this thesis. First of all, the basic elements of any system identification method are outlined. Then, the prediction error method is discussed in detail, which is one of the main system identification methods. This discussion includes the aspects of model structure selection, linear regression model estimation, model validation, bias and variance properties and closed-loop identification.

3.2.2 Basic elements of system identification

System identification [64, 122, 123] is an empirical/experimental modeling technique where parameters of a set of *a priori* postulated physically meaningless model equations, *i.e.* a so called black-box model, are fitted to experimentally obtained data. System identification consists globally of the following three steps:

1. data acquisition
2. estimation of the model parameters
3. validation of the estimated model

Data acquisition is performed by means of one or more experiments applied to the plant to be modeled during which this plant is perturbed by means of user-defined excitation signals, also denoted as *test* signals. These identification experiments may be split up in preliminary and final data acquisition experiments where the first set of experiments is used to optimally design the test signals to be applied during the second one. The final data acquisition experiment then delivers the data actually used for estimation and, also, data used for validation. The process variables that are excited correspond to the inputs of the resulting model which, in case the purpose of the model is control design, are equal to the plant MVs. At the same time a number of other process variables are measured that correspond to the outputs of the resulting black-box model. In case this model is used for control design, these are equal to the CVs of the plant.

In the estimation phase, first a model structure is chosen, *i.e.* type and number of model equations and their parameterization, *i.e.* the way the parameters are contained in these model equations. After that, the parameters of the model equations are fitted to the experimentally obtained data. The methods that are used to perform this fit can be divided roughly in two groups: *prediction error methods* (PEM) [64, 97] and, here so called, *realization based methods*. With the PEM one aims to minimize the difference between measured and predicted model outputs, which (indeed) is denoted as the prediction error. The realization based methods are state space model based methods that employ linear algebra techniques, in particular projection techniques and the singular value decomposition, to obtain the parameter estimates. The latter group largely consists of so called *subspace methods* [109].

Validation refers to the process of determining the degree to which the model is an accurate description of the real plant dynamics from the perspective of the intended use(s) of this model. A typical validation method is the comparison of simulated outputs with their measured counterparts, with the observed gap being a measure for the model quality. Another typical method used for validation of estimated models employs correlation tests on model inputs and residuals. Preferably, to allow for an objective validation, an experimentally obtained data set is used for validation that has not been used for estimation.

System identification has been applied for the modeling of many systems, amongst which but certainly not only industrial processes. Mostly the aim of system identification has been control design. It has long been subject of academic research, which has disclosed many theoretical properties of this way of modeling, in particular with respect to the accuracy of the resulting models. From the latter point of view, aspects that have received much attention are model *bias*, *i.e.* model error of a persistent *c.q.* systematic character, and (asymptotic) *variance*, *i.e.* model error of an accidental character.

Because of its relevance for the discussion in the remainder of this chapter, the PEM is discussed now in more detail, thereby also illuminating aspects of system identification discussed in this section.

3.2.3 The prediction error method

The basic idea

The basis for the PEM is, as for any system identification method, a data set containing sampled in- and output data that are experimentally obtained from the process or system to be modeled. Such a data set is denoted as

$$Z^N = \{u(1),\ y(1),\ \dots,u(N),\ y(N)\} \tag{3.1}$$

with $u(t)$, $t = 1, ..., N$, an m-dimensional column vector containing the input signals at sample instant t and $y(t)$ a p-dimensional column vector containing the output signals at this sample instant.

In order to obtain a model for the process that has produced the data set Z^N, first an assumption is made on what the real dynamics of this process actually looks like. In terms of the PEM one makes an assumption on the so called *true system*, denoted as $\mathcal{S}$, that produces the data set. The PEM generally assumes this system $\mathcal{S}$ to be causal, linear, time invariant (LTI) and to be given as

$$y(t) = G_o(q)u(t) + v(t) \qquad v(t) = H_o(q)e(t) \tag{3.2}$$

where $\{e(t)\}$ $(p \times 1)$ is a zero mean white noise process with covariance matrix Λ_o and bounded moments of order 4 + δ, for some $\delta > 0$, [28] and $H_o(q)$ is an inversely stable, monic filter. Both $G_o(q)$ and $H_o(q)$ represent discrete-time transfer function matrices of dimension $p \times m$ resp. $p \times p$ with q being the (forward) shift operator which is defined as $qs(t) = s(t+1)$ and $q^{-1}s(t) = s(t-1)$, with $s(t)$ an arbitrary discrete-time signal. The transfer function matrix $G_o(q)$ is often called the *plant* dynamics whereas $H_o(q)$ is called the *disturbance* dynamics. The latter is inversely stable meaning that its inverse is stable. The monicity of this filter means that $H_o(0) = I_p$ with I_p representing a unity matrix of dimension $p \times p$. The assumption on the true *c.q.* data generating system is schematically depicted in figure 3.1.

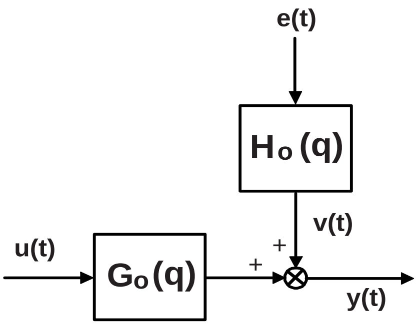

Figure 3.1: The true c.q. data generating system according to the PEM.

Having made an assumption on the true plant and disturbance dynamics, one has a means of modeling these dynamics. When one employs the assumption on $\mathcal{S}$ given by eqn. (3.2), a logical choice for the structure of the model to be estimated, which in the PEM literature is denoted by $\mathcal{M}$, is

$$y(t, \theta) = G(q, \theta)u(t) + H(q, \theta)e(t) \tag{3.3}$$

with both $G(q, \theta)$ and $H(q, \theta)$ LTI transfer function matrices similar to their true counterparts $G_o(q)$ resp. $H_o(q)$, but now parameterized by parameters contained in the parameter vector θ. Modeling according to the PEM corresponds to estimating the values for this parameter vector. For that purpose, the parameterized model given by eqn. (3.3) is used to form a so called *prediction* model, *i.e.* a model that predicts the outputs $y(t)$ one or more (sample) time steps ahead in the future. Typically, a *one-step-ahead* prediction model of the outputs is used. For a model of the form (3.3) it can be shown that the optimal one-step-ahead predictor for $y(t)$ is equal to

$$\hat{y}(t,\theta) = [I_p - H^{-1}(q,\theta)]y(t) + H^{-1}(q,\theta)G(q,\theta)u(t) \tag{3.4}$$

Here, $\hat{.}$ is used to denote an estimate or, as in this case, prediction. With the predictor (3.4), the (one-step-ahead) *prediction error* is defined as

$$\epsilon(t,\theta) = y(t) - \hat{y}(t,\theta) = H^{-1}(q,\theta)[y(t) - G(q,\theta)u(t)] \tag{3.5}$$

Typically, the parameters in θ are obtained by minimizing a scalar valued criterion that is a function of the prediction error, using some numerical optimization routine:

$$\hat{\theta}_N = \min_{\theta}\ V_N(\theta) = \min_{\theta}\ f(\epsilon(t,\theta)) \tag{3.6}$$

Most of the times a least squares (LS) criterion is chosen:

$$\hat{\theta}_N = \arg\min_{\theta} V_N(\theta) = \arg\min_{\theta} \frac{1}{N}\sum_{t=1}^{N} \varepsilon^T(t,\theta)\Lambda^{-1}\varepsilon(t,\theta) \tag{3.7}$$

Here, Λ is a weighting matrix (often chosen equal to I_p). Another way of estimating the parameters in θ is through so called instrumental variables (see *e.g.* [64]) but here only the optimization based way of parameter estimation is considered with a LS criterion of the form (3.7).

Model structures and linear regression model estimation

The choice of the model structure or model *set* $\mathcal{M}$, *i.e.* type, size (*e.g.* model order) and parameterization of the transfer function matrices $G(q,\theta)$ and $H(q,\theta)$, has been a major issue in the system identification literature for many years for such reasons as unique identifiability of the parameters contained in θ and fastness of estimation. The model structures encountered in the system identification literature are mostly rational transfer functions in the shift operator q. Model structures that are particularly popular within the field of process identification are the ARX (Auto-Regression with eXtra inputs/eXogeneous variables) and FIR (finite impulse response) model structures. Because of their relevance for the MSWC plant system identification methodology to be proposed here, these model structures are discussed here in more detail.

More specific, the ARX model structure is given by

$$G(q,\theta) = A(q,\theta)^{-1}B(q,\theta) \qquad\qquad H(q,\theta) = A(q,\theta)^{-1} \tag{3.8}$$

with the matrix polynomials $A(q,\theta)$ and $B(q,\theta)$ given as

$$\begin{aligned} A(q,\theta) &= I_p + A_1q^{-1} + ... + A_{na}q^{-na} \\ B(q,\theta) &= B_0 + B_1q^{-1} + ... + B_{nb}q^{-nb} \end{aligned} \tag{3.9}$$

with the matrices A_i, $i = 1...na$ and B_j, $j = 0...nb$, containing the parameters to be estimated contained in θ. The FIR model structure is a more simple variant of the ARX model structure with $A(q,\theta) = I_p$.

The popularity of the ARX and FIR model structures within the field of process identificaton is due to the fact that the corresponding models can be estimated fastly. This is due to the fact that these model structures are *linear in the parameters*, which provides the corresponding models the name of *linear regression* models. The solution to the optimization problem underlying the linear regression model estimation problem is well known (see *e.g.* [64]) and is provided, for completeness, in appendix B.

Below, the choice of model structure, in particular the choice between FIR or ARX model structure, is discussed in more detail when discussing model bias.

Model validation

Common ways of validating estimated PEM models are

- by comparing *simulated* model outputs (eqn. (3.3) with $e(t) = 0 \; \forall \; t$) with their measured counterparts
- by comparing *predicted* model outputs (eqn. (3.4)) and their measured counterparts
- by evaluating
  - the cross-correlation function between the inputs $u(t)$ and the model residuals $\varepsilon(t, \hat{\theta}_N)$
  - the auto-correlation function of the model residuals $\varepsilon(t, \hat{\theta}_N)$

Because of its relevance for the discussion later on in this chapter, the correlation functions based way of validating is discussed here in more detail. With the cross-correlation function between $u(t)$ and $\varepsilon(t, \hat{\theta}_N)$ the estimated model $G(q, \hat{\theta}_N)$, also refered to as the *G-estimate* or plant model, is validated. It is based on the fact that this correlation function is zero for all lags when the G-estimate is equal to the true plant dynamics $G_o(q)$. The validation, as a consequence, consists of checking whether this correlation function is sufficiently close to zero for all lags. With the auto-correlation function of $\varepsilon(t, \hat{\theta}_N)$, both the estimated plant and disturbance model ($H(q, \hat{\theta}_N)$ / *H-estimate*) are validated: in case the latter estimates perfectly represent their true counterparts, the model residuals $\varepsilon(t, \hat{\theta}_N)$ represent the white noise inputs $e(t)$ of the disturbance model. Hence, the auto-correlation then represents that of white noise, being zero for all lags except for a zero lag. The validation here consists of checking whether this auto-correlation sufficiently close approximates these values.

Bias, model structure selection and closed-loop identification

Apart from evaluating the quality of the estimated model through the validation techniques discussed above, this quality is often also evaluated through a closer investigation of analytically derived asymptotic properties of models obtained with the PEM, in particular their bias and asymptotic variance properties. Asymptotic here refers to the condition that the number of estimation data, N, goes to infinity. Bias is a systematic deviation between the estimated quantity and its true counterpart, more specific a deviation that hypothetically would be obtained even if the estimation would repeatedly be performed on the basis of an infinitely long estimation data set. Asymptotic (co)variance is the (co)variance (matrix) of the probability distribution of the estimated quantity when N would go to ∞.

For the discussion later on in this chapter, it is important to discuss the properties of the PEM with respect to the bias and asymptotic variance of the G-estimate, which in this thesis represents the model for the furnace and boiler part of the MSWC plant. In this (sub)subsection and the next, the bias properties are discussed whereas the asymptotic variance properties are considered in the final (sub)subsection of this (sub)section.

More specific, here, the influences of

- the conditions during the identification experiment performed for data acquisition
- the model structure

on the bias of the G-estimate are disclosed, with particular focus on FIR and ARX model structures. The experimental conditions here refer to whether the identification experiment has been performed in open-loop, *i.e.* as in figure 3.1, or in closed-loop, *i.e.* as in 3.2 where $K(q)$ represents a controller and $r_1(t)$ and $r_2(t)$ represent user-defined test signals. The influences refered to above can be made explicit through a specific

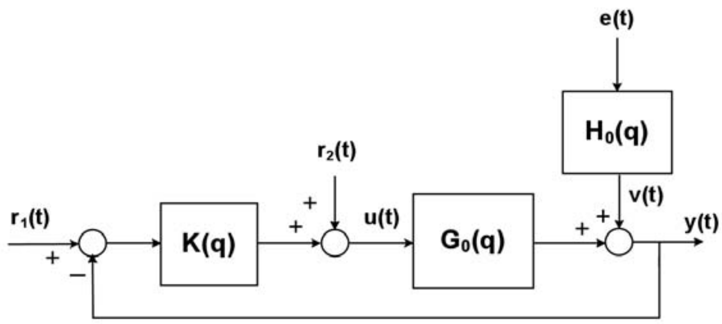

Figure 3.2: Closed-loop identification experimental setting.

convergence result for the PEM that can be found in [28] and states that, under fairly general conditions (to be found in [28, 64]), the parameter estimate $\hat{\theta}_N$ converges as N tends to ∞ with probability 1 to an element of the set

$$\begin{aligned} D_c &= \arg\min_{\theta} \int_{-\pi}^{\pi} tr\left[[(G_o + B_G - G_\theta)\Phi_u(G_o + B_G - G_\theta)^* \right. \\ &\qquad \left. +(H_o - H_\theta)\Phi_e^r(H_o - H_\theta)^*](H_\theta \Lambda H_\theta^*)^{-1}\right] d\omega \end{aligned} \tag{3.10}$$

with

$$B_G = (H_o - H_\theta)\Phi_{eu}\Phi_u^{-1} \tag{3.11}$$

(the so called ”bias-pull”) and

$$\Phi_e^r \quad = \quad \Lambda_o - \Phi_{eu}\Phi_u^{-1}\Phi_{ue} \tag{3.12}$$

Here, B_G characterizes the amount of bias that is obtained for the G-estimate due to the controller induced correlation between the white noise terms e and inputs u. In the open-loop case, *i.e.* with no controller being active, $\Phi_{ue} = 0$, $\Phi_{eu} = 0$ and, consequently, $B_G = 0$ and (3.10) reduces to

$$\begin{aligned} D_c \quad = \quad \arg\min_{\theta} \int_{-\pi}^{\pi} & tr\,[[(G_o - G_\theta)\Phi_u(G_o - G_\theta)^* \\ & +(H_o - H_\theta)\Lambda_o(H_o - H_\theta)^*](H_\theta \Lambda H_\theta^*)^{-1}\big]\, d\omega \end{aligned} \tag{3.13}$$

From this integral it can be deduced that, under the assumptions that (i) the identification experiment has been *informative* [64] *c.q.* $\Phi_u > 0 \;\forall\; \omega$ and (ii) G_o is contained in the model set for G_θ (denoted as $G_o \in \mathcal{G}$), the G-estimate is estimated without bias irrespective whether the disturbance model *c.q.* the H-estimate is estimated with bias or not as long as the G- and H-estimate are independently parametrized. The latter implies that the G-estimate can even be obtained unbiasedly if the disturbance model is chosen equal to the unity matrix I_p, in which case one speaks of an *output error* model structure. Hence, focusing on the influence of the model structure on the bias for the G-estimate and restricting the attention to the FIR and ARX model structure, one can conclude for the open-loop case that

- an unbiased G-estimate can always obtained with the FIR model structure
- an unbiased G-estimate can be obtained with the ARX model structure only if the order (na) of this model structure is chosen high, *i.e.* such high that the disturbance model is contained in the corresponding model set (*i.e.* $H_o \in \mathcal{H}$).

The second conclusion follows from the fact that the $A(q, \theta)$ matrix polynomial is contained in both the plant and disturbance model, see eqn. (3.8), and that, hence, the G- and H-estimate are not independently parametrized and, thereby, the order for the $A(q, \theta)$ matrix polynomial must be chosen high to ensure that the corresponding disturbance model $H_o(q)$ is truly contained in the model set for $H(q, \theta)$.

In the closed-loop experimental situation, $\Phi_{ue} \neq 0$, $\Phi_{eu} \neq 0$ and, consequently, $B_G \neq 0$ and conclusions with respect to bias for the G-estimate must be derived from (3.10) instead of (3.13). The main conclusion is that an unbiased G-estimate is obtained only if $H_o \in \mathcal{H}$ (even if and additional to $G_o \in \mathcal{G}$ and independent from the parametrization). Hence, now, the FIR model structure cannot be used. For the ARX model structure it follows, again, that the order must be chosen high in order to obtain an unbiased G-estimate.

A two stage method for closed-loop identification without restriction on the model structure

The restriction to specific model structures, ARX model structures in particular, may not be desirable. A first reason for that is that one may arrive at a more compact model with one model structure compared to the other, thereby leading to lower computational times and model variance (due to a lower number of parameters). For example, the usage of an FIR model structure may lead to more compact models due to the absence of a disturbance model. (It is noted, though, that the usage of a FIR model structure does not automatically guarantee a compact model (see *e.g.* [123, 65])). A second reason is that one model structure may lead to models with undesired properties whereas the other may not. For example, estimated ARX models may be unstable whereas estimated FIR models may not. On the other hand, estimated FIR models may exhibit highly oscillatory behavior (which may be artificially removed through regularization of the optimization problem underlying the estimation) whereas estimated ARX models may not. Additionally, considerations of *bias tuning* may lead to prefer an FIR over an ARX model structure, *i.e.* the ability to tune the bias through *e.g.* the input spectrum such that it is minimal in some desired dynamical range (see *e.g.* [102]).

Summarizing, it may be desirable to remove the restriction on the choice of model structure present when estimating a model with the PEM on the basis of closed-loop data as this allows for (i) an increased control over the number of model parameters and, thereby, over model variance and computation time, (ii) avoidance of model structures with undesired properties such as instability or oscillatory behavior and (iii) bias tuning. In particular, it allows for usage of the FIR model structure instead of the ARX model structure, although from the point of view of points (i) and (ii) this does not necessarily guarantee an improvement. To remove the restriction on the choice of model structure in case of closed-loop estimation data and thereby to obtain the mentioned advantages, an alternative system identification method for dealing with such data has been proposed in [102]. More specific, this method aims to remove the requirement of having to estimate a full disturbance model when aiming for an unbiased G-estimate on the basis of closed-loop data, albeit that it takes *two* consecutive open-loop identification steps to achieve this goal. Because of its relevance for the discussion in this chapter and the next one, this so called *two stage method* is briefly outlined here. Consider the closed-loop setting in figure 3.2 with, without loss of generality, $r_1(t) = 0$. Also, denote $r(t) := r_2(t)$. The in- and outputs for this setting are given as

$$\begin{aligned} y(t) &= G_o(q)S_i(q)r(t) + S_o(q)H_o(q)e(t) \\ u(t) &= S_i(q)r(t) - S_i(q)K(q)H_o(q)e(t) \end{aligned} \tag{3.14}$$

with

$$\begin{aligned} S_i(q) &= [I_m + K(q)G_o(q)]^{-1} \\ S_o(q) &= [I_p + G_o(q)K(q)]^{-1} \end{aligned} \tag{3.15}$$

Note now that the description for $y(t)$ in (3.14) implies that, if

$$u_r(t) := S_i(q)r(t) \tag{3.16}$$

is known, one can obtain an estimate for $G_o(q)$ by identifying a model between this signal $u_r(t)$ and $y(t)$. The two stage method of [102] exploits this fact by first reconstructing $u_r(t)$ and then, at the second stage of the method, perform the latter identification step. For the reconstruction of $u_r(t)$, an estimate $\hat{S}_i(q)$ for $S_i(q)$ is first obtained after which the reconstructed version of $u_r(t)$ is obtained as the simulated outputs of this model with the known signals $r(t)$ as the inputs, as implied by eqn. (3.16). For obtaining $\hat{S}_i(q)$ the description for $u(t)$ in (3.14) is exploited, *i.e.* this estimate is obtained as the estimated transfer function matrix from $r(t)$ to $u(t)$. Hence, the two stage method consists of two open-loop estimation steps, *i.e.* one for obtaining $\hat{S}_i(q)$ and one for obtaining the actual G-estimate, implying that at both these steps an unbiased G-estimate can be obtained without estimating a full disturbance model. Hence, now also a FIR model structure can be used for obtaining the G-estimate.

Because of its indirect way of providing the G-estimate, the two stage method is also refered to as an *indirect closed-loop identification method* [101]. In contrast, applying the PEM directly to closed-loop identification data is refered to as *direct identification*.

Asymptotic variance

Asymptotic variance expressions for models estimated with the PEM can be found in [66] for the SISO (single input single output) case and in [28] for the more generic MIMO case. Here, for ease of explanation, the SISO results of [66] are considered. More specific, a main result in this reference states that, when $\mathcal{S} \in \mathcal{M}$ (*i.e.* when both $G_o \in \mathcal{G}$ and $H_o \in \mathcal{H}$), the asymptotic (co)variance of the (SISO) G-estimate is given as

$$cov\,(G_\theta) \sim \frac{n}{N}\frac{\Phi_v}{\Phi_u^r} \tag{3.17}$$

where Φ_u^r represents the part of the input spectrum due to the external excitation signals r_1 and r_2: $\Phi_u = \Phi_u^r + \Phi_u^e$ (*i.e.* the part due to the disturbance sources e). In words and as intuitively may be expected, asympotically, the variance of the G-estimate is proportional to the 'signal-to-noise ratio' $\frac{\Phi_v}{\Phi_u^r}$ and the model order n *c.q.* number of model parameters, and inversely proportional to the number of data N. In the remainder of this thesis, only the proportionality between model variance and the number of parameters is considered.

3.3 Elements of process identification

3.3.1 Introduction

The MSWC plant system identification methodology to be proposed here contains many elements from common process identification practice. These elements are solutions chosen to optimally handle specific experiment and estimation related issues that are characteristic for this branche of system identification. These issues and corresponding solutions are discussed now, thereby also providing an idea of the state-of-the-art of process identification.

3.3.2 Experimental issues

Identification experiments at industrial processes are particularly characterized by:

- *Large disturbances*
- *Constraints on process variables* to minimize economic losses, in particular due to off-spec production
- *Limited experimentation time*, where the limitation is due to *e.g.* economic or safety reasons
- *Closed-loop identification*

If not properly accounted for, these characteristics may result in the following problems:

- highly uncertain models in the sense of having a high variance, in particular with respect to the slowest dynamic behavior of the plant to be modeled
- model bias
- long estimation times

More specific, large unmeasured disturbances during the identification experiments result in low signal-to-noise estimation data, which may result in uncertain models *c.q.* models with a high variance. Also, constraints on process variables may lead to high variance models due to the corresponding restrictions on the test signals and, thereby, on the signal-to-noise ratio. High model variance also results from usage of a high order model structure, which is typically chosen in process identification applications to avoid bias.

Additionally, large disturbances may result in fracturization of the experimentally obtained data set(s) due to (user-defined) removal of one or more large pieces of bad data from the originally obtained data set(s), *e.g.* pieces of data where the process under consideration has drifted too far away from the operating point at which the model is to be valid. A first problem associated with the resulting fractured data sets is that the estimation method to be employed may not be not equipped with the capability of properly handling these multiple data sets, *i.e.* it is only capable of handling one data set. Stacking up of the experimentally obtained multiple data sets into one data set and using this data set for estimation leads to model errors of the bias type.

Both fracturization of the data set(s) and a limited experimentation time may lead to the resulting data set(s) to be relatively short compared to the dominating time constant of the plant to be modeled. As a result, the resulting one or multiple data sets may not contain sufficiently well the slowest dynamical behavior of this plant, which leads then to models with inaccurate prediction capabilities of this behavior.

As discussed in section 3.2.3, estimating a model on the basis of data that have been obtained under closed-loop conditions may lead to this model to be biased. In particular, a bias results when specific model structures are chosen. More specific, the usage of an output error model structure such as the FIR model structure leads to a biased model.

The usage of a high order model structure may not only lead to a high model variance but also to long estimation times.

The solutions that are employed within the field of process identification to overcome these problems are:

- Dedicated design of the test signals, denoted as *input* or *experiment design*, such that (at least) the process dynamics relevant for the purpose of the model *c.q.* control design are excited with sufficient high input power, *i.e.* with sufficiently high signal-to-noise ratio. Moreover, experiment design also aims to take into account operational constraints such as constraints on experimentation time and constraints to avoid off-spec production. See *e.g.* [9, 122, 123] for currently employed experiment design approaches.

- Enforcement of *a priori* determined static gains on the model to be estimated to improve the quality of the model with respect to the slow dynamics. See *e.g.* [65, 103]. The static gain values to be enforced can be obtained *e.g.* through offline analysis of historic process data or from a (steady-state) first-principles model. Enforcement of the static gains can be done by extending the numerical optimization method underlying the estimation with constraints that are translations of the equalities that define the static gain enforcement problem. This extension is a fairly standard procedure. Nevertheless, for completeness, it is discussed in more detail in appendix B for linear regression model structures.

- Employment of a multiple data set identification method, *i.e.* a method that allows for proper model estimation on the basis of more than one data set. Such methods are discussed in *e.g.* [64, 103]. The key to proper multiple data set estimation is to avoid that the output predictors (3.4) are a function of more than one data set. This interference between multiple data sets happens *e.g.* when simply stacking these data sets on top of each other and using a single data set based identification method. Multiple data set identification is presently also a fairly standard method. Nevertheless, for completeness, it is elaborated into more detail in appendix B for the case of linear model structures.

- Estimation of a full disturbance model to avoid bias in case of direct identification on the basis of closed-loop data, in particular through the usage of a high order ARX model (see *e.g.* [122]).

- Usage of the two stage method of [102] to both avoid bias and retain the flexibility in choice of model structure and, thereby, the corresponding advantages (outlined in section 3.2.3). See *e.g.* [21] for an application of this method to a continuous crystallization process.

Other typical characteristics of process identification methods are related to the estimation phase and are discussed now.

3.3.3 Estimation issues

Issues that are particularly characteristic for the estimation of models of industrial processes are:

- Obtaining a model with a minimal combined bias and variance error

- Avoidance of long estimation times
- Avoidance of a difficult and, thereby, time-consuming choice of model structure
- Obtaining a numerically reliable estimation method

The first issue here is a translation of the requirement to obtain a model that is as accurate as possible. In system identification theory, where bias and variance are considered the two main model error sources, this translates to the requirement of a minimal combined bias and variance error. The last three issues above are related to the objective of having a fast and easy applicable estimation method. It is noted that estimation times may easily become large for industrial processes due to the many in- and outputs and, thereby, large amount of parameters to be estimated. The choice of model structure also concerns the aim of having as few parameters to be estimated as possible, without introducing a significant bias error, to reduce the variance of the model.

The main solution used in the field of process identification to overcome the issues above is a two-step strategy where in the first step a high order linear regression model, often an ARX model, is estimated and then, in the second step, model reduction is applied to obtain a lower order approximation of the high order model. See *e.g.* [19, 120, 122]. A first advantage of this method is that overall estimation time is low. This is particularly due to the usage of a linear regression model structure in the high order modeling step. Secondly, the often difficult choice of model structure is greatly simplified. This simplification is, first of all, due to the absence of a preliminary step where parameters are estimated that characterize a suitable model structure with a minimal number of parameters. Also, the model order choices to be made in the high order modeling and model reduction steps typically are easy or automated. A third advantage of the two step strategy is its numerical reliability. Finally, the two step approach provides a simple solution to the problem of arriving at a model with a minimal combined bias and variance error. More specific, the usage of a high order model in the first step ensures a low bias, though at the cost of a high variance due to the high number of parameters, whereas the model reduction step allows for a reduction of this variance to an acceptably low level.

Other approaches estimate the model directly from the data. An example is the approach proposed in [123, 65] where, motivated by the many parameters that typically need to be estimated when using an FIR model structure, a compact so called Minimal Polynomial and Start Sequence of Markov Parameters (MPSSM) model structure is used to minimize the number of model parameters and, thereby, the variance of the estimated model. The model structure is not of the linear regression type but the specific model structure is exploited in the optimization to reduce the computation time to a minimum.

The process identification issues discussed here are also relevant for the identification of MSWC plants. As a consequence, the solutions discussed here can also be used for the identification of these plants. In fact, a number of these solutions are chosen as part of the MSWC plant system identification methodology proposed here and discussed later on in this chapter. First, now, two issues are discussed that are particularly relevant for the identification of MSWC plants but which are not explicitly dealt with in the process identification literature (and, consequently, are not discussed in this section). These issues are related to the validation of estimated MSWC plant models and to the partial closed-loop identification setting that may be encountered at MSWC plants during the identification experiment. These two issues are discussed in the following two sections together with solutions proposed here to overcome them.

3.4 A note on the validation of estimated MSWC plant models

A particular characteristic of MSWC plant system identification is that, in the typical case that no on-line disturbance *c.q.* waste composition estimator is present at the MSWC plant to be modeled, proper validation of the estimated plant model(s) is impossible through a simple comparison of simulated outputs and their measured counterparts. More specific, the presence of large nonmeasured disturbances causes a large gap between these variables which cannot be interpreted as an indication of the plant model being inaccurate as this large gap may as well be present when the model is accurate. As a consequence, one lacks an important tool for validating MSWC plant models. However, system identification theory provides alternative tools that are suitable for the validation of such models, as already outlined in section 3.2.3. In particular, these tools are (i) comparing the gap between *predicted* model outputs and their measured counterparts and (ii) analyzing the inputs-residuals and residuals-residuals correlation functions. In the MSWC plant system identification methodology to be proposed here these validation tools are employed. In particular, validation through the mentioned correlation functions has proven to be a useful tool for the validation of estimated MSWC plant models (see *e.g.* the next chapter).

3.5 Partial closed-loop identification

3.5.1 Introduction

Identification experiments at MSWC plants may be performed in open-loop but often also have to be performed in closed-loop out of economic, safety or other reasons. In particular, the situations depicted in figure 3.3 are encountered at MSWC plants during identification experiments. Note that the first two of these settings correspond to the open-loop configuration of figure 3.1 resp. the closed-loop configuration of figure 3.2. The third setting is characterized by inputs and outputs that are not contained in a feedback control loop. An open-loop input is *e.g.* ambient air temperature or an on-line estimate of the waste composition as obtained by *e.g.* the calorific value sensor discussed in this thesis. An open-loop output is *e.g.* the gas phase temperature, which one may want to have as a model output to maintain its value, within a model based controller, below some upper limit. In system identification terms, this third setting corresponds to the one depicted in figure 3.4. Note the notation that is used in this figure, in particular the partitioning of the signals in open-loop inputs $u_1(t)$ and closed-loop inputs u_2, etc. The closed-loop setting of figure 3.4 is different from the

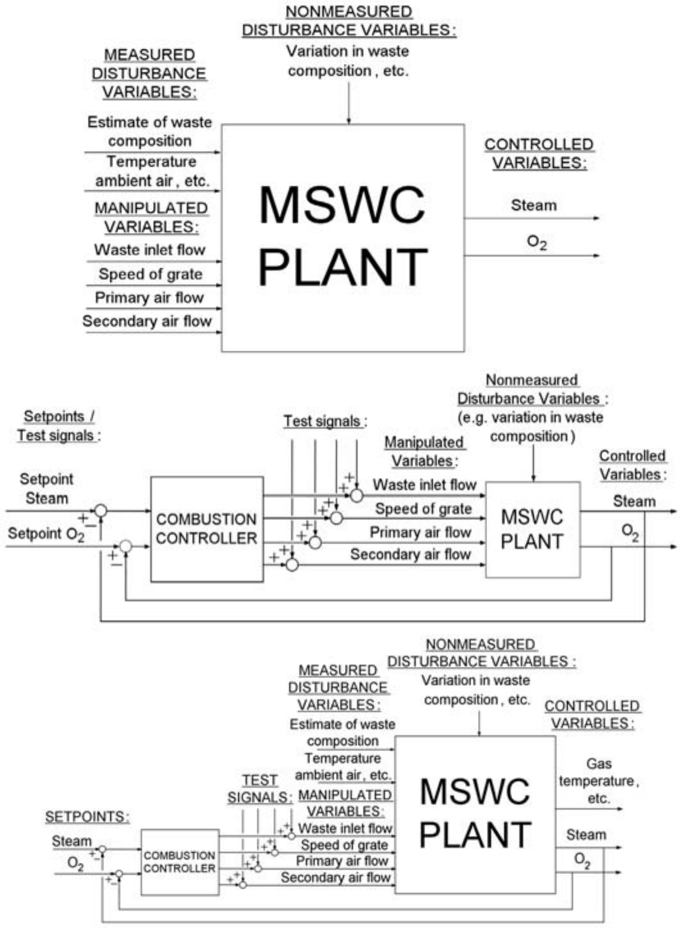

Figure 3.3: Experimental settings at MSWC plants.

closed-loop setting generally encountered in the literature, which is of the type of figure 3.2, and is refered to here as a *partial closed-loop* setting. In contrast, the closed-loop setting of figure 3.2 is refered to here as a *completely* closed-loop setting.

Model estimation on the basis of partial closed-loop identification data has not explicitly been studied before in the system identification literature. Considering the issues that are relevant for the estimation of models on the basis of completely closed-loop data, in particular with respect to the restriction this may impose on the choice of model structure (and, thereby, on the reduction of model variance and computation time and on the ability to tune the bias; see the discussion in section 3.2.3), relevant questions here are:

- Does the usage of partial closed-loop identification data imposes restrictions on the choice of model structure when aiming for an unbiased G-estimate?

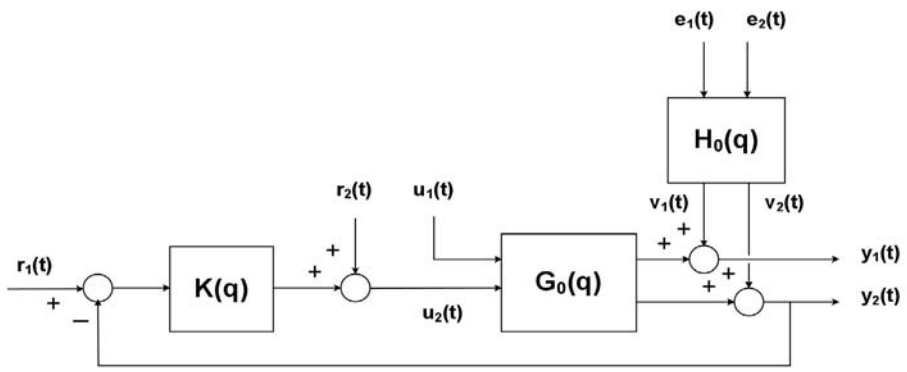

Figure 3.4: Partially closed-loop identification setting.

- How can these possible restrictions be overcome?

Because of their relevance for MSWC plant system identification, these questions are addressed now below.

3.5.2 Analysis of the bias in the presence of partial closed-loop data

The bias properties of a model estimated on the basis of partial closed-loop identification data can be derived from the general convergence result (3.10) - (3.11) by substituting $\Phi_{e_1u_1} = 0$ and $\Phi_{e_2u_1} = 0$ into

$$\Phi_{eu} = \begin{bmatrix} \Phi_{e_1u_1} & \Phi_{e_1u_2} \\ \Phi_{e_2u_1} & \Phi_{e_2u_2} \end{bmatrix} \tag{3.18}$$

The zero entries $\Phi_{e_1u_1} = 0$ and $\Phi_{e_2u_1} = 0$ are due to the fact that, in case of partial closed-loop identification, there is no correlation between the white noise signals entering the disturbance model and the open-loop inputs. More specific, the bias properties for the partially closed-loop case follow from analyzing whether substitution of $\Phi_{e_1u_1} = 0$ and $\Phi_{e_2u_1} = 0$ renders elements of B_G (structurally) zero. If so, the related entries of the G-estimate may be identified without bias irrespective of the choice for the disturbance model. Performing this bias analysis leads to the following proposition:

Proposition 3.5.1 *Consider the PEM setting of section 3.2.3 and assume the estimation data to have been obtained through an informative experiment under the partially closed-loop setting formulated above (see figure 3.4). In this partially closed-loop setting, none of the entries of G_o can be identified without bias independent of the choice of the model structure for H_o.*

Proof 3.5.2 *Substituting $\Phi_{e_1u_1} = 0$ and $\Phi_{e_2u_1} = 0$ into the expression (3.11) for the bias pull B_G leads to*

$$B_G = \begin{pmatrix} B_G^{11} & B_G^{12} \\ B_G^{21} & B_G^{22} \end{pmatrix} \tag{3.19}$$

with

$$
\begin{aligned}
B_G^{11} &= -(H_o^{11} - H_\theta^{11})\Phi_{e_1u_2}\Phi_{u_2}^{-1}\Phi_{u_2u_1}\Delta^{-1} - (H_o^{12} - H_\theta^{12})\Phi_{e_2u_2}\Phi_{u_2}^{-1}\Phi_{u_2u_1}\Delta^{-1} \\
B_G^{12} &= (H_o^{11} - H_\theta^{11})\Phi_{e_1u_2}(\Phi_{u_2}^{-1} + \Phi_{u_2}^{-1}\Phi_{u_2u_1}\Delta^{-1}\Phi_{u_1u_2}\Phi_{u_2}^{-1}) + \\
&\quad (H_o^{12} - H_\theta^{12})\Phi_{e_2u_2}(\Phi_{u_2}^{-1} + \Phi_{u_2}^{-1}\Phi_{u_2u_1}\Delta^{-1}\Phi_{u_1u_2}\Phi_{u_2}^{-1}) \\
B_G^{21} &= -(H_o^{21} - H_\theta^{21})\Phi_{e_1u_2}\Phi_{u_2}^{-1}\Phi_{u_2u_1}\Delta^{-1} - (H_o^{22} - H_\theta^{22})\Phi_{e_2u_2}\Phi_{u_2}^{-1}\Phi_{u_2u_1}\Delta^{-1} \\
B_G^{22} &= (H_o^{21} - H_\theta^{21})\Phi_{e_1u_2}(\Phi_{u_2}^{-1} + \Phi_{u_2}^{-1}\Phi_{u_2u_1}\Delta^{-1}\Phi_{u_1u_2}\Phi_{u_2}^{-1}) + \\
&\quad (H_o^{22} - H_\theta^{22})\Phi_{e_2u_2}(\Phi_{u_2}^{-1} + \Phi_{u_2}^{-1}\Phi_{u_2u_1}\Delta^{-1}\Phi_{u_1u_2}\Phi_{u_2}^{-1}) \qquad (3.20)
\end{aligned}
$$

and $\Delta = \Phi_{u_1} - \Phi_{u_1u_2}\Phi_{u_2}^{-1}\Phi_{u_2u_1}$, *with each of the block elements* B_G^{11}, B_G^{12}, *etc. representing the bias pull for a corresponding block entry* G_o^{11}, G_o^{12}, *etc. of*

$$
G_o = \begin{bmatrix} G_o^{11} & G_o^{12} \\ G_o^{21} & G_o^{22} \end{bmatrix} \qquad (3.21)
$$

which has been partitioned according to the split in open- and closed-loop in- and outputs (see figure 3.4). The fact that all of the entries of B_G *remain non-zero (One might have expected that at least some part of* B_G*, e.g.* B_G^{11} *would have become zero) and dependent on the bias of some part of the noise model, which has been partitioned in the same way as the plant model in (3.21), proves proposition 3.5.1.*

From proposition 3.5.1 it follows immediately that, for the considered partial closed-loop setting, one is restricted to model structures that estimate a full disturbance model, hence to the ARX model structure when considering only linear regression models, if one is to avoid bias for the G-estimate. Note that this conclusion is exactly the same as for the completely closed-loop case. Proposition 3.5.1 can be shown to hold also for the case that no open-loop inputs ($dim(u_1) = 0$) or no open-loop outputs ($dim(y_1) = 0$) are present (these special cases are elaborated in [58]), which also are partial closed-loop settings that can be encountered at MSWC plants. Also, it can be shown that under specific conditions parts of G_o *can* be identified unbiased irrespective of the choice of disturbance model. These conditions are (i) uncorrelated disturbances v_1 and v_2, and (/or) (ii) $G_o^{21} = 0$. However, these latter two conditions are not relevant for MSWC plants where generally the output disturbances are highly correlated (the variation of the waste composition has a simultaneous (large) effect on all output variables of the MSWC plant) and $G_o^{21} \neq 0$.

The bias that can be obtained in case of estimating a model on the basis of partial closed-loop identification data is illustrated here through a simulation example. More specific, this simulation example involves an LTI process G_o with 3 inputs and 2 outputs where one output is connected to one input through a feedback controller. In fact, this process is part of a model that has been identified from real-life data obtained in a partial closed-loop setting from a large-scale MSWC plant, using the identification methodology proposed in this chapter. Also, the simulation example involves a full disturbance model H_o for simulating the disturbances. This model was derived from the same real-life MSWC plant data as G_o. A simulation was performed with the closed-loop system excited by two open-loop input signals u_1 and one external signal r_1 (r_2

$= 0$); see figure 3.4. All these signals were uncorrelated white noise signals of length $N = 120000$. The experiment length was chosen this (very) large in order to minimize as much as possible any accidental (variance) error in the estimated model(s), thereby properly disclosing any bias. Both 5 full order output error (OE) models and 5 full order Box-Jenkins (BJ) models were estimated ($G_0 \in \mathcal{G}$) from the resulting partial closed-loop data, with the main difference between these two types of models being that the latter contain a full disturbance model. Step responses of these models are depicted in figure 3.5. It clearly can be seen from this figure that a bias is obtained for

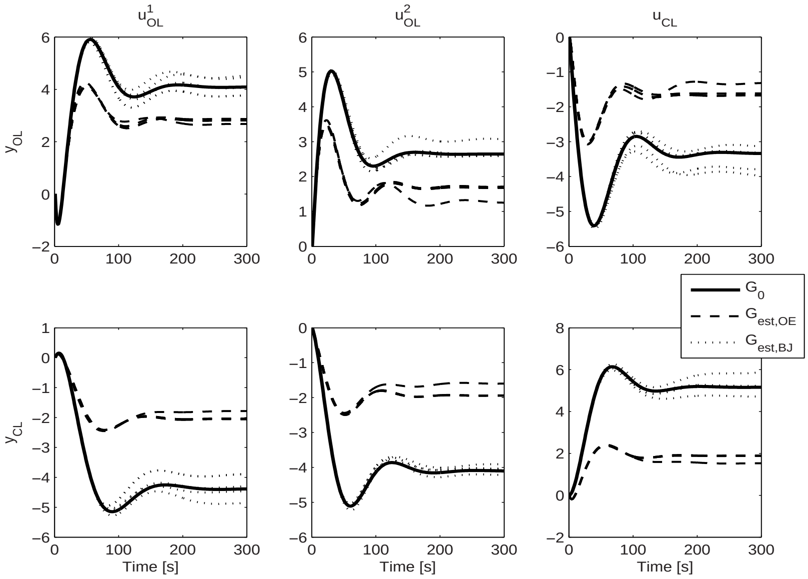

Figure 3.5: Step responses of G_0 (solid), five realizations of OE estimates (dashed) and of BJ estimates (dotted) for data generating system with full noise model H_0 [58]. 'OL' here stands for open-loop and 'CL' for closed-loop.

all entries of G_o in case of the OE estimates (even though only one loop is closed), whereas no bias is present in case of the BJ estimates.

3.5.3 A two stage method for partially closed-loop identification

The restriction on the choice of model structure present when estimating a model directly on the basis of partial closed-loop identification data can again be removed by the two stage method of [102], *i.e.* in (almost) the same way as is the case for completely closed-loop identification data. Consider the setting of figure 3.4 with, without loss of generality, $r_1 = 0$ and $r := r_2$. The first stage of the partial closed-loop two stage method then consists of estimating the transfer from both r and u_1 to u_2 and subsequently constructing the noise free part $\hat{u}_2$ of the latter signal(s). In the second stage, the G-estimate is then obtained unbiasedly via estimating the transfer from

$\hat{u} := [u_1^T\ \hat{u}_2^T]^T$ to the outputs y. Both estimation steps are open-loop and, hence, can be performed with any model structure without introducing bias in the G-estimate.

3.6 A system identification methodology for MSWC plants

3.6.1 Introduction

In this section, a specific MSWC plant system identification methodology is proposed. This methodology is a collection of solutions for issues that play a dominant role at system identification of such plants. These issues are related to the identification experiment, estimation phase and validation phase and have been discussed in the preceding sections together with potential solutions to overcome them. These solutions have been proposed in the process identification literature or are solutions to issues that have not explicitly been considered so far in this literature. In this section, the issues that play a dominant role at MSWC plant system identification are briefly outlined again and solutions are proposed to overcome them to, finally, arrive at a full MSWC plant system identification methodology.

First, the experimental, estimation and validation related issues and solutions are outlined here. After that, the specific MSWC plant system identification methodology is summarized.

3.6.2 Experimental issues and solutions

The identification experiments at MSWC plants are characterized by the same elements as discussed in sections 3.3.2 and 3.5: (i) large disturbances, (ii) constraints in process variables to minimize economic losses, (iii) limited experimentation time and (iv) closed-loop identification. With respect to the latter characteristic it is noted that MSWC plant estimation data may also have been obtained in a partial closed-loop setting, see section 3.5, which situation has not explicitly been dealt with before in the process identification literature. When not properly accounted for, these characteristics may result in the following problems: (i) models with a high variance, in particular with respect to the slowest dynamical behavior, (ii) model bias, (iii) a bad prediction of the slowest dynamics of the plant to be modeled and (iv) long estimation times.

As an illustration of the large disturbances acting on an MSWC plant, see figure 3.6. In this figure, the two most commonly encountered types of MSWC plant disturbances are depicted:

- temporarily large upsets (peaks, outliers)
- randomly varying persistent disturbances (of a stationary stochastic nature)

Temporarily large upsets are the main causes for fractured data sets whereas the disturbances of a randomly varying nature are a major cause for low signal-to-noise-ratio estimation data. It is noted that usage of the calorific value sensor (see section 2.3.3) may significantly decrease the nonmeasured disturbance contribution in MSWC plant identification data but the (CO_2 and H_2O) sensors required for this sensor are often

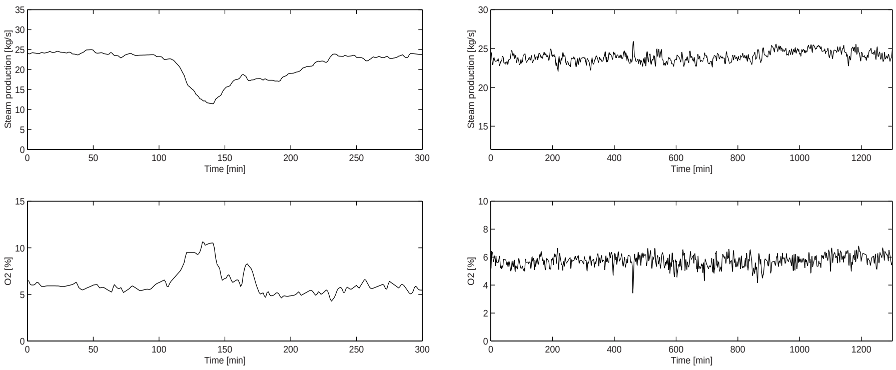

Figure 3.6: MSWC plant disturbances on steam production $[kg/s]$ and O_2-concentration $[\%]$ as measured at a large-scale MSWC plant during normal operation. Left: temporarily large upset. Right: stationary stochastic type of disturbances.

not present.

It is proposed here to use the following solutions to overcome the experiment related MSWC plant identification problems mentioned above, *i.e.* from those solutions discussed in sections 3.3.2 and 3.5: (i) dedicated *experiment design* to maximize the signal-to-noise-ratio of the identification data while minimizing economic losses during experimentation, (ii) enforcement of *a priori* determined static gains on the model to be estimated to improve the quality of the model with respect to the slow dynamics, (iii) employment of a multiple data set identification method and (iv) usage of the two stage method of [102] to avoid bias of the plant model in case of closed-loop identification data, either of the partial or complete type, while also retaining flexibility in the choice of model structure and, thereby, retaining the corresponding advantages of increased control over the model variance and estimation time and ability to tune the bias (see section 3.2.3). Additionally, it is proposed to derive the static gains to be enforced on the model to be estimated from a steady-state first-principles model of the MSWC plant under consideration as this has been found to provide better results than through estimating these gains from experimentally obtained data. Enforcement of the static gains on the model to be estimated can be done either through a commercially available package or by means of the method proposed in appendix B for linear regression type of models. This also applies to the choice of multiple data set identification method.

3.6.3 Estimation and validation related issues and solutions

The estimation issues characteristic to system identification of MSWC plants are the same as those discussed in section 3.3.2, *i.e.* as those encountered at the identification of other industrial processes. These issues are (i) arriving at a model with a minimal combined bias and variance error, (ii) avoidance of long estimation times, (iii) easy choice of model structure and (iv) the estimation method must be numerically reliable. To solve these issues, it is proposed here to use, as part of the MSWC plant system identification methodology considered here, the most commonly applied solution of those used for the identification of other industrial processes, as discussed in section 3.3.2, *i.e.* to use a two-step strategy where in the first step a high order linear regression model, *i.e.* FIR or ARX model, is estimated and then, in the second step, model reduction is applied to obtain a lower order approximation of the high order model. The main reason for this choice are the good and well established properties of this strategy. For the model reduction step it is proposed to use the step response based approximate realization algorithm of [104] (see also [19]). The main reason for that is that it best captures the slow dynamics of the estimated high order model, thereby preventing degradation of the slow dynamic behavior of the finally resulting low order model in comparison to the estimated high order model on which the static gains have been enforced. Because the number of transfer functions in MSWC plant models used for combustion control design typically is relatively low (often in the order of 10), it is proposed to apply the method of [104] to each SISO transfer function of the high order model separately as this has been found to provide overall more accurate approximation results than full MIMO model reduction. For completeness, the method of [104] is provided in appendix A, where, as a novelty, also a link is established between this method and a major group of subspace methods.

As discussed in section 3.4, the main issue with respect to validation of estimated MSWC plant models is that this cannot be done through comparison of simulated and measured outputs. To overcome this problem, validation through input-residual and residual-residual correlation functions is proposed, which is a standard validation technique in the field of system identification.

3.6.4 Summary of the system identification methodology

Summarizing, the MSWC plant system identification methodology proposed here consists of the following steps:

- Dedicated experiment design to obtain a model with a minimum variance as possible while also taking into account constraints on experimentation time and on process variables.

- Estimation of a high order model with the PEM using an FIR or ARX model structure. Enforcement of static gains on this model that have been derived from a steady-state first-principles model of the MSWC plant under consideration. Usage of a multiple data set identification method to use all fractured data sets for estimation. In case of closed-loop data, usage of the two stage method of [102]. Validation of the estimated model(s) with input-residual and residual-residual correlation functions.

- Application of the step response based approximate realization algorithm of [104] to each of the (SISO) transfer functions of the estimated high order model to arrive at a lower order approximation of the high order model.

3.7 Conclusions

It has been shown that sufficient opportunities *c.q.* solutions are available for overcoming all potential obstacles for arriving at a model suitable for model based MSWC plant combustion control through system identification. From these solutions, a specific MSWC plant system identification methodology has been derived, which is the main contribution of this chapter.

Another contribution of this chapter is an analysis of and solution approach to the occurrence of bias in case of estimating a model with the so called direct method on the basis of partial closed-loop identification data. This partial closed-loop identification problem has been investigated because of the fact that it can be encountered at system identification of MSWC plants. Within the prediction error framework, it has been shown that, whereas with the direct method (particular transfers in) multivariable plant models can be identified consistently without consistent noise modeling if the data are obtained in open-loop, in a general closed-loop situation this method looses this property already if one single control loop is present. The loss of this property, which also occurs at MSWC plants due to the specific experimental conditions at these plants, imposes restrictions on the model structure that may have significant negative consequences for the model variance, computation time and ability to tune the bias. A method that allows for overcoming these issues is the so called two stage method.

Chapter 4

An assesment of MSWC plant modeling approaches

4.1 Introduction

In this chapter, the exploration of the opportunities of both first-principles modeling and linear system identification for obtaining a model suitable for model based MSWC plant combustion control is finalized. More specific, through an application of these modeling approaches, in particular those proposed in chapters 2 and 3, on data experimentally obtained from a large scale Dutch MSWC plant a final assessment is made with respect to their ability of delivering such a model. This is the main contribution of this chapter.

The objective of the application considered here was parameter estimation and validation of a chapter-2-type first-principles model of the considered MSWC plant. For that purpose, the system identification methodology proposed in chapter 3 was used. The resulting system identification based method of validating first-principles MSWC plant models is another major contribution of this chapter.

The contents of this chapter are as follows. First, in section 4.2, the parameter estimation and validation oriented application refered to above is discussed in detail and motivated. This also includes an outline of the proposed and applied first-principles MSWC plant model validation method. Additionally, the results are discussed of the final assessment on first-principles modeling and system identification with respect to their ability to deliver models suitable for model based MSWC plant combustion control. Subsequently, in section 4.3, the conclusions of this chapter are provided.

4.2 Closed-loop identification and first-principles modeling of a large-scale MSWC plant

4.2.1 Introduction

With the aim to disclose the dynamics of MSWC plants, in 1998 a project was carried out to validate and adapt a model of the type of chapter 2 to data experimentally obtained from the large scale Dutch MSWC plant Huisvuilcentrale N-H (HVC) at Alkmaar, the Netherlands. In fact, the model that was validated was a simplified version of the model of [41] discussed in section 2.3.2, with the simplification being a linearly approximated reaction rate, and formed the basis for the postulation of this model. The HVC model validation and adaptation project resulted not only in the disclosure of the MSWC plant dynamics but also provided an opportunity for assessing both system identification and first-principles modeling as tools for deriving models suitable for model based MSWC plant combustion control. The assessment of system identification was made possible by the enforced usage of the MSWC plant system identification methodology proposed in chapter 3 to allow for validating the considered first-principles model. More specific, a major result from the HVC project was a specific system identification based method for validating first-principles MSWC plant models that allows for overcoming a typical MSWC plant model validation problem.

In this section, the results of the HVC project are discussed in more detail. The main aim is to provide the results of the assessment of system identification and first-principles modeling as tools for deriving models suitable for model based MSWC plant combustion control. Apart from that, the specific system identification based first-principles MSWC plant model validation method is outlined and motivated and, also, specifics of the application of the MSWC plant system identification methodology proposed in chapter 3 to the (closed-loop) HVC data are provided to illustrate its applicability.

The contents of this section are as follows. First, in section 4.2.2, the specific system identification based validation procedure employed at the HVC is outlined and motivated. Then, in section 4.2.3, details about the application of this procedure to the HVC data are discussed, including details about the corresponding application of the system identification methodology proposed in the previous chapter. Subsequently, in section 4.2.4, results of this application are discussed and conclusions are made with respect to the suitability of first-principles modeling and system identification as tools for deriving models for model based MSWC plant combustion control. Finally, in section 4.2.5, other system identification results obtained with the HVC application are used to provide an idea about the non-measured disturbance contribution in the MSWC plant variables and, thereby, to provide a further motivation for the proposed validation procedure.

4.2.2 A system identification based method for validation of first-principles MSWC plant models

The problem with validating first-principles models is that validation of such models directly on the basis of comparing simulated model outputs with their measured counterparts is not possible while, simultaneously, other validation methods are not available. In fact, the validation problem is the same as encountered at the validation of models obtained with system identification, as discussed in section 3.4, *i.e.* the presence of large nonmeasured disturbances causes a large gap between simulated and measured outputs which cannot be interpreted as an indication of the model being inaccurate as this large gap may as well be present when the model is accurate. The results discussed later on in this section demonstrate this. Whereas system identification theory provides alternative validation methods such as input-residual and residual-residual correlation functions to validate the resulting models, these are not directly applicable to first-principles models. One way to overcome the first-principles model validation problem is to use system identification to provide an alternative model of the MSWC plant and, thereby, to provide alternative simulations of the outputs of the first-principles model. Validation, and possibly parameter adaptation, of the first-principles model is then performed on the basis of comparison of simulated first-principles model outputs with simulated system identification model outputs, both obtained for the same inputs. The latter method has been used to validate the HVC first-principles model considered here.

4.2.3 Experiments, estimation and validaton

MSWC plant HVC is one of the 11 MSWC plants currently present in the Netherlands. It was built in 1995 and has three lines (*i.e.* each one corresponding to a chute and furnace of the form depicted in figure 1.4, etc.) each one of which has a throughput of approximately 18.5 tons of waste per hour by means of which approximately 63 tons of steam (400 °C, 42 bar) is produced per hour (per line).

In July and August of 1998, closed-loop identification experiments were performed at this plant at two distinct operating points, each characterized by a specific primary air flow temperature to (also) examine the influence of this temperature on the waste combustion process *c.q.* $T_{prim} = 70$ $[°C]$ and $T_{prim} = 120$ $[°C]$. Here, only the results for the operating point with $T_{prim} = 70$ $[°C]$ are discussed. The results for the other operating point are similar. The closed-loop experimental setting during the identification experiments at the HVC was the completely closed-loop one depicted in the middle of figure 3.3, with the four MVs waste inlet flow, speed of grate, primary air flow and secondary air flow actually representing setpoints of corresponding lower level controllers. During the identification experiments, these setpoints were excited by user-defined test signals (r_2 in figure 3.4) of the pseudo-random binary sequence (PBRS) type [64], whereas the steam production and O_2-concentration setpoints of the combustion controller were not changed ($r_1 = 0$ in figure 3.4). At first, some short preliminary experiments were performed in order to determine the frequency contents of the PRBS test signals to be used during the final data acquisition experiments. With the latter data acquisition experiments, a total of 5 data sets were obtained with a length of upto 25 $[hrs]$. The sample time was 60 $[s]$, which was thought to be sufficiently small to capture all relevant MSWC plant dynamics. One of the obtained data sets was used for validation only. The operating point values for the resulting data are given in table 4.1. After the data acquisition experiments, a number of pre-treatment steps were

	Variable	*Value*	
Inputs	Waste inlet flow	53.2	$[\% \ (controller \ \ scale)]$
	Speed of grate	79.9	$[\% \ (controller \ \ scale)]$
	Primary air flow	52028	$[Nm^3/h]$
	Secondary air flow	10580	$[Nm^3/h]$
Outputs	Steam	16.3	$[kg/s]$
	O_2	5.7	$[Vol. \%]$

Table 4.1: Operating point values of the HVC plant during the identification experiments ($T_{prim} = 70$ $[°C]$).

applied to the experimentally obtained data to render a proper estimation result. These included subtraction of the means and scaling.

After that, the identification methodology proposed in this thesis was applied to the resulting data, including the two stage method of [102] to account for the completely closed-loop situation. At both steps of the latter method, a high order ARX model was estimated to obtain a model with as small a bias as possible. These estimations were performed with the constrained multiple data estimation method for linear regression models discussed in apendix B. The static gains were enforced (only) on the high order ARX model estimated at the second stage of the two stage method. These static gains were obtained from the first-principles model to be validated after this model was adapted *statically* to the operating point values given in table 4.1. Order selection for and validation of both estimated high order ARX models were done on the basis of the validation data set. The model orders were chosen according to a simple output-error based procedure that is, for completeness, outlined in appendix C. The choices for model orders turned out not to be critical, however, for the final estimation results. The orders for the two resulting ARX models were $n = na = nb = 23$ at the first stage of the method and $n = 18$ at the second stage (It is noted that these numbers comply with those encountered in the literature for similar high order modeling approaches, which were found to range from $n = 10$ to 50. See *e.g.* Zhu and Backx [122, 123]).

Validation of the estimated models was done by means of the correlation tests described in section 3.2.3. The corresponding correlation functions were found to be good, see figure 4.1 for an example, indicating that the estimated models were of good quality. Confidence intervals computed for responses of the resulting high order models indicated a high variance for these models, in particular in the slower dynamical range. As discussed in the previous chapter, this is subscribed to the low signal-to-noise estimation data and to the relatively short length of the estimation data sets. It must be noted, though, that, when bias *is* present, the usage of an ARX model structure penalizes the misfit in the fast dynamical range much more than in the slow dynamical range [64], which may be an additional explanation for the high uncertainty of the estimated models in the slower dynamical range.

The high variance of the model exhibited itself also by the responses being noisy.

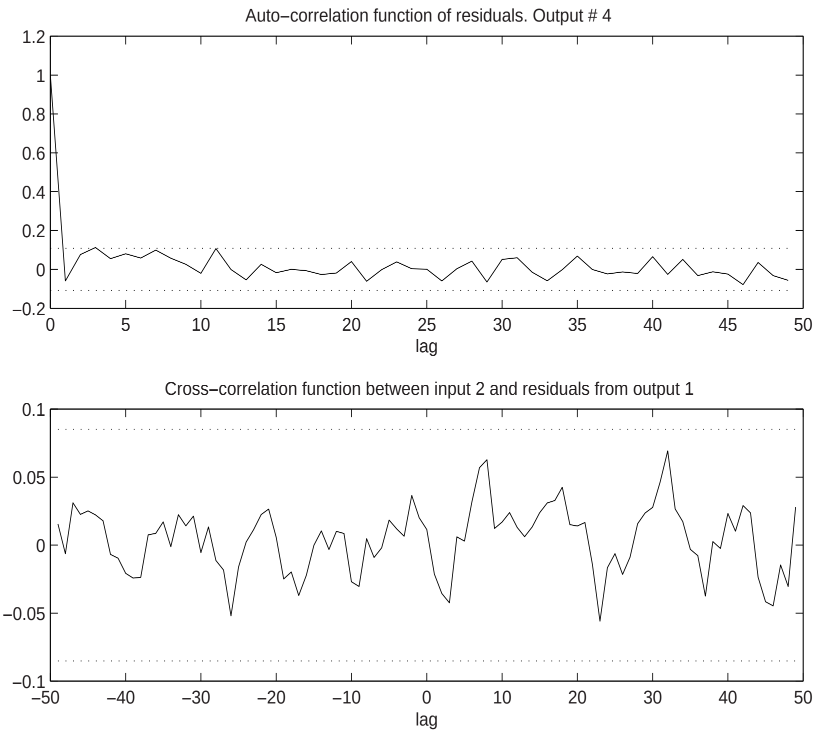

Figure 4.1: An auto- and cross-correlation function for the high order ARX model estimated at the first stage of the two stage method of [102], computed on the basis of the validation data set. Dotted lines = 99 % confidence limits.

To reduce the model variance and the noise on the model responses, model reduction was applied to the high order model with the realization based method of [104]. As discussed earlier, this was done to each SISO transfer function of the high order model separately. Examples of the model reduction results are given in figure 4.2. The model reduction step resulted in 8 low order SISO models (with, typically, $n = 2$ or 3). These SISO models were collected into one MIMO model, representing the dynamics of the MSWC plant HVC.

Responses of this model were subsequently used to simulate the dynamic behavior of the large scale MSWC plant under consideration and validate the first-principles model. For that purpose, the inputs used to produce the responses of the estimated model were given to the first-principles model and the gap between the resulting responses was investigated. At first, this gap was large. However, after several manual (first-principles model) parameter adaptation steps a good agreement was obtained between the estimated and first-principles model responses.

These responses are discussed below.

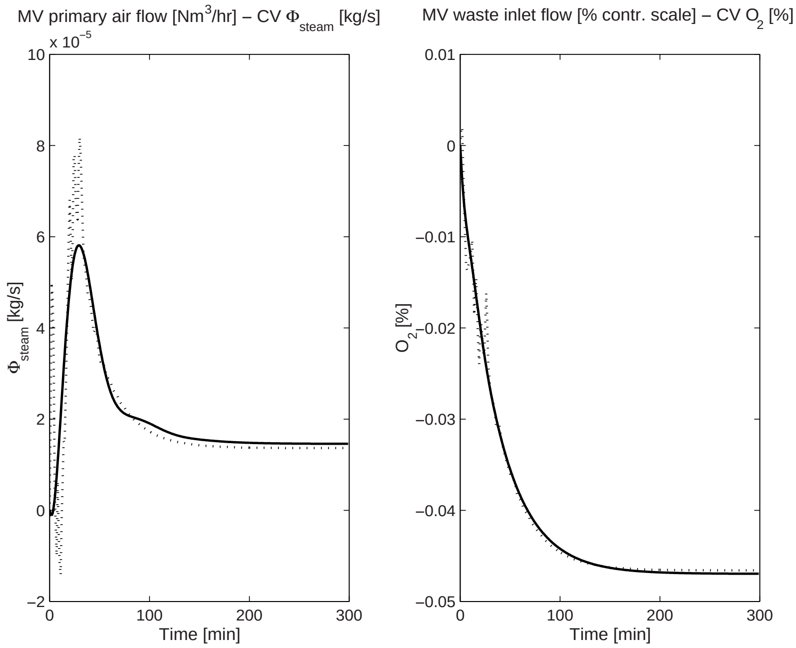

Figure 4.2: Model reduction with the realization based method of [104]. Dotted line = (noisy) high order model step response, solid line = (smooth) reduced order model step response.

4.2.4 Assessment of modeling approaches

In figures 4.3 to 4.6 simulated responses are shown of both the steam production and O_2-concentration as obtained with the estimated and first-principles model finally resulting from the estimation and validation procedure discussed in the preceding sections. These simulations were made with only one input signal excited at a time while the other input signals were kept constant (in case of the first-principles model) or zero (estimated model). The excited input signals are depicted in the lowest part of the figures (These have been chosen here equal to the test signals of the validation data set after having been filtered by the (so called) input sensitivity transfer function matrix obtained in the first stage of the two stage method of [102]).

Figure 4.3 shows the simulated output signals when only the waste inlet flow is excited. One can see that both the simulated responses of the steam production and those of the O_2-concentration coincide very well. A similar good agreement between the simulated outputs of the estimated model and the first-principles model is seen in figures 4.4 and 4.5. A less well agreement is found for the simulated outputs of the steam production as a function of the secondary air flow input signal (upper part of figure 4.6). In contrast, the corresponding simulated responses of the O_2-concentration coincide rather well.

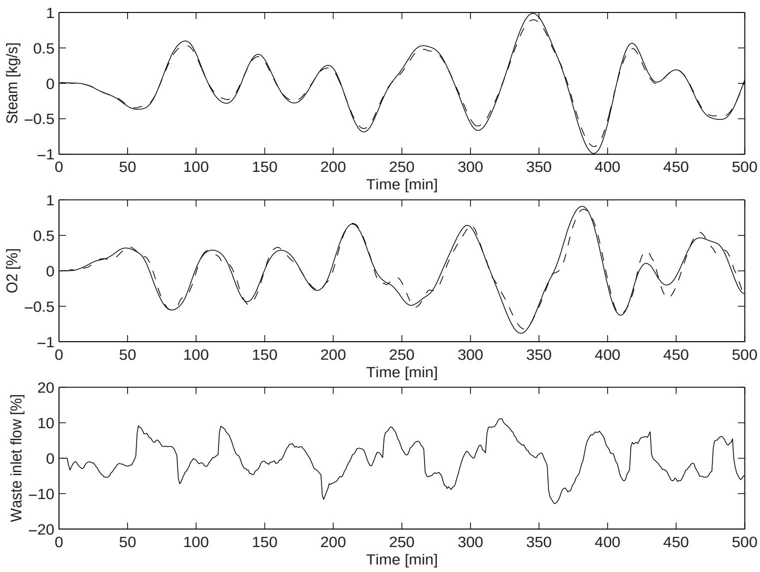

Figure 4.3: Simulation of the steam production and O_2-concentration obtained with the estimated model (— —) and the first-principles model (—) when only the waste inlet flow (lowest part of figure) is excited.

The main and important conclusion that can be made from the good agreement results here, together with the good validation results for the estimated high order model, is that these indicate that *both first-principles modeling and system identification can be used to obtain a model suitable for model based MSWC plant combustion control, i.e. of low order and of sufficient accuracy. In particular, the first-principles modeling and system identification procedures proposed in this thesis allow for the derivation of such a model.*

It is noted that the dynamics of both the estimated and first-principles model considered here are more closely investigated in chapter 5 in terms of their step responses.

4.2.5 A motivation for system identification based validation of first-principles MSWC plant models

From figure 4.7 an idea can be obtained of the contribution of the disturbances to the CVs steam production and O_2-concentration. In this figure these two CVs are depicted together with estimates of their disturbance free counter parts as obtained by simulation using the estimated plant model as filter and the experimentally obtained input signals as input signals to this filter (*i.e.* as $G(q,\hat{\theta})u(t)$ with $G(q,\hat{\theta})$ the estimated plant model. Note that, when this model is equal to the real plant ($G(q,\hat{\theta}) = G_o(q)$), the difference between the measured and simulated output signals is given by the real plant disturbances ($v(t) = y(t) - G_o(q)u(t)$). From figure 4.7 it can be seen that the contribution

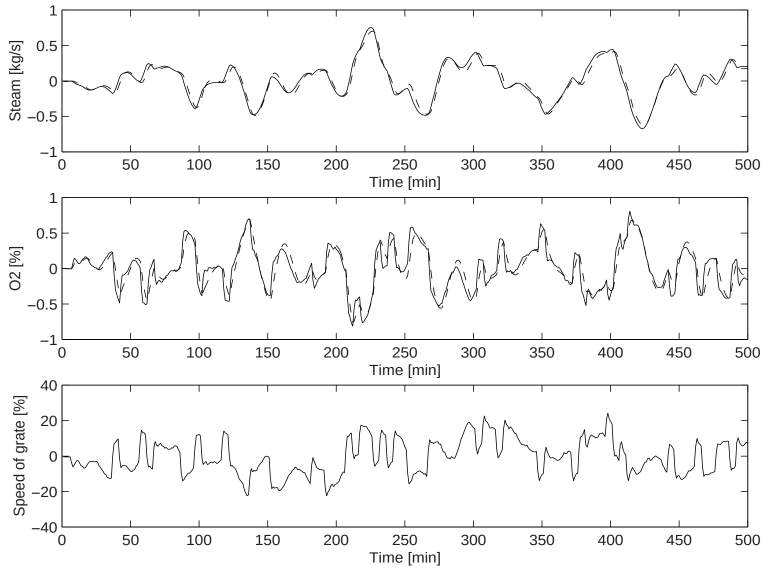

Figure 4.4: Simulation of the steam production and O_2-concentration obtained with the estimated model (− −) and the first-principles model (−) when only the speed of the grate is excited.

of the disturbances to the CVs is large. This is confirmed by signal-to-noise ratios computed on the basis of standard deviations: the contribution of the disturbances to the steam production was 33 $\%$ and even 49 $\%$ for the O_2-concentration. In other words, the gap between simulated model outputs and their measured counterparts is large. However, this does not imply that the estimated model is highly erroneous as the results in the preceding sections indicate the opposite. This observation motivates the earlier proposed validation methods for both system identification and first-principles models, *i.e.* to not use the gap between simulated and measured MSWC plant outputs as a measure for validation but to choose instead another measure for validation.

4.3 Conclusions

Through an application of the earlier proposed first-principles modeling and system identification approaches on data experimentally obtained from a large scale Dutch MSWC plant, a final assessment has been made with respect to the ability of such approaches to deliver a model suitable for model based MSWC plant combustion control (*i.e.* of low order and sufficient accuracy). The main conclusion from this assessment is that both these modeling approaches allow for delivering such a model. In particular, the first-principles modeling and system identification procedures proposed in the previous two chapters of this thesis allow for the derivation of such a model.

Additionally, a new method of validating first-principles MSWC plant models has

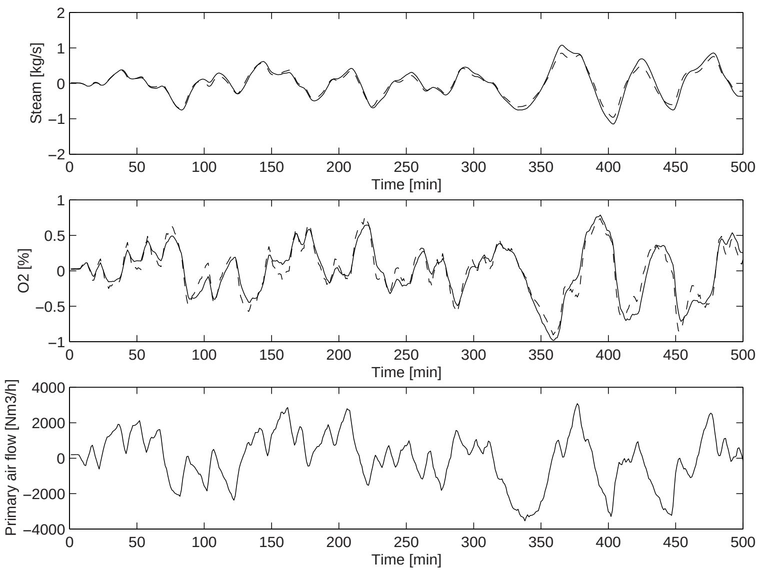

Figure 4.5: Simulation of the steam production and O_2-concentration obtained with the estimated model (− −) and the first-principles model (−) when only the primary air flow is excited.

been proposed. This method aims to overcome the problem that validation of such models directly on the basis of comparing simulated model outputs with their measured counterparts is not possible while, simultaneously, other validation methods are not available. In fact, the presence of large nonmeasured disturbances causes a large gap between simulated and measured outputs which cannot be interpreted as an indication of the model being inaccurate as this large gap may as well be present when the model is accurate. The validity of the proposed validation method has been demonstrated with results obtained from the discussed application.

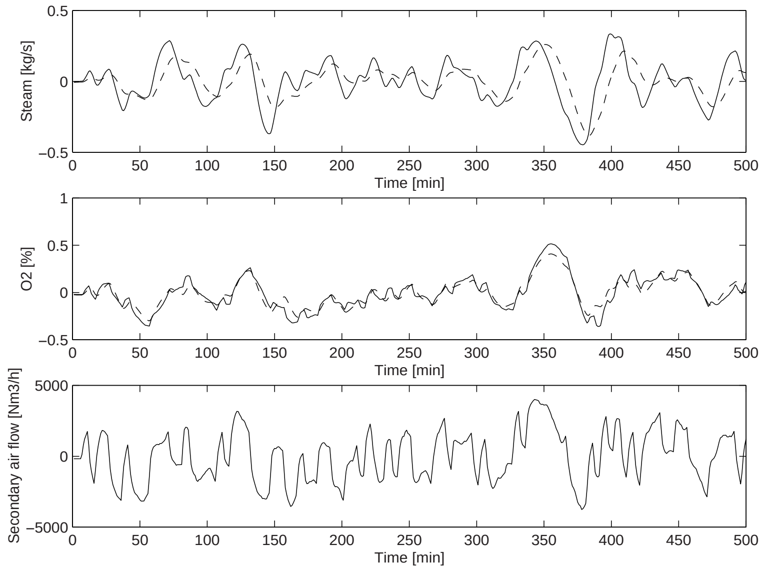

Figure 4.6: Simulation of the steam production and O_2-concentration obtained with the estimated model (– –) and the first-principles model (–) when only the secondary air flow is excited.

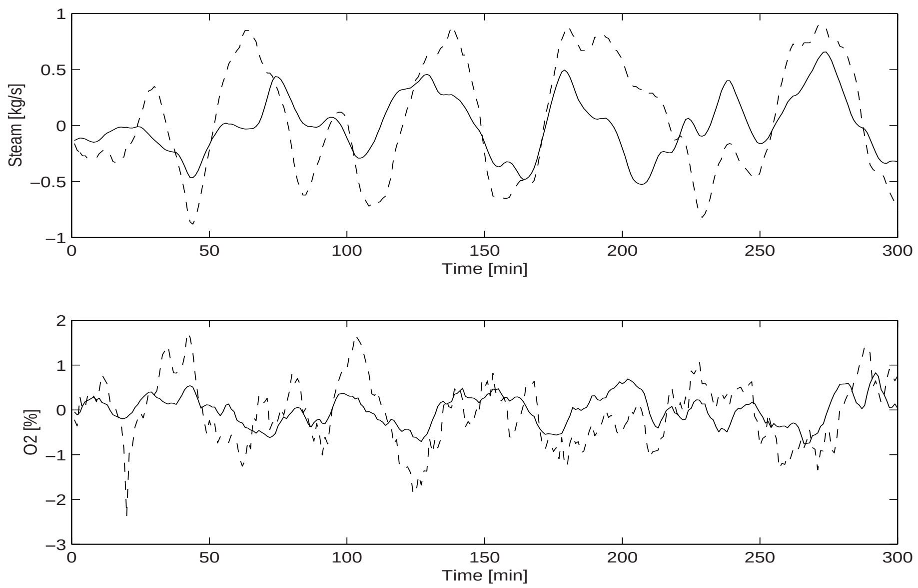

Figure 4.7: Top figure: measured (– –) steam production and disturbance free, simulated, steam production (–) (see text). Bottom figure: measured (– –) O_2-concentration and disturbance free, simulated, O_2-concentration (–).

Chapter 5

Performance assessment of PID based MSWC plant combustion control

5.1 Introduction

In this chapter, the opportunities are explored for improving MSWC plant combustion control strategies of the proportional-integral-derivative (PID) type (which are the mostly applied control strategies in practice) through similar (PID type of) combustion control strategies. The main motivation for exploring these opportunities is that these combustion control strategies are cheaply implementable in practice as the corresponding control soft- and hardware and architecture is generally already in place. Hence, the benefit-cost ratio may be high for this type of combustion control strategy and may cause it to be more beneficial than advanced type of combustion control strategies, even when the latter perform better. Another motivation is that a significant margin for performance improvement could well be present for PID type of combustion control strategies due to the fact that these strategies have not been designed on the basis of an explicit description of the plant dynamics in the form of a dynamic model. In this chapter, this margin for improvement is explored, thereby exploiting the availability of explicit descriptions of the dynamics of MSWC plants in the form of models that have been obtained through the approaches discussed in chapters 2 and 3.

Starting point here is a PID type of combustion control strategy that is applied in practice. This control strategy is compared to a new PID combustion control strategy proposed here that is derived from a closer investigation of the MSWC plant dynamics, in particular the dynamics exhibited by models that have been validated on and adapted to real-life MSWC plant data by means of the procedures of chapters 2 and 3. The results from this comparison are used to assess the margin for improvement present in PID based MSWC plant combustion control [1].

[1] A further assessment of PID combustion control strategies, in particular the one proposed in this chapter,

The contents of this chapter are as follows. First, in section 5.2, a brief review of PID based MSWC plant combustion control is given to provide preliminaries important for the remainder of this chapter, in particular (i) the considered combustion control problem, (ii) an outline of the control strategy to be compared with the new one proposed here and (iii) performance aspects of currently used PID combustion control strategies. After that, in section 5.3, the dynamics of MSWC plants are discussed, in particular through analysis of step responses of the models obtained with the application discussed in section 4.2. Subsequently, in section 5.4, the proposed new PID based MSWC plant combustion control strategy is outlined and motivated, using the results from the analysis of the MSWC plant dynamics made in section 5.3. After that, in section 5.5, this new control strategy is compared to the one used in practice and an assessment is made of the margin for improvement present in this control strategy and in PID based MSWC plant combustion control in a general sense. This assessment also includes an evaluation of the margin for improvement in an overall economic MSWC plant performance sense. Finally, in section 5.6, the conclusions of this chapter are given.

5.2 A review of PID based MSWC plant combustion control

Typically, an MSWC plant combustion controller is of the type schematically depicted in figure 5.1 with the controller being of the multivariable PID type, *i.e.* consisting of a network of PID blocks. The aim of such controllers is to maintain the CVs as close to

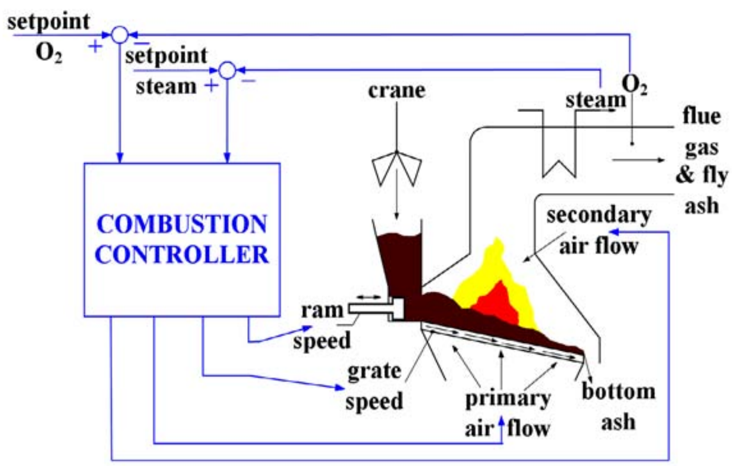

Figure 5.1: A typical MSWC plant combustion controller.

is made in chapter 7. In particular, in this chapter this control strategy is compared to model predictive control based combustion control strategies.

their setpoints as possible, with the deviations typically being large and caused by the variation in waste composition. In other words, these controllers are particularly fit for setpoint tracking and process variation minimization purposes and, hence, for handling an overcapacity market situation (see sections 1.1.3 and 1.2.1). Nevertheless, these controllers may be, and are, also used for constrained maximization/constraint pushing type of MSWC plant combustion control problems to deal with an undercapacity market situation. The setpoints are then chosen such that the MSWC plant operates with maximum possible steam production and waste throughput within the boundaries set by environmental and maintenance related constraints. Most of the times this corresponds to (i) choosing the setpoint of the steam production such that the actual steam production remains as close as possible to an *a priori* specified upper limit (while exceeding it as few times as possible), and to (ii) choosing the setpoint of the O_2-concentration such that the actual O_2-concentration operates as close as possible to an *a priori* specified lower limit (This lower limit is typically enforced by law and is, in the Netherlands, equal to 6 $\%$. Its aim is to ensure complete combustion and, thereby, a minimal amount of toxic emissions).

Note that, irrespective a process variation minimization/overcapacity market situation application or a (pseudo-)constrained maximization/undercapacity application, the PID combustion control system must have good disturbance rejection properties. If not, there is no optimal suppression of variations and relatively large (on average) distances between CVs and the imposed (environmental and maintenance related) constraints must be maintained, degrading the overall economic performance of an MSWC plant for an over- resp. undercapacity market situation.

It is also noted that for an undercapacity market situation the overall economic performance of an MSWC plant may significantly benefit from good setpoint tracking properties of the combustion controller in the sense of offset-free control for all CVs required to follow a setpoint. This is due to the fact that, when the latter is not the case, in this market situation the steam production setpoint may have to be set at a lower than optimal value to account for other CVs not being controlled without offset, leading to a(n on average) lower steam production and waste throughput and, thereby, lower revenues. More specific, a lower steam production setpoint may have to be accepted to avoid frequent violation of a constraint imposed on another, not-offset-free-controlled, CV such as *e.g.* O_2. For an overcapacity market situation, where process variation minimization is more important than revenue maximization, the benefits of offset-free control are lower to nothing due to the fact that a lower steam production (setpoint and) level and corresponding lower waste throughput level may be acceptable and not lead to a significant degradation in overall economic performance. An important observation in this respect is that current PID type of combustion control strategies typically contain integral action for offset-free control of one CV only, which most of the times is the steam production. Apparently, the operators and/or designers of these control strategies do not consider offset-free control of both steam and O_2 to be feasible or useful and accept a loss in overall economic performance in case of an undercapacity market situation.

In the remainder of this chapter, the same type of controller and control problem as discussed in this section are considered, *i.e.* with the same MVs, CVs and setpoints as in figure 5.1. The existing controller to be compared with the new one is commercially available and is for reasons of confidentiality not outlined here in detail. It has, however, a much different structure than the new control strategy to be proposed here. Also, its properties are considered representative for other PID combustion control strategies currently in use and includes no offset-free control of the O_2-concentration. The disturbance rejection and setpoint tracking properties of both the new and existing PID controller are compared and the results from that are translated to conclusions on the opportunities for improving the overall economic performance of MSWC plants with the new controller and, additionally, with PID based combustion control in general.

5.3 The dynamics of MSWC plants

5.3.1 Introduction

The dynamics of MSWC plants is discussed here via step responses of steam and O_2 obtained with the system identification and first-principles model considered in the application discussed in section 4.2. These step responses are, hence, valid representations of the dynamics of the considered large scale Dutch MSWC plant at the operating point given in table 4.1. Also, the dynamics of the disturbances acting on MSWC plants are discussed here, in particular through step responses of a linear black-box time series model obtained also from the data used at the application discussed in section 4.2.

It is noted that the dynamics at another operating point of the considered MSWC plant and the dynamics at other such plants, as observed from other applications of the system identification and first-principles modeling procedures proposed in this thesis to large scale MSWC plant data, were found to exhibit different time constants and static gains in comparison to those of the dynamics to be discussed here but were also found to be of the same global shape as that of these dynamics. As a consequence, the PID combustion control system proposed here, which structure is based on this global dynamic shape rather than the specific time constants and static gains, also can be used at other operating points and plants, though with different controller settings.

All dynamics discussed here are used to form a new PID MSWC plant combustion controller in section 5.4.

5.3.2 Plant dynamics

Responses of steam and O_2 on step on ram speed

See figure 5.2. It is observed from this figure that an increase in the ram speed (/waste inlet flow) results, at steady state, in an increase in steam and a decrease in O_2. This behavior is explained from the fact that an increase in ram speed is accompanied by an increase in combusting fuel (*i.e.* waste) on the grate, which results in a higher conversion of the waste energy content into steam but also in a higher O_2 consumption. It can also be seen that steam and O_2 reach their steady state in approx. 100 minutes, which is a typical (dominant) time constant for large scale MSWC plants. Also, before steam rises, it first approximately remains constant for about 10 minutes. Similarly, O_2 remains approximately constant for several minutes before decreasing to its steady state value. In fact, these transfer functions are known to typically exhibit an inverse response (see *e.g.* [67]) due to evaporation consuming thermal energy resulting from the waste chemical energy conversion before combustion takes over and causes an effectively positive thermal energy output *c.q.* steam production. The delays here can

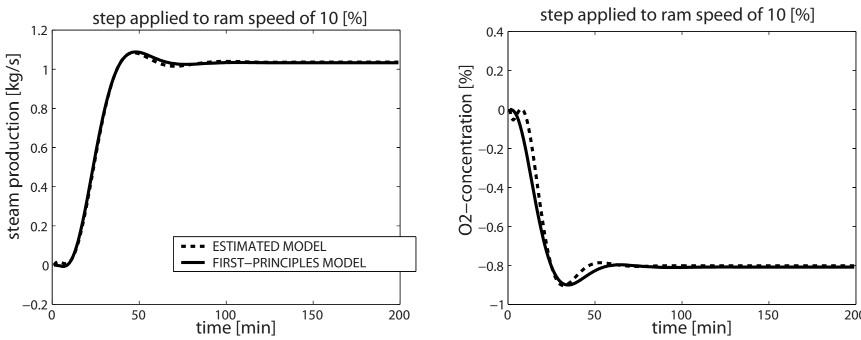

Figure 5.2: Responses of steam and O_2 to step on ram speed.

be interpreted as (almost) flat inverse responses with an explanation for this flatness being that the evaporation has only a small effect due to a relatively high primary air flow temperature. Note that the steam and O_2 responses are, more or less, mirrored versions of each other, with the time axis being the mirror axis.

Responses of steam and O_2 on step on grate speed

See figure 5.3. One can observe in both steam and O_2 a temporary large effect but no change in steady-state value. An explanation for the latter is that a change in grate speed does not deliver new fuel or O_2 to the furnace. The temporary large effect can be attributed to a temporary increase in effective area over which the air reaches the combustible part of the waste. This causes a temporary increase in combusted waste and, as a result, a temporary increase in energy output and a temporary decrease in in O_2. This effect is similar to the the so called poke-effect, *i.e.* the effect observed when

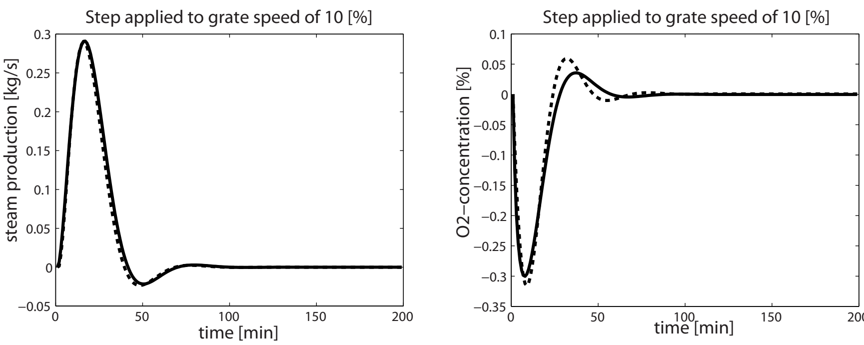

Figure 5.3: Responses of steam and O_2 to step on grate speed.

poking with a stick in the fire of the stove back home. Note that, again, steam and O_2 are, more or less, mirrored versions of each other with the time axis as the mirror axis.

Responses of steam and O_2 on step on primary air flow

See figure 5.4 for the primary air flow step responses. An eventual net increase in

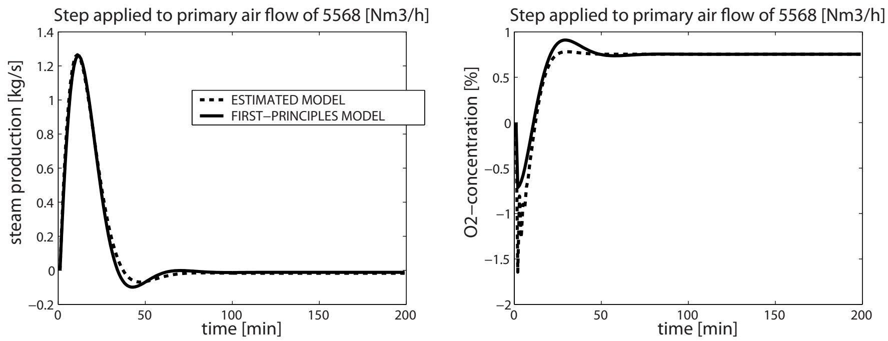

Figure 5.4: Responses of steam and O_2 to step on primary air flow.

O_2 is observed, preceded by a sharp inverse response. The steam response exhibits a large upset similar to those encountered at the grate speed responses, before reaching a slightly positive steady state value. The positive sign of this value is the result of an increased heat flow to the boiler. This increased heat flow is due to an increase in flue gas flow induced by the increase in primary air flow. At the same time, notably, the furnace (*i.e.* waste layer and flue gas) temperatures drop due to the cooling effect of the primary air flow (which temperature is lower than those in the furnace). Nevertheless, the heat flow to the boiler, which is a function of the product of the flue gas flow and temperature, increases as the increase in flue gas flow is larger than the decrease in its temperature. The size of the steam production steady-state value depends on the specific combustion conditions and the boiler efficiency and typically is small. Here, as can be observed from figure 5.4, this steady-state value is apparently very small. This low steam production steady-state value is an important assumption to be fulfilled for a proper operation of the new PID combustion control strategy to be presented here.

Responses of steam and O_2 on step on secondary air flow

The responses of steam and O_2 to a step applied to the secondary air flow are to a large extent similar to those for the primary air flow though with much lower amplitude. See figure 5.5.

5.3.3 Disturbance dynamics

For the discussion of the disturbance dynamics of MSWC plants, step responses of a black box (discrete-time linear time-invariant auto-regressive type of) time series disturbance model are used. This model was estimated in a similar way as the earlier

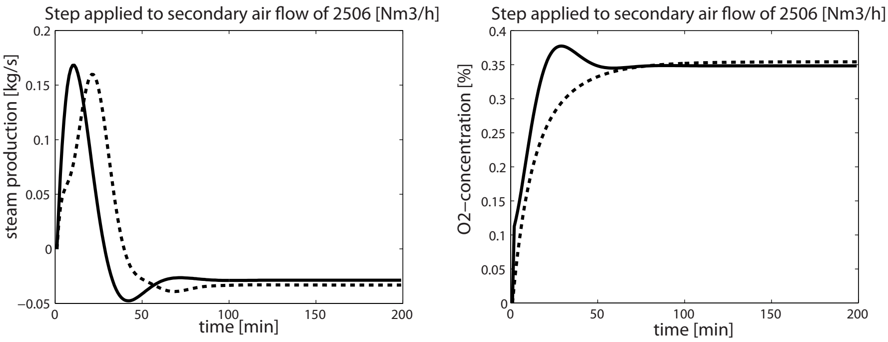

Figure 5.5: Responses of steam and O_2 to step on primary air flow.

discussed black-box model from closed-loop data experimentally obtained from a large scale Dutch MSWC plant. Correlation functions were (again) used for validation and indicated that the model is of good quality. With this model the disturbances can be simulated that are acting on steam and O_2 as a function of (two) inputs that can be interpreted as the (white noise) stochastic sources for these disturbances, though without having any physical meaning. When a (unit) step is applied to one of the inputs of the model, steam and O_2 respond as depicted in figure 5.6. The main thing to note here

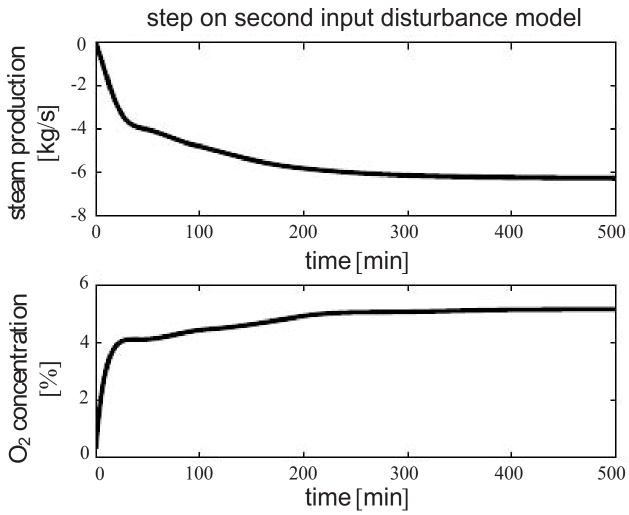

Figure 5.6: Step responses of time series disturbance model.

is that the responses are, more or less, mirrored, again. In fact, from simulations with the model, it was found that this type of disturbances accounts for most of the disturbances acting on steam and O_2: approx. $74\ \%$ for steam and $93\ \%$ for O_2, in variance sense. Hence, suppression of these disturbances would mean suppression of most of the MSWC plant disturbances, which observation is the main pillar for the disturbance rejection part of the new control strategy to be proposed here.

The most likely physical source of the mirror disturbances discussed here is the waste composition as a step applied to this disturbance also leads to a mirrored response of steam and O_2, as confirmed by simulations with the first-principles MSWC plant model proposed in chapter 2.

5.4 A new PID combustion control strategy

5.4.1 Introduction

The controller proposed here is derived from a closer study of the dynamics discussed above. The following four main parts can be distinguished:

- A part dealing with offset-free setpoint tracking of both steam and O_2
- A part for suppressing the disturbances by manipulating the grate speed
- A part for suppressing the disturbances by manipulating the primary and secondary air flow
- A (small) part steering the grate speed setpoint

Each of these four parts will now be motivated and discussed in more detail. Figure 5.7 presents a schematic overview of the resulting MSWC plant PID combustion control strategy.

5.4.2 Offset-free setpoint tracking of steam and O_2

From the step responses of the steam production given above, see figures 5.2-5.5, one can deduce that only the ram speed has a large static effect on this CV. This implies that the ram speed is effectively the only candidate for steering steam to its setpoint, whereas the usage of other MVs for that purpose leads to excessively large MV amplitudes. To obtain offset-free setpoint tracking for steam, a PI controller is proposed to connect, on the one hand, the error between steam and its setpoint and, on the other hand, the ram speed. The integral action in this steam-ram speed control loop ensures offset-free control of the steam production. Note that the PI controller cannot be tuned too tight when a significant inverse response is present in the ram speed-steam transfer function, as this introduces undesired oscillations. Compensating for this inverse response will lead to a loss in control performance in the sense of a slower closed-loop response for this control loop. This loss will, however, not be too severe as the closed-loop response is slow anyway due to the slow dynamics involved in the ram speedsteam transfer function.

When using only this PI steam-ram speed control loop, O_2 cannot be steered independently to its setpoint. Rather, its steady-state value depends on the steady-state value for the ram speed determined by this loop. In order to also obtain offset-free control of O_2, one needs to use one or more additional MVs that have a significant static

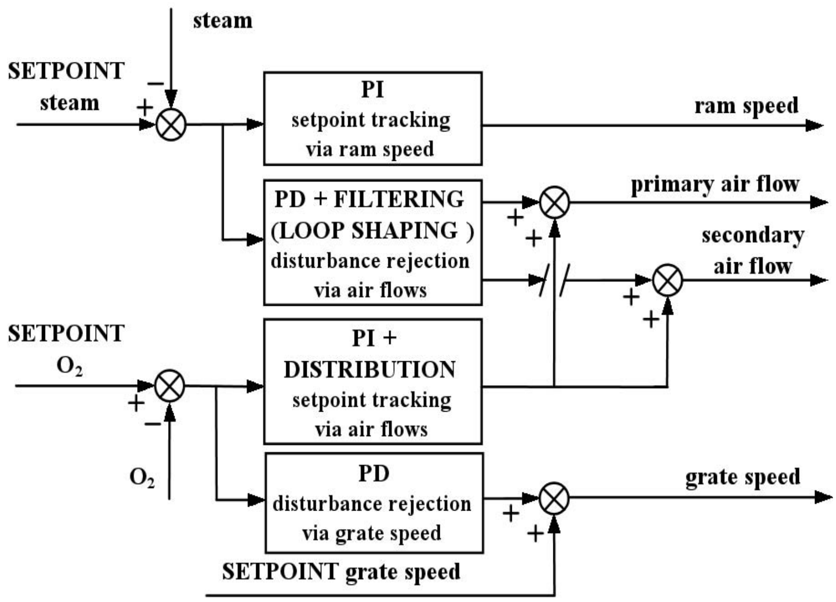

Figure 5.7: The new PID combustion control strategy.

influence on this CV. From the step responses discussed above, see figures 5.3-5.5, one can see immediately that this leaves only the two air flows as candidate MVs as the grate speed - O_2 transfer function, see figure 5.3, has a zero static gain. Note that a control loop containing these MVs would hardly interfere with the already discussed steam - ram speed control loop because of the low static gains of these MVs towards steam. The proposed O_2 - air flows control loop(s) can be implemented by means of one PI controller acting on the error between measured O_2 and its setpoint to steer the air flows according to some user-defined distribution.

5.4.3 Disturbance rejection via the grate speed

At the discussion of the MSWC plant disturbance dynamics it was observed that the major part of the disturbances acting on steam and O_2 are of the mirror type. It was also observed that these CVs respond in a mirror way to manipulation of the grate speed. See figure 5.3. Because of this, the grate speed MV is a good candidate for simultaneously rejecting this type of disturbances in both CVs. More specific, it is an ideal MV for the reduction of the middle and higher frequency mirror disturbances due to the small effect this MV has in the low frequency range. The grate speed control loop proposed here may be implemented either using the error between steam and its setpoint or between O_2 and its setpoint to steer, via some PD action, the grate speed. The choice for the O_2 error has the advantage of allowing the suppression of the highest frequency disturbances. This is due to the disturbances acting on O_2 having a larger bandwidth than those on steam, which can be subscribed to the low pass filtering effect of the boiler dynamics on the steam signal. However, one may also choose to suppress these highest frequency disturbances by means of the air flows disturbance rejection loop to be discussed below. Here, arbitrarily, the O_2 error has been chosen for the grate speed disturbance rejection loop and the steam error has been chosen for the air flows disturbance rejection loop.

5.4.4 Steering the setpoint of the grate speed

As the grate speed has no static effect on steam and O_2, the control loop for disturbance rejection via this MV discussed just above also has no static effect on these CVs. As a result, the setpoint of the grate speed is an extra degree of freedom that can be used to control, in a static or low frequency manner, the combustion process in another way than via steam and O_2, *e.g.* to keep the fire properly positioned on the grate. This may be done manually or automatically, *e.g.* through a camera monitoring the waste on the grate.

5.4.5 Disturbance rejection via the air flows

With the grate speed disturbance rejection control loop one is already able to obtain a significant disturbance rejection. The sole usage of this control loop may lead, though, to an unacceptably large variation in the grate speed. A solution to that is to also use the air flows for disturbance rejection. In that way, the energy in the MVs resulting from disturbance rejection may be spread over both the grate speed and the air flows, thereby reducing the variation in the first of these MVs. One particular way of implementing the air flows disturbance rejection control loops is by using a PD controller in combination with loop shaping. The idea here is to shape, via (pre-)filters, the transfer functions from the primary and secondary air flow towards steam and O_2 such that, when given the same MV input, the resulting responses of steam and O_2 are, as much as possible, mirrored with respect to each other. See figure 5.8 for an example result. A PD controller then simultaneously manipulates the input of the resulting pre-filtered transfer functions, as if it were an MV that has a mirroring effect on the two CVs, in response to either the steam or O_2 error signal. The aim of the resulting control loop is then to counteract the mentioned mirror type of disturbances. As mentioned above, the steam error signal has arbitrarily been chosen here as the input for the control loop while the O_2 error signal has been chosen for the grate speed disturbance rejection control loop. A final important note here is that proper operation of the loop shaping control loop proposed here depends on the availability of accurate air flow transfer functions, which may not always be the case.

5.5 Performance assessment

Via simulations on system identification and time series MSWC plant resp. disturbance models the new controller was compared with the existing, commercially available one. These simulations and comparison were performed for both disturbances of a stationary stochastic nature and for temporarily large upset type of disturbances. The reason for

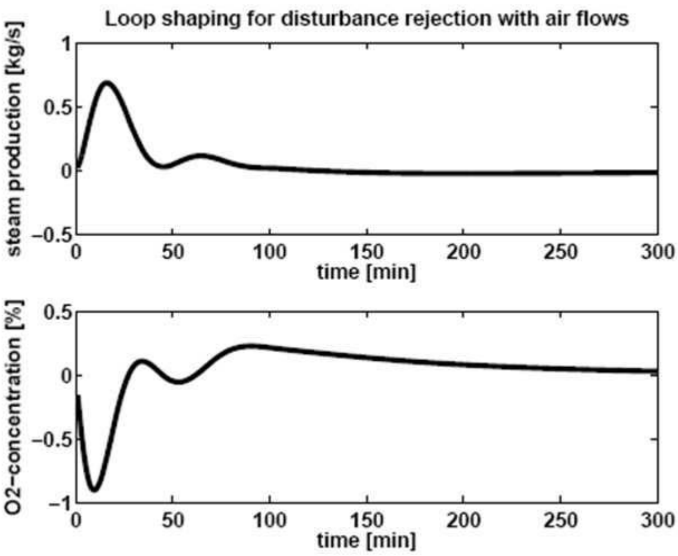

Figure 5.8: An example result of loop shaping: mirrored CV responses to same step applied to pre-filtered air flow inputs.

that was that these have been identified as two of the most commonly encountered types of disturbances acting on MSWC plants (see figure 3.6). It is noted that, by lack of a systematic tuning procedure, both considered controllers were tuned via trial-and-error.

In figures 5.9 and 5.10 typical simulation results are shown that were obtained from the comparison. Figure 5.9 depicts closed-loop responses obtained with the new and existing controller for the case that disturbances of a stationary stochastic nature are acting on the MSWC plant. It clearly can be seen that, as expected, the new controller is capable of offset-free setpoint tracking of both steam and O_2, whereas the existing controller is not (which is only capable of offset-free setpoint tracking of steam). From figure 5.9 it can also be seen that both controllers are capable of significant disturbance rejection: compare the variation in steam and O_2 before and after the controller is activated at $t = 25000\ [min]$. In fact, the simulation results showed that both controllers exhibit the same disturbance rejection performance, with a recorded maximum reduction in standard deviation of $70\ \%$ for steam and $30\ \%$ for O_2, implying that no significant improvement can be made with the new controller in disturbance rejection sense. As an example simulation result, standard deviations for this maximum reduction case are summarized in table 5.1.

Figure 5.10 depicts closed-loop responses obtained with the new and existing controller for the case that a temporarily large upset type of disturbance is acting on the MSWC plant. Note that both controllers perform similarly in response to this type of disturbance except from an O_2 setpoint tracking (offset-free control) performance point of view.

Simulations with the new and existing combustion controller on other MSWC plant

	open-loop	*PID current*	*PID new*
STD(ϕ_{steam}) $[kg/s]$	0.81	0.23	0.23
STD(O_2) $[\%]$	1.07	0.72	0.72
STD($waste\ inlet\ flow$) $[\%]$	0	6.19	6.94
STD($speed\ of\ grate$) $[\%]$	0	15.36	15.28
STD($\phi_{prim.air}$) $[Nm^3/h]$	0	4951	4825
STD($\phi_{sec.air}$) $[Nm^3/h]$	0	1238	1192

Table 5.1: Open and closed-loop standard deviations of the MVs and CVs. Closed-loop results for both current and new PID MSWC plant combustion control strategy and for stationary stochastic type of disturbances.

and disturbance models led to similar results as discussed above, indicating both (i) the robustness of the new controller against variability in plant dynamics and (ii) that the disturbance rejection properties of the existing PID combustion controllers cannot be improved significantly through the same type of controller. Rather, to improve on the disturbance rejection properties, other, non-PID, type of combustion controllers are required.

The results here also indicate that in an overcapacity market situation, where the focus is on process variation minimization, an improvement in overall economic MSWC plant performance is likely to be small when using the new PID combustion controller as this controller is not able to deliver a significantly improved disturbance rejection performance and offset-free CV following properties may not have a large effect on this performance (see also the discussion in section 5.2). In an undercapacity market situation, however, the improved setpoint tracking properties of the new controller do allow for a signficant improvement in overall economic MSWC plant performance. More specific, the usage of this controller prevents the need for decreasing the steam setpoint, and thereby the average waste throughput and steam production itself, to a much lower value to prevent O_2 from violating its law enforced lower bound ($= 6\ \%$) too many times.

5.6 Conclusions

A new PID-type of combustion control strategy for MSWC plants has been derived from a closer investigation of the MSWC plant dynamics, as exhibited by recently derived black and white box models. This new control strategy has improved setpoint tracking properties compared to PID combustion control strategies typically encountered in the industry, but also equal disturbance rejection properties. Due to the improved setpoint tracking properties, the usage of the new controller can lead to a substantially improved overall economic performance for MSWC plants when these plants are subject to an undercapacity market situation and, consequently, the combustion controller is used in a constrained maximization/constraint pushing manner. To improve upon the disturbance rejection properties of the currently used PID combustion constrol strategies, non-PID type of combustion control strategies are required.

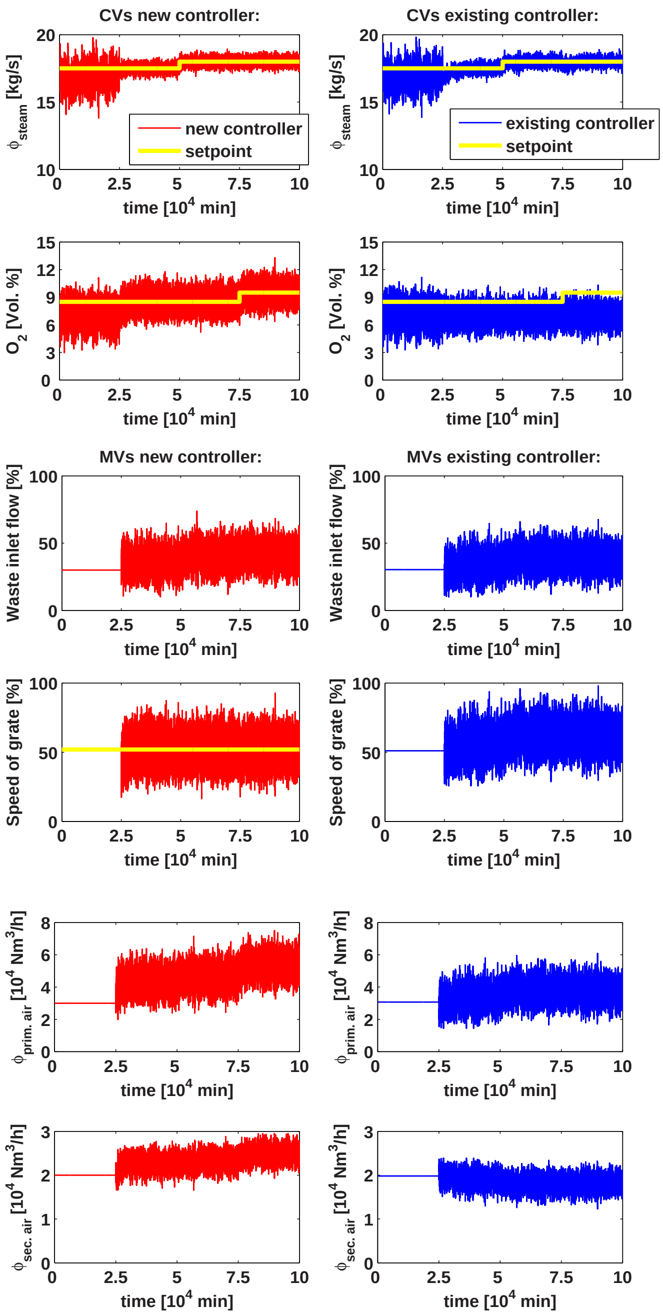

Figure 5.9: New (left) versus existing (right) controller: response to the same stationary stochastic type of disturbance. Upper two rows: CVs. Lower four rows: MVs.

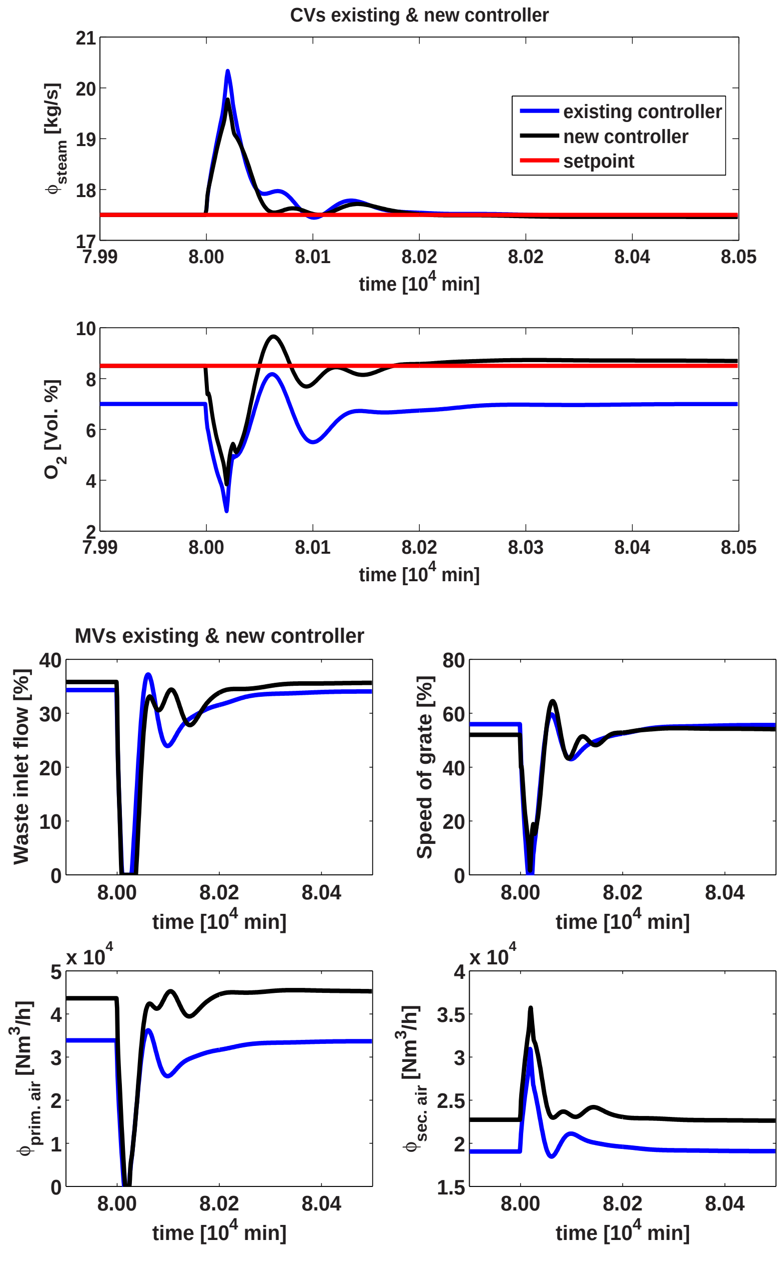

Figure 5.10: New versus existing controller: response to the same temporarily large upset type of disturbance. Upper part: CVs. Lower part: MVs.

Chapter 6

A linear and nonlinear model predictive control strategy for MSWC plants

6.1 Introduction

In this and the next chapter, results are presented from the exploration of the opportunities for improving the combustion control performance of MSWC plants through model predictive control (MPC), given a suitable model of the MSWC plant combustion process. In this chapter, the focus is on implementational issues related to this control strategy, when applied to MSWC plant combustion control, while (actual) performance improving opportunities are discussed in the next chapter. More specific, important MPC related implementational issues are identified and solutions to these issues are presented. In particular, implementational issues of NMPC based MSWC plant combustion control strategies are considered as these are the most challenging to resolve. The results are both an LMPC and NMPC strategy that have been found suitable for MSWC plant combustion control. The main contribution(s) of this chapter are these control strategies and the solutions proposed to the implementational issues.

The contents of this chapter are as follows. First, in section 6.2, the global layout is presented of both the linear and nonlinear MSWC plant MPC strategy proposed here. This section also provides a more general introduction into MPC, both of the linear and nonlinear type. Then, in section 6.3, the nonlinear variant of the proposed control strategies is elaborated through presenting solutions to identified main implementational issues for this variant. Subsequently, in section 6.4, the linear variant is presented. Finally, in section 6.5, the conclusions of this chapter are given.

6.2 Global layout of the model predictive control strategies

6.2.1 Introduction

To be able to provide the global layout of the MSWC plant MPC strategies to be proposed here, the main idea behind MPC and the main implementational aspects of this control strategy are discussed first. The proposed global layout follows then from making specific choices with respect to these implementational aspects.

Starting point here is the availability of a low order sufficiently accurate MSWC plant model of the form

$$\begin{aligned} x(k+1) &= f(x(k), u(k), d_m(k), d_{nm}(k), \theta) \\ y(k) &= h(x(k), u(k), d_m(k), d_{nm}(k), \theta) \end{aligned} \quad (6.1)$$

This specific format is chosen here because it agrees with the sampled-data nature of MSWC plant controller implementations, with k representing discrete-time, and because it incorporates both the first-principles type of MSWC plant model given by eqns. (2.45), *i.e.* after integration over a sample time T_s under *e.g.* zero-order-hold conditions, and linear MSWC plant models obtained through system identification. As a consequence, the global layout proposed here applies to both the linear and nonlinear MSWC plant MPC strategy to be proposed here.

It is noted that T_s is most of the times chosen equal to 60 $[s]$ at MSWC plant control applications and that this value is also chosen here as it typically is sufficiently fast for capturing the fastest MSWC plant dynamics.

6.2.2 The main idea behind model predictive control

MPC refers to a class of control strategies that explicitly use predictions of the model of the plant to be controlled for the computation of the MVs for this plant. In contrast, other types of control *implicitly* use model predictions to determine the MVs, *i.e.* via a fixed control law into which the dynamic relations between these predictions and the MVs are condensed. MPC translates its model predictions to MVs by solving an optimization problem known as an *optimal control* problem (which is a special type of *dynamic optimization* problem). This optimal control problem typically is a special case or closely related version of

$$
\begin{array}{rll}
\min_{u(0),\ldots,u(N_p-1)} & V(u(0),...,u(N_p-1)) = & \\
& \sum_{i=0}^{N_p-1} l_{stage}(x(i),u(i),d_m(i),d_{nm}(i),\theta) + l_{terminal}(x(N_p)) + & \\
& \sum_{i=0}^{N_p-2}(u(i+1)-u(i))^T W_u(i)(u(i+1)-u(i)) & \\
s.t. & x(i+1) = f(x(i),u(i),d_m(i),d_{nm}(i),\theta); & \forall\ i = 0...(N_p-1); \\
& & x(0) = x_0 \\
& g_{stage}(x(i),u(i),d_m(k),d_{nm}(k),\theta) \leq 0; & \forall\ i = 0...(N_p-1) \\
& g_{terminal}(x(N_p)) \leq 0; & \\
& h_{stage}(x(i),u(i),d_m(k),d_{nm}(k),\theta) = 0; & \forall\ i = 0...(N_p-1) \\
& h_{terminal}(x(N_p)) = 0; & \\
& u(i+1) = u(i); \qquad \forall i = N_c...(N_p-2) &
\end{array}
\tag{6.2}
$$

with $U(N_p) := \{u(0),...,u(N_p-1)\}$ the MV trajectory to be computed for given initial state x_0 and disturbance trajectories $D_m(N_p) := \{d_m(0),...,d_m(N_p-1)\}$ and $D_{nm}(N_p) := \{d_{nm}(0),...,d_{nm}(N_p-1)\}$. Here, N_p is called the *prediction horizon* and N_c the *control horizon, i.e.* the time horizon over which the optimization effectively takes place (over the remainder of the input trajectory the inputs are held constant). The matrices $W_u(i)$, $i = 0\ ...\ N_p-1$, represent input weighting matrices (typically chosen constant over time i) that allow for suppression of the variations in the MVs in a user-defined manner. The objective function and constraints of the optimal control problem (6.2) reflect the control objectives for the plant to be controlled. MSWC plant combustion control related examples of optimal control problem formulations are given later on in this chapter.

For ease of explanation, the optimal control problem formulation (6.2) considered here is of a generic form. However, depending on the specific form of model, constraints and objective function, different MPC forms are distinguished in the literature. The motivation for that are the corresponding theoretical and implementational differences. In particular, linear and nonlinear MPC (LMPC and NMPC) are distinguished. With LMPC the model and constraints that are employed are linear in the MVs and the objective function is linear or quadratic in these variables. With NMPC the model and/or constraints are nonlinear and/or the objective function is nonlinear and nonquadratic. In this thesis, both an LMPC and NMPC based MSWC plant control strategy are proposed with the main difference being the usage of a linear resp. a nonlinear model.

Irrespective the specific form of MPC, this control strategy proceeds by first determining the values for x_0, $D_m(N_p)$ and $D_{nm}(N_p)$ and then, given these values, by computing the optimal MV trajectory $U(N_p)$ via solving the optimal control problem. The resulting MV trajectory is subsequently implemented on the plant to be controlled until at a later time instant, typically the next sampling instant, a new optimal MV trajectory is implemented that has been computed along the same steps. This repetition of steps is performed to handle uncertainties *c.q.* errors in x_0, $D_m(N_p)$ and $D_{nm}(N_p)$ and in the model equations $f(.)$ and, when used, $h(.)$. Because of these errors, one-time application of the mentioned steps would eventually result in a large deviation of the model state $x(k)$ from its true plant counterpart and, thereby, in a large performance degradation and possibly instability. The repetition of the mentioned steps ensures that the model state remains as close as possible to the actual plant state and is also refered to as *receding* or *moving horizon control* (MHC).

6.2.3 Implementational aspects

The ideal MPC implementation

The ideal MPC implementation would involve

- perfect information on the current state vector x_0
- perfect information on the current and future disturbances $D_m(N_p)$ and $D_{nm}(N_p)$
- the solution $U(N_p)$ to be immediately available and implemented on the plant to be controlled (in particular the first element $u(0)$ of $U(N_p)$)
- the prediction and control horizons to be infinite
- the MPC control strategy optimally taking into account the presence of feedback in its future predictions

However, this ideal implementation is not or at most only partly encountered at real-life MPC applications due to the presence of *e.g.* unresolvable obstacles or of more easy to implement, though less optimal, alternatives. Here, each of the mentioned aspects of the ideal implementation is discussed in more detail and, briefly, alternative solutions are outlined that allow to overcome corresponding obstacles or that allow for a more easy implementation. Additionally, one of the considered solutions is chosen as part of the MSWC plant MPC strategies to be proposed here. Together, these choices form the basic layout of these strategies, which for ease of reference is summarized in section 6.2.4.

State feedback versus output feedback MPC

Perfect information on x_0 may not be available. In particular, the values of this state vector may not be measured or, when these are measured, the corresponding measurements may be very noisy. The solution that is applied to overcome this problem is to infer the uncertain state values from measured data using a *state estimator*. It is noted that a control strategy with perfect measurements of the state vector is called a *state feedback* control strategy whereas its state estimator counterpart is denoted as an *output feedback* control strategy because of the usage of output (CV) measurements.

The MSWC plant MPC control strategies considered here employ a state estimator because of the absence of perfect state measurements, not only because these measurements are simply not available or are noisy but also because of the computional delays typically present in MPC control strategies. The latter is elucidated further on in this section.

Handling present and future disturbances

Future random disturbances cannot perfectly be predicted and, hence, no perfect information on the corresponding elements in $D_m(N_p)$ and $D_{nm}(N_p)$ is available. Additionally, the (only remaining) current disturbance values $d_m(0)$ and $d_{nm}(0)$ of $D_m(N_p)$ resp. $D_{nm}(N_p)$ are either not measured or may be contaminated with a severe level of noise.

The standard and simple solution applied to overcome the problem of no perfect future disturbance predictions is that of simply assuming these equal to the current values $d_m(0)$ and $d_{nm}(0)$, thereby accepting the resulting loss in control performance. More systematic ways of incorporating future disturbances within MPC are currently being investigated, see *e.g.* [17], but are not commonly applied yet.

The absence of perfect measurements for $d_m(0)$ and $d_{nm}(0)$ is typically overcome by using, again, a state estimator. For that purpose, the state vector of the dynamic model of the plant to be controlled ($x(k)$ in (6.1)) is typically augmented with the uncertain disturbance variables using a description of the disturbance dynamics.

The MSWC plant MPC control strategies considered here employ the assumption of constant future disturbances and a state estimator for inferring the values for $d_m(0)$ and $d_{nm}(0)$ from measured data. The usage of a state estimator here is motivated by the inevitable presence of noise, absence of measurements and, in particular also, the presence of computational delays, as with x_0 and as elucidated below. Dynamic models are incorporated in this state estimator that describe the dynamics of the measured and nonmeasured disturbances to allow for the inference of $d_m(0)$ resp. $d_{nm}(0)$. The specific type of these *disturbance* models is discussed further on in this chapter.

Handling of computational delays

MPC control strategies are typically subject to computational delays, *i.e.* to nonzero computation times in the solution of the optimal control problem and/or in the state estimation problem. As a consequence, the solution for $U(N_p)$ cannot be immediately implemented at the same sampling instant k_0 at which the state estimation and/or optimal control problem computations commence and needs to be implemented at a later sampling instant (*i.e.* in a sampled-data setting), thereby inevitably leading to a performance degradation. For example, $U(N_p)$ can be implemented at $k_0 + 1$ in case the sum of computational delays does not exceed one sampling period. However, this solution implicitly assumes that x_0 obtained for k_0 is the best prediction for x_0 at $k_0 + 1$. A better solution *c.q.* a smaller degradation in MPC performance would be obtained when using a better prediction for the latter state vector. Such an improved prediction can be obtained with a state estimator.

The MSWC plant MPC control strategies considered here contain a non-zero sum of computational delays that does not exceed one sampling period. To overcome the presence of these delays, these strategies employ a state estimator that, given the available plant data up till and at time k_0, provide an optimal prediction of x_0 at the next sampling instant $k_0 + 1$. Before this sampling instant is reached, the proposed MPC strategies have computed the optimal solution $U(N_p)$ for this predicted state vector which is directly implemented at this next sampling instant.

It is noted that MPC control strategies without computational delays do exist. These consist of one or more *a priori* calculated fixed control laws. Two types of such MPC strategies are presently available but these are either of limited use or, having only recently been developed, are not commonly applied yet. More specific, a first such type of MPC control strategy is the well known linear quadratic regulator (LQR) or linear quadratic Gaussian (LQG) controller (see *e.g.* [93]), which employs an *a priori* calculated control law and which use is restricted due to the assumptions of a linear model, quadratic objective function and absence of constraints. A second type of MPC control strategy without computational delays is *multi-parametric* or *explicit* MPC (see *e.g.* [23]), which recently has been developed and involves the *a priori* computation of several fixed control laws (each for a specific state space region).

Infinite horizon MPC

With respect to the choice for the values of N_p and N_c, the best MPC performance would be obtained with infinite values for these time horizons as the corresponding control strategy would then take into account the future plant behavior up till the end of time. Also, the resulting control strategy would have good stability properties, *i.e.* it would at least be stabilizing for the state feedback case and in the presence of a perfect plant model (see *e.g.* [72]). However, the usage of infinite time horizons would lead to infinitely long computation times and, hence, a computationally infeasible MPC control strategy. Therefore, the values for N_p and N_c are given finite values. It is noted that a feasible infinite horizon MPC strategy can be obtained in the case of a linear model, quadratic objective function and the absence of constraints as then an analytical solution can be obtained for the MPC control strategy, *i.e.* in the form of an LQR or LQG controller. The MSWC plant MPC control strategies considered here, however, are not of this type, *i.e.* use at least a nonlinear model and/or take into account constraints, and therefore employ finite values for N_p and N_c.

Open- versus closed-loop MPC

Typically, for ease of implementation, the presence of feedback is not taken into account in the solution to the MPC optimal control problem, *i.e.* in the corresponding prediction of the future plant behavior. This leads to a degradation in control performance as feedback is inevitably present in the future plant behavior due to the presence of the MPC control strategy itself. Recently, MPC formulations have been proposed and investigated that do take into account this presence of feedback. See *e.g.* [5] and [17]. Such MPC strategies are refered to as *closed-loop* MPC strategies. However, for ease of implementation and because of the already acceptable resulting control performance, the MSWC plant MPC control strategies considered here are of the standard open-loop type.

6.2.4 Summary of the global layout

The MSWC plant MPC strategies proposed here follow a standard moving horizon control strategy where the MVs are computed by repeatedly solving a finite-horizon open-loop optimal control problem for newly determined plant state and disturbance values. A state estimator is employed to predict the values for these states and disturbances for the next sampling instant $k_0 + 1$ from measurements available until and at the current sampling instant k_0. The optimal control problem is subsequently solved for these values with the future disturbance values, *i.e.* for $k > k_0 + 1$, chosen constant and equal to the predicted ones at $k_0 + 1$. At the latter sampling instant, the resulting optimal MV trajectory is then implemented on the plant to be controlled until at the next sampling instant a new MV trajectory is implemented that has been computed for newly predicted state and disturbance values.

6.3 A nonlinear model predictive control strategy for MSWC plants

6.3.1 Introduction

In this section, the outline of the MSWC plant NMPC strategy proposed here, which global layout has been provided in the previous section, is completed. In particular, solutions are presented to the following main implementational issues identified for this control strategy:

- the formulation of the optimal control problem
- the numerical solution to the optimal control problem
- the solution to the state and disturbance estimation *c.q.* prediction problem
- the enforcement of offset-free control in the presence of model error and/or persistent unmodeled disturbances
- the enforcement of stability on the resulting closed-loop MSWC plant

From the presented solutions, specific ones are chosen that are considered optimal for the MSWC plant NMPC strategy proposed here, with optimality here referring to an optimal compromise between control performance (degree of fulfillment of combustion control objectives and constraints), computational speed and ease of implementation. For ease of reference, the resulting strategy is summarized at the end in section 6.3.7. In fact, this NMPC strategy is fairly generic and may well be applicable to many other processes or systems for which a low order sufficiently accurate dynamic model is available and for which the sampling time is in the order of a minute. A final note here is that the discussions in this section also provide a state-of-the-art of NMPC, albeit limited.

6.3.2 Formulation of the optimal control problem

As mentioned in section 1.2.1, the aims (objectives and constraints) to be fulfilled by an MSWC plant combustion control system are to maximize waste throughput and steam production (to maximize the economic income), to minimize the variations in various process variables (to minimize operational and maintenance costs), and to fulfill constraints imposed out of environmental and (again) maintenance considerations. A preferred option is to collect all these aims into one, generic, NMPC optimal control problem formulation. Such a formulation can take into account any type of MSWC plant market situation, in particular the over- and undercapacity one identified earlier (see *e.g.*, again, section 1.2.1), each of which corresponding to (particularly) different weighting factors in the optimal control problem objective function. In fact, this ability of using one and the same control strategy for dealing with multiple market situations may be regarded an important benefit of MPC type of MSWC plant combustion control strategies as it allows for a relatively easy and, thereby, cheap adaptation of the plant operation to different market situations.

Rather than attempting here to define a fully comprehensive MPC optimal control problem formulation for MSWC plant combustion control applications, two separate more easy to implement non-generic such formulations are proposed and discussed here, one specifically aimed at handling the overcapacity market situation and another one aimed at handling the undercapacity market situation. These formulations may be viewed as 'asymptotic'/extreme versions of the comprehensive one. Vice versa, the comprehensive formulation may be obtained by combining elements of the separate two formulations, which are discussed now.

In the first market situation of *overcapacity* the priority is on cost reduction which typically translates to the objective of minimizing process (variable) variations in variance/standard deviation sense, in particular those present in steam production and O_2-concentration with respect to their setpoints, and in the sense of maintaining process variables within bounds, *e.g.* temperatures below upper limits to increase the lifetime of furnace components. Violation of such limits may be due to *e.g.* the temporarily large disturbances discussed in section 3.6.2. Ignoring so far the issue of choosing $l_{terminal}(.)$, $g_{terminal}(.)$ and $h_{terminal}(.)$, see for that further on in this section, the NMPC optimal control problem could be formulated for this market situation as in (6.2) with

$$l_{stage}(.) = W_{\phi_{st},varmin}(\phi_{st}^{sp} - \phi_{st}(i))^2 + W_{O_2,varmin}(O_2^{sp} - O_2(i))^2 + ... \quad (6.3)$$

or, in a more generic form,

$$l_{stage}(.) = (y^{sp} - y(i))^T W_y (y^{sp} - y(i)) + (x^{sp} - x(i))^T W_x (x^{sp} - x(i)) \quad (6.4)$$

where y^{sp} and x^{sp} represent setpoints and W_y and W_x weighting matrices, with possibly terms of the quadratic form

$$\begin{aligned}\sum_i (y(i+1) - y(i))^T W_{\Delta y}(i)(y(i+1) - y(i)) + \\ \sum_i (x(i+1) - x(i))^T W_{\Delta x}(i)(x(i+1) - x(i))\end{aligned} \quad (6.5)$$

added to the objective function ($V(.)$) to account for (further) process variable variation minimization objectives, and with

$$h_{stage}(.) = y(i) - h(.) \tag{6.6}$$

with $h(.)$ as defined in eqn. (6.1), and with

$$g_{stage}(.) = \begin{bmatrix} y(i) - y_{max} \\ y_{min} - y(i) \\ x(i) - x_{max} \\ x_{min} - x(i) \\ u(i) - u_{max} \\ u_{min} - u(i) \end{bmatrix} \tag{6.7}$$

defining environmental and maintenance related constraints (with y_{max}, y_{min}, x_{max}, ... representing the actual limits not to be violated).

In the second common MSWC plant market situation of *undercapacity*, the aim is to maximize the income through maximizing the steam production and waste throughput until a constraint is approximated to a minimum *c.q.* closest allowed level. For this market situation, ignoring again still the choices to be made for $l_{terminal}(.)$, $g_{terminal}(.)$ and $h_{terminal}(.)$, the NMPC optimal control problem may be formulated as above for the first formulation but with a different stage cost

$$l_{stage}(.) = -W_{\phi_{st},max}\ \phi_{st}(i) - W_{\phi_{w,in},max}\ \phi_{w,in}(i) + ... \tag{6.8}$$

or, in a more generalized setting,

$$l_{stage}(.) = -W_y^T y(i) - W_x^T x(i) - W_u^T u(i) \tag{6.9}$$

with the weightings W_y, W_x and W_u now representing column vectors rather than matrices as in the first formulation and with, notably, the minus signs rendering the minimization problem effectively a maximization problem.

It is noted that, in order to take into account the possibility of frequent violation of the constraints due to disturbances, model error, delays and other sources of imperfect control [3], the constraints (6.7) (used in both MPC formulations) may also be formulated in an alternative manner as

$$g_{stage}(.) = \begin{bmatrix} y(i) - (y_{max} - \delta_{y_{max}}) \\ (y_{min} + \delta_{y_{min}}) - y(i) \\ x(i) - (x_{max} - \delta_{x_{max}}) \\ (x_{min} + \delta_{x_{min}}) - x(i) \\ u(i) - u_{max} \\ u_{min} - u(i) \end{bmatrix} \tag{6.10}$$

where the so called *back-off* terms $\delta_{...}$ are introduced to reduce the constraint violations to a specified minimum amount. For a fully optimal implementation, then, these terms should be determined in a systematic and optimal way, taking into account information on the disturbances and (disturbance rejection) control performance. See *e.g.* [105]

for such an approach. Alternatively, one may simply set these terms manually which would lead to loss of control performance but only if these terms would have to be reset frequently. At MSWC plants, this may not be the case and one could choose the latter, implementationally more simple option.

With respect to the choice for the inequality constraints $g_{stage}(.)$ it is also noted that infeasibility may occur when such - relatively tight - constraints have to be fulfilled for (too) early time instants (*i.e.* $i = 0$, 1, ...). Such infeasibility problems may be removed simply by removing the(se) earliest time instants from the period over which the constraints have to be fulfilled, which may be done with little loss in control performance as observed from simulations.

The terminal cost function $l_{terminal}(.)$ and constraint functions $g_{terminal}(.)$ and $h_{terminal}(.)$ are assumed to be chosen such that stability is guaranteed. This is discussed in more detail in section 6.3.6.

MPC applications with both considered optimal control problem formulations are discussed in chapter 7.

6.3.3 Solution to the optimal control problem

Introduction

In this section a numerical solution method to the optimal control problem(s) of the proposed MSWC plant NMPC strategy (*i.e.* applicable to any of the optimal control problem formulations discussed in section 6.3.2) is presented. For that purpose, first, an overview of available solution methods for optimal control problems is given. From these methods, then, one is chosen to be used by the proposed MSWC plant NMPC strategy and subsequently further elaborated.

Overview of available solution methods

For the solution to optimal control problems, the following methods are available [6, 7, 12, 40, 94]

- *Methods based on the Principle of Optimality*
  These methods employ Bellman's Principle of Optimality ("*An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision*" [50]).

- *Indirect* or *variational methods*
  These methods solve the optimal control problem indirectly by computing the solution to the corresponding (first-order) necessary conditions.

- *Direct methods*
  These methods solve the optimal control problem directly by transforming it into a nonlinear programming problem (NLP)

$$\begin{aligned} &\min_{\underline{x}} \ \underline{f}(\underline{x}) \\ &s.t. \quad \underline{g}(\underline{x}) \leq 0 \end{aligned}$$

$$\underline{h}(\underline{x}) = 0 \tag{6.11}$$

and solve the resulting problem. Here, $\underline{x}$[1] represent the variables to be optimized *c.q.* represent $U(N_p)$.

Methods based on the Principle of Optimality solve a partial differential equation (PDE) known as the *Hamilton-Jacobi-Carathéodory-Bellman PDE* or a closely related recurrence relation that is solved using a method known as (*iterative) dynamic programming*.

The necessary conditions to be solved by the *indirect methods* follow from *Pontryagin's Maximum principle*[2] (or *minimum* principle, depending on the reference considered) and leads to a so called *two-point boundary value problem* (TPBVP). This name refers to the fact that the boundary conditions for the corresponding differential equations partly are given as initial values (first point) and partly as final values (second point). A variety of methods are available to solve the resulting TPBVP, such as *gradient* or *control vector iteration methods*, (*single* and, the more stable, *multiple*) *shooting* or *boundary condition iteration methods*, *invariant embedding* and *collocation methods*.

The *direct methods* employ a finite-dimensional parametrization/discretization for the manipulated and/or state variables (*e.g.* piecewise constants) to transform the optimal control problem into an NLP. The resulting NLP is then typically solved by means of a *sequential* or *successive quadratic programming* (SQP) method (see later on in this thesis), which solves the NLP by repeatedly solving a quadratic programmming (QP) problem. The indirect methods can be divided into three groups:

- *Sequential* or *control variable parametrization methods*
- *Simultaneous methods*
- the *(direct) multiple shooting method*

Here, ”sequential” and ”simultaneous” refer to the order with which the optimization is performed with respect to the solution of the model and constraint equations. Sequential methods only parametrize the MVs, hence the name control variable parametrization method, and employ a strict separation between the (*outer*) optimization loop and the (*inner*) model and constraint solution/integration loop. With the inner loop, the objective and constraint values are computed to be used by the outer NLP solver loop for obtaining the solution to the optimal MVs. In contrast, simultaneous methods employ a full parameterization of both the manipulated and state variables (using *collocation*) to transform the optimal control problem into one (large) NLP by means of which, when being solved, the solutions to both the optimal MVs and the model and constraint equations are simultaneously obtained. As a sort of hybrid approach, the (direct) multiple

[1]The 'underbar' notation is used here to distinguish the NLP variables and function names from their counterparts in the optimal control problem.

[2]The Maximum Principle can be viewed as an extension of the so called *Calculus of Variations* to inequality constrained optimal control problems, which determines the necessary conditions by considering infinitesimally small variations around the optimal trajectory. This also explains the name *variational methods*.

shooting also employs a full parameterization of the MVs but employs a mixed sequential/simultaneous approach to the discretization of the state variables. More specific, it splits the state trajectory into subintervals of which the initial state values are added as optimization variables to the NLP, thereby mimicking the simultaneous approach, and during which the model and state equations are solved exactly (*i.e.* without parameterization of the states), thereby mimicking the sequential approach. (It must be noted that, although the (direct) multiple shooting method also inherits characteristics from the sequential approach, some prefer to see this method as a full simultaneous method: see *e.g.* [12]).

Choice of method

The characteristics and relative advantages and disadvantages of the considered optimal control problem solution methods have been discussed intensively in the literature, with the main focus being on their applicability to large scale problems where reduction of the computation time is the main issue. For small scale problems, computation time is typically much less of an issue. For such problems, only the methods based on the Principles of Optimality may not be feasible, *i.e.* these methods are feasible only for problems with model order $n \leq 3$ [7]. Criteria for comparison of the optimal control solution methods other than computation time are *e.g.* the handling of unstable models and the ability to impose state constraints on a more fine grid than used for parametrization of the MVs.

The optimal control problem to be solved by the MSWC plant NMPC strategy proposed here is of the small scale type with $n \geq 4$. Following the discussion above, both direct and indirect methods should be feasible from the point of view of computation time. Of these remaining methods, it is proposed here to use the sequential method. The first reason for that is that, through simulations, this method indeed has been shown to allow for a sufficiently fast computation of the MVs, *i.e.* within the imposed sampling period (of, typically, 60 $[s]$). Secondly, this method is relatively easy to implement, *e.g.* avoiding the formulation of the necessary conditions needed for the indirect method and the sophistication needed for exploiting the structure in the (direct) multiple shooting and simultaneous method. An additional advantage of the sequential method is that it is a *feasible path* method, meaning that at each iteration of the optimal control problem solver a feasible solution to the model is obtained. Disadvantages are *e.g.* the inability to handle unstable models and the inability to handle state constraints on a more refined grid than the grid used for parameterization of the MVs. These disadvantages are, however, considered of no or minor importance for MSWC plant applications.

The sequential method proposed for the MSWC plant NMPC strategy considered here is now discussed in more detail.

Global layout of the sequential method

As mentioned above, the global layout of a sequential method consists of an outer optimization loop and an inner model and constraint solution loop. With the inner loop, the objective and constraint values are computed to be used by the outer NLP solver loop for obtaining the solution to the optimal MVs. These loops form a closed loop that is iterated until some termination criterion within the NLP solver is fulfilled. For the MSWC plant NMPC strategy proposed here this global layout translates to the one schematically depicted in figure 6.1, where for ease of explanation the optimal control problem (6.2) and NLP (6.11) are considered without equality constraints ($h_{stage}(.)$, $h_{terminal}(.)/\underline{h}(.)$). Note that the sequential method here also requires gradient infor-

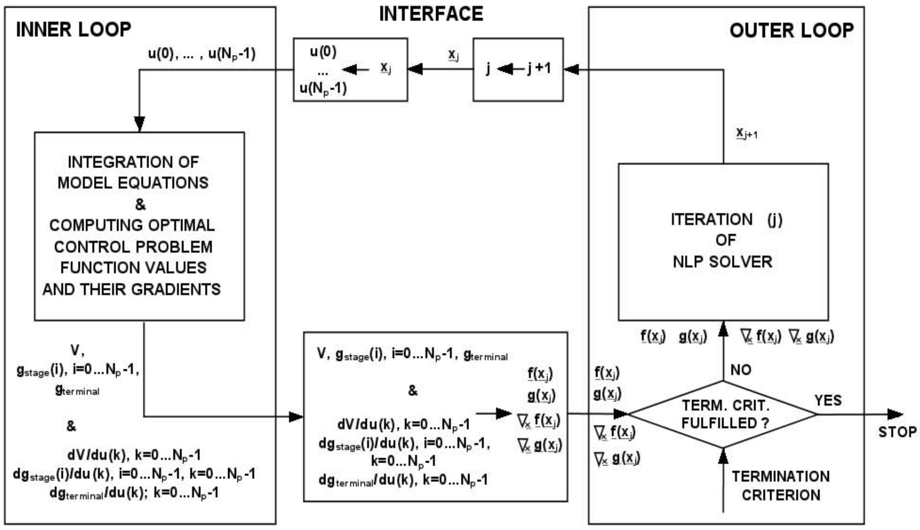

Figure 6.1: The sequential method for the proposed MSWC plant NMPC strategy.

mation $\nabla_{\underline{x}} f(\underline{x}_j)$, $\frac{\partial V}{\partial u(k)}(.)$, etc.

To complete the presentation of the MSWC plant NMPC strategy sequential method proposed here, two of its most important elements are elaborated now in detail. These elements are (i) the computation of the derivatives of the objective and constraint functions of the optimal control problem and (ii) the solution to the NLP.

Derivative computation

The computation of derivatives in optimal control problems is an important issue as it determines the accuracy of the solution and the computational speed of the chosen solution method. The following four ways of computing optimal control problem derivatives can be distinguished:

- via *finite differences*
- using *sensitivity equations*
- by solving *adjoint equations*
- via *automatic differentiation*

With *finite differences* (FD) the desired derivative is estimated from function values only. These function values are obtained at the point at which the derivative needs to be calculated and at one or more slightly perturbed versions of this point. The computationally most cheap FD methods use only two function evaluations. One such cheap method is the *forward difference* method which, for example, computes the derivative of the objective function $V(.)$ of the optimal control problem (6.2) towards some piecewise constant scalar MV $u(j) \in \{u(0),\ \ldots, u(N_p-1)\}$ as

$$\frac{\partial V}{\partial u(j)}(u(0),\ \ldots, u(N_p-1)) \approx \tag{6.12}$$
$$\frac{V(u(0),\ \ldots, u(j)+h,\ \ldots, u(N_p-1)) - V(u(0),\ \ldots, u(j),\ \ldots, u(N_p-1))}{h}$$

with h the applied perturbation.

Another approach is to compute the derivatives via integration of a set of equations (of differential and/or algebraic type) that represent the sensitivity of the states of the model used in the optimal control problem towards the parameter for which the derivative needs to be computed (*c.q.* $u(j)$). These equations are known as the *sensitivity equations* (SE; see *e.g.* [52, 88]) and are derived by taking the derivative of the model equations towards the parameter of interest. Also, for the integration of these equations, the model equations need to be integrated before or simultaneously. The SE approach will be demonstrated for the computation of, again, $\frac{\partial V}{\partial u(j)}(.)$ for scalar $u(j)$ (The extensions to vector-valued $u(j)$ and other optimal control problem functions are straightforward). For that purpose, note that the definition of $V(.)$ as given by eqns. (6.2) implies that the derivative of this objective function towards $u(j)$ is equal to

$$\begin{aligned} \frac{\partial V}{\partial u(j)}(u(0),\ \ldots, u(N_p-1)) \quad = \quad & \sum_{i=0}^{N_p-1} \{ (\frac{\partial l_{stage}}{\partial x(i)}(x(i), u(i), \ldots))(\frac{\partial x(i)}{\partial u(j)}) + \\ & (\frac{\partial l_{stage}}{\partial u(i)}(x(i), u(i), \ldots))(\frac{\partial u(i)}{\partial u(j)}) \} + \\ & (\frac{\partial l_{terminal}}{\partial x(N_p)}(x(N_p)))(\frac{\partial x(N_p)}{\partial u(j)}) \end{aligned} \tag{6.13}$$

which follows from applying the chain rule of differentation. The derivatives $\frac{\partial l_{stage}}{\partial x(i)}(x(i), u(i), \ldots)$, $\frac{\partial l_{stage}}{\partial u(i)}(x(i), u(i), \ldots)$ and $\frac{\partial l_{terminal}}{\partial x(N_p)}(x(N_p))$ here are generally easily computed. Also, $\frac{\partial u(i)}{\partial u(j)}$ is straightforwardly computed (as either 0 or 1 here). It, therefore, remains for the state derivatives $\frac{\partial x(i)}{\partial u(j)}$, $i = 0\ \ldots\ N_p$, to be computed in order to calculate $\frac{\partial V}{\partial u(j)}(.)$. These derivatives need to be computed via the SE, which follow from taking the derivative of the model equations $f(.)$ used in the optimal control problem (6.2) towards $u(j)$:

$$\frac{\partial x(i+1)}{\partial u(j)} = (\frac{\partial f}{\partial x(i)}(x(i), u(i), \ldots))(\frac{\partial x(i)}{\partial u(j)}) + (\frac{\partial f}{\partial u(i)}(x(i), u(i), \ldots))(\frac{\partial u(i)}{\partial u(j)}) \tag{6.14}$$

Assuming the derivatives $\frac{\partial f}{\partial x(i)}(x(i), u(i), ...)$ and $\frac{\partial f}{\partial u(i)}(x(i), u(i), ...)$ here to be easily computable, which generally is the case, and realising again that $\frac{\partial u(i)}{\partial u(j)}$ is easily computable too, these equations allow for the computation of the desired dervatives $\frac{\partial x(i)}{\partial u(j)}$, $i = 0 \; ... \; N_p$ by integrating them over the interval $i = 0 \; ... \; N_p - 1$ with initial condition

$$\frac{\partial x(0)}{\partial u(j)} = 0 \tag{6.15}$$

and simultaneously with or *a posteriori* the model equations to obtain the required state values $x(i)$.

The *adjoint equations* (AE) approach (see *e.g.* [88]) computes, as the name suggests, the derivatives as a function of a set of so called adjoint variables, which can be interpreted as the dynamic versions of the Lagrange multipliers for static optimization problems. These variables are computed by solving a set of differential equations backward in time that follow from deriving the first-order necessary conditions for the optimal control problem under consideration.

The aim of *automatic differentiation* (AD) is to use software to analyze the formulas that compute the value of the function for which the derivative is to be calculated and to produce from this analysis formulas to compute the value of the derivative of this function [71]. It is assumed that the function evaluation has been decomposed into a sequence of simple calculations, each one of which containing only one or two variables. The AD software package then analyzes these calculations to produce a similar set of simple calculations, but now containing a low number of partial derivatives, that, combined, efficiently evaluate the derivative of the considered function.

The *pros* and *cons* of the various derivative computation methods have been discussed intensively in the literature, in particular with respect to the ease of implementation, accuracy of the solution, ill-conditioning, stability of the solution and computational speed. Particularly the FD, SE and AE approach have been compared in the literature, see *e.g.* [88] Here, a comparison is given that heavily relies on the results presented in the latter reference.

With respect to the ease of implementation, the AD method is probably the preferred choice because of the software package doing all the work. However, such packages may not be available. Of the remaining methods, the FD method is the preferred choice from an implementation point of view, avoiding *e.g.* the computation of model Jacobians.

With respect to accuracy, the FD method typically performs worst due to the heavy dependency of the result on the choice for the perturbation size h. The other methods exhibit a high accuracy (limited by the integration tolerance).

The AE approach may suffer from ill-conditioning and unstable solutions for the adjoints, even if the model itself is stable. The other approaches do not seem to suffer from these disadvantages.

When comparing the FD, SE and AE approaches from a computation time point of view, the last one of these methods seems fastest for most problems due to the fact that the corresponding number of equations to be integrated is independent of the number of parameters towards which the sensitivities need to be computed. See the comparison results in [88]. However, for a proper comparison of these methods, one should also take into account the combined ability of (i) solving the underlying equations simultaneously and of (ii) exploiting this ability by exploiting *vectorization* / *parallel computing* capability. When taking into account this factor, it is important to note that a large reduction in computational speed can be obtained for the FD and SE approaches as these allow for simultaneous integration of the equations. The AE approach, on the other hand, is bound to a serial integration of equations, as a result of which, when exploiting parallel computation capability, the first two methods may exhibit for many cases a computational speed comparable to that of the AE approach. Also, the number of constraints is a factor to be taken into account when comparing the three methods here. More specific, the number of integrations to be performed by the AE approach depends on this number, whereas this is not the case for the other two methods. Hence, for problems with many constraints, which are not uncommon in NMPC applications, the FD and SE approaches approach may be faster. Finally, also the memory storage capacity and the speed with which this memory can be accessed may be factors to be reckoned with as the speed of the AD approach depends on these factors, in particular when the number of intermediate calculations needed for evaluating a function and its derivative is large.

For the small scale (MSWC plant) NMPC applications considered here, computational speed is not the dominant criterion on the basis of which to choose the derivative computation method as then typically all available methods are sufficiently fast. Instead, other factors such as ease of implementation, availability of the necessary software, accuracy, possible ill-conditioning and instability are dominating the choice of method. From an accuracy point of view, the FD method is less favorable. From the point of view of possible ill-conditioning and stability, the AE method is less favorable. The AD method is less favorable from the point of view of availability of the necessary software. Therefore, the SE approach has been chosen to compute the derivatives for the sequential method used for solving the optimal control problem of the MSWC plant NMPC strategy proposed here. More specific, the chosen SE approach is as described in this section with the model equations simultaneously solved with the sensitivity equations (and not *a priori*).

Solving the NLP

In the literature on solution methods for NMPC optimal control problems, the *sequential* or *successive quadratic programming (SQP)* method is the dominant NLP solver. In fact, it is a particularly popular method for solving any type of NLP and, additionally, one of the most successful ones [71]. The SQP method derives its name from the fact that it solves the NLP *c.q.* the optimization problem (6.11) by sequentially / successively (iteratively) solving a quadratic program (QP)

$$\begin{aligned} \min_x \ & \tfrac{1}{2}x^T H x \ + \ f^T x \\ s.t. \quad & A_{ineq} x \leq b_{ineq} \\ & A_{eq} x = b_{eq} \end{aligned} \tag{6.16}$$

A QP is itself an NLP due to the presence of a quadratic objective function, but for this type of optimization problem efficient and reliable solvers are widely available. In fact, one of the motivations for applying the SQP method is the availability of such solvers. The QP follows from applying Newton's method (see section D.2.3), in a QP manner (rather than solving it via a set of linear equalities), to the set of equations consisting of the first-order necessary conditions and constraints of the NLP to be solved. The QP manner of applying Newton's method is motivated by the presence of inequality constraints. In that case, Newton's method cannot be applied in the conventional way, *i.e.* by solving linear sets of equalities, but can be applied in a generalized way by means of solving a QP.

Because of its popularity the SQP method was chosen for the MSWC plant NMPC strategy proposed here, *i.e.* to solve the sequential method NLP used for solving the corresponding optimal control problem. Computational results obtained with realistic simulation based MSWC plant NMPC test problems showed this choice to be well valid. The specific SQP method employed here is an extension towards inequality constraints of an SQP method given in [8] that deals only with NLPs with equality constraints. For completeness, it is discussed in full detail in appendix D.

6.3.4 Solution to the state estimation problem

Introduction

The MSWC plant NMPC control strategy proposed here requires a state estimator to deliver sufficiently accurate predictions of the states and disturbances at the next sampling instant from measurements available until and at the current sampling instant. In this section, the state estimator chosen as part of this strategy is outlined and its choice is motivated. For that purpose, first, the corresponding state estimation problem is specified. After that, an overview is given of methods that can be used to solve this problem, where the set of considered methods is confined to those typically employed currently for industrial processes. From these methods then a choice is made.

Formulation of the MSWC plant state estimation problem

Starting point here is a model of the type of eqns. (6.1). To allow for estimation of the disturbances, it is assumed that this model is extended with dynamic models for the measured and nonmeasured disturbances, *e.g.*

$$\begin{aligned} x^{d_m}(k+1) &= f^{d_m}(x^{d_m}(k), v^{d_m}(k)) \\ d_m(k) &= h^{d_m}(x^{d_m}(k), w^{d_m}(k)) \end{aligned} \tag{6.17}$$

for the measured disturbances and

$$\begin{aligned} x^{d_{nm}}(k+1) &= f^{d_{nm}}(x^{d_{nm}}(k), v^{d_{nm}}(k)) \\ d_{nm}(k) &= h^{d_{nm}}(x^{d_{nm}}(k)) \end{aligned} \tag{6.18}$$

for the nonmeasured disturbances. Here, $v^{d_m}(k)$, $w^{d_m}(k)$ and $v^{d_{nm}}(k)$ are zero-mean white noise (ZMWN) signals representing *system noise* in case of $v^{d_m}(i)$ and $v^{d_{nm}}(i)$

and *measurement noise* in case of $w^{d_m}(i)$ (Note the absence of measurement noise in the model for $d_{nm}(i)$ which is motivated by the fact that these disturbances are not measured). The measured disturbance models can be obtained through time series analysis performed on historic data. The nonmeasured disturbance model (unless somehow *a priori* available) is assumed here to be of the integrating type to introduce offset-free control, as is discussed in section 6.3.5. Combining the disturbance models with the original model (6.1) results in the augmented model

$$\begin{aligned}
x^{d_m}(k+1) &= f^{d_m}(x^{d_m}(k), v^{d_m}(k)) \\
x^{d_{nm}}(k+1) &= f^{d_{nm}}(x^{d_{nm}}(k), v^{d_{nm}}(k)) \\
x(k+1) &= f(x(k), u(k), h^{d_m}(x^{d_m}(k), w^{d_m}(k)), h^{d_{nm}}(x^{d_{nm}}(k)), \theta) \\
d_m(k) &= h^{d_m}(x^{d_m}(k), w^{d_m}(k)) \\
y(k) &= h(x(k), u(k), h^{d_m}(x^{d_m}(k), w^{d_m}(k)), h^{d_{nm}}(x^{d_{nm}}(k)), \theta)
\end{aligned} \tag{6.19}$$

The state estimation problem considered here is then to deliver optimal predictions $\hat{x}(k+1)$, $\hat{x}^{d_m}(k+1)$ and $\hat{x}^{d_{nm}}(k+1)$ from measurements of $u(k)$, $d_m(k)$ and $y(k)$ available up till and at current time k, using model (6.19) and *a priori* information (*c.q.* covariance matrices) on the disturbances $v^{d_m}(k)$, $w^{d_m}(k)$ and $v^{d_{nm}}(k)$ (this may not be available in case of which this information is obtained indirectly through tuning of the state estimator). Then, when the augmented model (6.19) is also used in the NMPC optimal control problem, the resulting values for $\hat{x}(k+1)$, $\hat{x}^{d_m}(k+1)$ and $\hat{x}^{d_{nm}}(k+1)$ are used as the initial values to this problem to compute the optimal MV values $u(k+1)$. Note that for the solution to this problem, also predictions $\hat{v}^{d_m}(k+1)$, $\hat{w}^{d_m}(k+1)$ and $\hat{v}^{d_{nm}}(k+1)$ are required (further future values, *i.e.* for $k+2$ and further, are assumed to be equal to those for $k+1$ within the proposed MSWC plant NMPC strategy; see section 6.2). Given the assumed ZMWN nature of the corresponding disturbance sources, these futher future predictions are simply chosen here equal to 0. When the original non-augmented model (6.1) is used in the NMPC optimal control problem, the values for $\hat{x}(k+1)$, $\hat{x}^{d_m}(k+1)$ and $\hat{x}^{d_{nm}}(k+1)$ are used to deliver the initial *c.q.* one-sample-instant-ahead-predicted values for $x(k)$ and (initial and future values) for $d_m(k)$ and $d_{nm}(k)$. Here, also predictions $\hat{v}^{d_m}(k+1)$, $\hat{w}^{d_m}(k+1)$ and $\hat{v}^{d_{nm}}(k+1)$ are required which, again, are set to 0.

Available methods and choice of method for MSWC plant NMPC applications

A large variety of state estimation techniques is found in the literature, see *e.g.* [10] for an overview, which can be divided in groups in several ways. A first such division is in *stochastic* versus *deterministic* approaches with the difference being that the first group takes the characteristics of the disturbance terms in the model (such as *e.g.* $v^{d_m}(k)$ in (6.19)) in a systematic manner into account whereas the latter group ignores these characteristics. Another division is that between approaches that solve the estimation problem by means of numerical optimization and approaches that do not. To limit the scope in the search for the best state estimation technique for the proposed MSWC plant NMPC strategy, only techniques have been considered that are currently most popular for process control applications. These state estimation techniques are

- the following three nonlinear extensions of the linear *Kalman filter* (KF) [39]:
  - the *Extended Kalman filter* (EKF) [2, 37]
  - the *Ensemble Kalman filter* (EnKF) [24]
  - the *Unscented Kalman filter* (UKF) [38, 113])

  which are stochastic approaches.
- the numerical optimization based *moving-horizon estimator* (MHE; see *e.g.* [84, 87, 91] and appendix E)

For ease of explanation of KF types of state estimators a simplified yet comprehensive model

$$\begin{aligned} x(k+1) &= f(x(k), v(k)) \\ y(k) &= h(x(k), w(k)) \end{aligned} \tag{6.20}$$

is assumed with $v(k)$ and $w(k)$ ZMWN signals with covariance matrices $Q(k)$ resp. $R(k)$. (Note the absence of inputs/MVs $u(k)$ here, stressing that typically the main contribution of state estimators is the incorporation of output measurements rather than input measurements, which may be regarded as known - time varying - model parameters and thereby can be left out of the discussion).

The aim of KF types of state estimators, then, is to provide a stochastically optimal estimate for $x(l)$, with optimal defined in *minimum mean square error* sense (see *e.g.* [2, 38, 113]), or at least one that approximates this optimal estimate, using (i) measurements $y(k)$, $y(k-1)$, ..., (ii) knowledge of the plant dynamics *c.q.* $f(.)$ and $h(.)$ and (iii) knowledge of the covariance matrices $Q(k)$ and $R(k)$. Here, l may be larger than k, in which case one speaks of (state) *prediction*, or equal to k, in which case one speaks of *filtering*, or smaller than k, in which case one speaks of *smoothing*. Here, the prediction case with $l = k+1$ is of interest. The estimate for $x(l)$ is computed in a recursive manner where at each recursion an already available state estimate is updated to a new one using newly available measurements and the available knowledge about the model and about the stochastic properties of the noises. Before discussing this recursion in more detail, first the notation $\hat{x}(l|k)$ is introduced, meaning "optimal estimate of the state at time l given (information c.q.) measurements up till time k". Then, each recursion is triggered by the acquisition of new measurements which, in the setting considered here, occurs every sampling instant (k). At such an instant an estimate $\hat{x}(k|k-1)$ of the state at time k is already available from an earlier recursion or as an *a priori* best guess. The aim of a KF type of estimator is then to provide $\hat{x}(k+1|k)$ at the end of the recursion, which at the next sampling instant, at a new recursion, is to be updated to a new optimal estimate $\hat{x}(k+2|k+1)$, etc. The update from $\hat{x}(k|k-1)$ to $\hat{x}(k+1|k)$ at the recursion at sampling instant k proceeds in two steps:

1. First, a *measurement update* (MU) step (also called *analysis step*) is performed where the already available state estimate $\hat{x}(k|k-1)$ is upgraded to an improved estimate $\hat{x}(k|k)$ using new information extracted from, as the name of this step suggests, newly available measurements $y(k)$.

2. Secondly, a *time update* (TU) step (or *forecast step*) is performed where the state estimate resulting from the MU step $\hat{x}(k|k)$ is upgraded to an improved estimate $\hat{x}(k+1|k)$. This step essentially is an open-loop prediction. Its name refers to the fact that it computes the evolution of the state estimate over the time interval (T_s) connecting sampling instants k and $k+1$.

The MU and TU steps, in fact, represent the recursive propagation of the conditional mean and covariance of $x(k)$ over time, with the conditional mean representing the optimal estimate of the state vector. KF types of estimators deliver optimal estimates if (i) the states $x(k)$ and measured outputs $y(k)$ are Gaussian distributed, (ii) the MU step is, with $E[.]$ representing the expectation operator, equal to (see *e.g.* [38, 113])

$$\begin{aligned}
\hat{x}(k|k) &= \hat{x}(k|k-1) + L(k)(y(k) - \hat{y}(k|k-1)) \\
\hat{y}(k|k-1) &= E[h(\hat{x}(k|k-1), w(k))] \\
L(k) &= P^{\tilde{x}\tilde{y}}(k|k-1)(P^{\tilde{y}\tilde{y}}(k|k-1))^{-1} \\
P^{\tilde{x}\tilde{y}}(k|k-1) &= E[(x(k) - \hat{x}(k|k-1))(y(k) - \hat{y}(k|k-1))^T] \\
P^{\tilde{y}\tilde{y}}(k|k-1) &= E[(y(k) - \hat{y}(k|k-1))(y(k) - \hat{y}(k|k-1))^T]
\end{aligned} \tag{6.21}$$

and (iii) the subsequent TU step is equal to

$$\begin{aligned}
\hat{x}(k+1|k) &= E[f(\hat{x}(k|k), v(k))] \\
P(k+1|k) &= E[(x(k+1) - \hat{x}(k+1|k))(x(k+1) - \hat{x}(k+1|k))^T]
\end{aligned} \tag{6.22}$$

Each KF type of estimator employs the linear update rule of the MU step, see the first of the equations (6.21) (which actually represents the best linear unbiased estimate (BLUE) [10]), but differs with respect to how the remaining quantities are computed. In particular, in case of a linear model, leading to the original Kalman filter [39], the *Kalman gain* $L(k)$, means and covariances can be directly expressed in terms of the system matrices that make up this model (see *e.g.* [2]). The EKF approximates the optimal KF equations by

$$E[h(\hat{x}(k|k-1), w(k))] \approx h(E[\hat{x}(k|k-1)], E[w(k)]) = h(\hat{x}(k|k-1), 0) \tag{6.23}$$

in the MU equations and by

$$E[f(\hat{x}(k|k), v(k))] \approx f(E[\hat{x}(k|k)], E[v(k)]) = f(\hat{x}(k|k), 0) \tag{6.24}$$

in the TU equations and, additionally, by using the model Jacobians

$$\frac{\partial f(x(k), v(k))}{\partial x(k)}|_{x(k)=\hat{x}(k|k), v(k)=0} \qquad \frac{\partial h(x(k), w(k))}{\partial x(k)}|_{x(k)=\hat{x}(k|k-1), w(k)=0} \tag{6.25}$$

for approximating the covariances, which in fact resemble the expressions of the (linear) KF (the covariance expressions are left out here for reasons of space). The EnKF and UKF approximate the optimal KF equations by replacing the expectations in (6.21) and (6.22) for *sample* means and sample covariances:

$$\begin{aligned}
E[h(\hat{x}(k|k-1), w(k))] &\approx \frac{1}{N}\Sigma_{t=1}^{N} h(\hat{x}^t(k|k-1), w^t(k)) \\
E[f(\hat{x}(k|k), v(k))] &\approx \frac{1}{N}\Sigma_{t=1}^{N} f(\hat{x}^t(k|k), v^t(k))
\end{aligned} \tag{6.26}$$

etc., where $\hat{x}^t(k|k-1)$, $w^t(k)$, $\hat{x}^t(k|k)$, $v^t(k)$, $t = 1...N$, are *a priori* chosen realizations of the corresponding random variables. The difference between the EnKF and the UKF lies in how the latter realizations are selected: the EnKF does that in a random way (via some random generator) to create an *ensemble* of such realizations, whereas the UKF does that in a specific deterministic way using the so called *unscented transform(ation)* [38].

When comparing the EKF, EnKF and UKF, the latter two generally deliver more accurate estimates, in particular for highly nonlinear plants. An illustrative and easy-to-reproduce example of that, more specific of the worse performance of the EKF compared to the UKF when applied to a significantly nonlinear system, can be found in [38]. Due to the similar way of computing the means and covariances by the latter two estimators, these are expected to deliver estimates of the same accuracy. Of these two, the UKF is faster than the EnKF for low order systems (more specific for systems (6.20) with approx. $nx + nv < 50$ with nx the dimension of $x(k)$, nv the dimension of $v(k)$ and a typical ensemble size of 100 assumed for the EnKF) but requires the tuning of three more parameters.

The MHE is the state estimation equivalent of MPC, *i.e.* it solves the state estimation problem in a receding horizon manner via a constrained optimal control problem formulation reflecting the aim of fitting past state and disturbance values to measured data. Here, the nonlinear type of MHE (NMHE) is considered, *i.e.* employing a nonlinear optimal control problem, in particular a nonlinear model.

Through simulations the (N)MHE has been tested on MSWC plant NMPC applications. From these tests, see appendix E and [57], it was observed that it allows for a good estimation performance while also fulfilling the (overall NMPC) computational requirements. In particular, it has been shown that it provides a significantly better estimation performance than an EKF, and thereby a significantly better closed-loop NMPC performance. However, it was also concluded that (i) constraints do not play a significant role in the MSWC plant estimation problem and that (ii) the improvement over the EKF performance is due to a better handling of plant nonlinearities. In addition, it was concluded from the comparison results in [38] that a similar improvement over the EKF is most likely to be obtained also with an UKF and EnKF. Because of this and because of (i) the much easier implementation of the latter two types of state estimators and (ii) the much lower computation time of these state estimators compared to the (N)MHE, it is proposed here to use one of the latter two KF types of state estimators. Note that the lower state estimation computation time is beneficial for the solution to the MV optimal control problem as it is allowed to (iteratively) approximate its optimum more closely.

Of the remaining two state estimators, *i.e.* the UKF and EnKF, the first is faster for the considered MSWC plant NMPC applications and is therefore in principle preferable as it also allows the MV optimal control problem solver to more closely approximate its optimum. However, the difference in computation time with the EnKF was found to be small and it requires more tuning. Therefore, the EnKF may as well be used for the proposed MSWC plant NMPC strategy.

6.3.5 Enforcing offset-free control

Introduction

Due to model errors and/or the presence of unmodeled (on average) constant disturbances, offsets may arise between setpoints and their CV counterparts, *e.g.* in case of a setpoint tracking type of control problem used in an overcapacity market situation (see section 6.3.2), or between the desired back-off to some dominant constraint to be fulfilled and its CV counterpart, *e.g.* in case of a constrained maximization/constraint pushing type of control problem used in an undercapacity market situation. The presence of these offsets may lead to a significantly degraded control performance and, thereby, overall economic performance and these should therefore be removed. In this section, methods are discussed and proposed by means of which offset-free control can be incorporated in an MPC strategy and, particularly, in the MSWC plant NMPC combustion control strategy considered here. First, methods for removal of offsets towards setpoints are discussed, then methods for removal of offsets towards dominating constraints.

Removal of offsets towards setpoints

Methods
The first method considered here is based on the observation that an offset on a CV can simply be removed by adding an equally large offset to the corresponding setpoint. More specific, the idea is to replace the original setpoint r by a new one $r' = r + \Delta r$ with Δr chosen such that $r - y(\infty) = 0$ where $y(\infty)$ represents the value to which the corresponding CV (on average) eventually converges. The computation of the offset Δr can be implemented by means of an integrator operating on the observed difference between r and $y(k)$. The output value of the integrator is then added to the setpoint. Further filtering can be added to this integrator to fine-tune *c.q.* optimize the resulting closed-loop performance. The most simple implementation performs each sample time k the following computation:

$$r'(k+1) = r'(k) + (r - y(k)) \tag{6.27}$$

which, for a stable closed-loop system, results for $k \to \infty$ in $y(\infty) = r$ as then $r'(k+1) = r'(k) = r'(\infty)$.

The second method is one that has been proposed and elaborated in the literature, albeit mainly only for LMPC. To obtain offset-free setpoint tracking, it employs (non-measured) disturbance models to incorporate integral action in the state estimator and, thereby, in the resulting NMPC controller. In order to explain this second method, assume a discrete-time plant

$$\begin{aligned} x(k+1) &= f(x(k), u(k)) \\ y(k) &= h(x(k), u(k)) \end{aligned} \tag{6.28}$$

to be controlled by an NMPC controller that employs a model

$$\begin{aligned} \underline{x}(k+1) &= \underline{f}(\underline{x}(k), u(k)) \\ \underline{y}(k) &= \underline{h}(\underline{x}(k), u(k)) \end{aligned} \tag{6.29}$$

An offset in the CVs $r - y(\infty) \neq 0$ can now be rewritten *c.q.* decomposed as

$$r - y(\infty) = (r - \underline{y}(\infty)) + (\underline{y}(\infty) - y(\infty)) \tag{6.30}$$

This description of the offset discloses the following two actual main potential sources for offset in the CVs:

- The inability of the NMPC optimal control problem solver to deliver the values for $u(\infty)$ that result in $\underline{y}(\infty) = r$.
- The inability of the state estimator to deliver values for $\underline{x}(\infty)$ that render $\underline{y}(\infty) = y(\infty)$.

The first source of offset is due to the NMPC optimal control problem solver getting stuck in a local rather than the global optimum or due to too many constraints being active at steady-state, in which case the steady-state r may not be reachable at all. It is noted that the problem of getting stuck in a local optimum is a typical NMPC problem whereas it is not a problem for LMPC controllers where convexity of the optimization problem always results in the global optimum (This may be the cause for this source of offset (seemingly) not yet having been identified in the literature). This first source of offset may be removed by properly choosing and tuning of the optimal control problem solver and by preventing steady-states with too many active constraints.

The second source for offset is the result of model errors and/or unmodeled disturbances and the absence of a mechanism to deliver values for $\underline{x}(\infty)$ that render $\underline{y}(\infty) = y(\infty)$, even in the presence of these errors and disturbances. Such a mechanism is the addition of so called *integrating disturbances*, which are of the form

$$d(k+1) \quad = \quad d(k) \tag{6.31}$$

and which introduce integral action into the state estimator *c.q.* constantly "push" the state estimate to close the gap between $\underline{y}(k)$ and $y(k)$. Following the LMPC work on this subject in [77] and [99], this addition of integrating disturbances may be done according to

$$\begin{aligned} \overline{d}(k+1) &= \overline{d}(k) \\ \overline{x}(k+1) &= \underline{f}(\overline{x}(k), u(k)) + \overline{C}^f \overline{d}(k) \\ \overline{y}(k) &= \underline{h}(\overline{x}(k), u(k)) + \overline{C}^h \overline{d}(k) \end{aligned} \tag{6.32}$$

with the size of the disturbance vector $\overline{d}(k)$ and the matrices $\overline{C}^f$ and $\overline{C}^h$ to be chosen. In order to see now how and under what conditions, for a model of the form of (6.32), integrating disturbance models introduce integral action in the state estimator, it is noted that the integrating disturbance equations are used, every time k, to perform the following computations:

$$\overline{d}(k+1|k) \quad = \quad \overline{d}(k|k-1) + L^d(k)(y(k) - \overline{y}(k|k-1)) \tag{6.33}$$

with $L^d(k)$ being the row(s) of the Kalman gain $L(k)$ corresponding to the integrating disturbance equation(s) for $\overline{d}(k)$. Now, when the closed-loop system is stable the values for $\overline{d}(.)$, $y(k)$ and $\overline{y}(k|k-1)$ converge to steady-state values $\overline{d}(\infty)$, $y(\infty)$ resp. $\overline{y}(\infty)$. Similarly, the gain $L^d(k)$ converges to a steady-state gain $L^d(\infty)$ and (6.33) becomes equal to

$$\overline{d}(\infty) \quad = \quad \overline{d}(\infty) + L^d(\infty)(y(\infty) - \overline{y}(\infty)) \tag{6.34}$$

When now, additionally, $L^d(\infty)$ is such that the last term on the right side of (6.34) can only be zero when the difference $y(\infty) - \overline{y}(\infty)$ is zero, it follows from this equality that $\overline{y}(\infty) = y(\infty)$.

Notes on disturbance modeling and on conditions for offset-free setpoint tracking
The choice of integrating disturbance model, *c.q.* the size of $\overline{d}(k)$ and the matrices $\overline{C}^f$ and $\overline{C}^h$ in (6.32), is an important issue in NMPC control design as this choice can have a large influence on the closed-loop performance of this controller (Note that multiple disturbance models can lead to offset-free setpoint tracking and that, hence, the disturbance model represents a degree-of-freedom in optimizing the closed-loop MPC performance). Several examples of that fact can be found in the literature, both for LMPC controllers [70, 77, 76] and NMPC controllers [99]. An erroneous choice of integrating disturbance model can even lead to an unstable closed-loop system.

With respect to integrated disturbance modeling, it, first of all, should be noted that the size of $\overline{d}(k)$ should not exceed the number of output measurements used by the state estimator to avoid problems due to non-uniqueness of the steady-state solution for this disturbance vector. A typical choice for this size would be the number of CVs for which the offsets need to be removed.

A particular observation that has brought up the discussion on integrating disturbance modeling is that most industrially applied MPC controllers use a so called *output* integrating disturbance model, with $\overline{C}^h = I$ and $\overline{C}^f = 0$, while at the same time the usage of this model is known to lead to sluggish rejection of disturbances entering the plant elsewhere. Because at most industrial processes the disturbances enter at the input rather than at the output, the literature on integrating disturbance modeling presently seems to favour *input* integrating disturbance models ($\overline{C}^h = 0$ and $\overline{C}^f \neq 0$). There do exist systematic approaches for determining the best integrating disturbance model, though. One such approach is given in [76]. The solution proposed in this reference assumes a robust control setting, with uncertainty bounds defined for the model employed by the MPC controller, and aims to minimize, over the integrating disturbance model parameters, the worst closed-loop control performance that is possible over the set of models defined by the uncertainty region. The proposed methodology solves a min-max optimization problem to obtain the corresponding integrating disturbance model parameters.

A note here is that no formal proofs have been encountered in the literature yet that fully disclose the conditions that guarantee the presence of integral action in NMPC controllers when using the second - disturbance modeling based - method discussed above. In contrast, for LMPC controller such proofs *can* be found, *e.g.* in [70] and [77]. These proofs provide similar conditions to those stated above for NMPC controllers. These involve, amongst others, stability of the closed-loop system, conditions on the number of integrating disturbances and *e.g.* conditions on $L^d(\infty)$ that guarantee that, see eqn. (6.34), $L^d(\infty)(y(\infty) - \overline{y}(\infty)) = 0$ if and only if $\overline{y}(\infty) = y(\infty)$ (which typically translates to a condition on the null space of this gain *e.g.* that this space is allowed to only contain the zero vector).

Choice of method
With respect to the choice of method for (setpoint) offset-free control in MSWC plant NMPC applications, the second *c.q.* disturbance modeling based method is proposed here. This choice is based on closed-loop NMPC simulations (discussed in the next chapter) where it has been found that the second method allows for a better overall control performance, in particular a better disturbance rejection performance.

More specific, using the model (6.1) as a starting point, good results have been obtained with a state estimator model of the form

$$\begin{aligned} \overline{d}_{nm}(k+1) &= \overline{d}_{nm}(k) \\ \overline{x}(k+1) &= \underline{f}(\overline{x}(k), \overline{u}(k), \overline{d}_m(k), \overline{d}_{nm}(k), \theta) \\ \overline{y}(k) &= \underline{h}(\overline{x}(k), \overline{u}(k), \overline{d}_m(k), \overline{d}_{nm}(k), \theta) \end{aligned} \tag{6.35}$$

and an optimal control problem model of the form

$$\begin{aligned} \overline{x}(k+1) &= \underline{f}(\overline{x}(k), \overline{u}(k), \overline{d}_m(k), \overline{d}_{nm}(k), \theta) \\ \overline{y}(k) &= \underline{h}(\overline{x}(k), \overline{u}(k), \overline{d}_m(k), \overline{d}_{nm}(k), \theta) + [\overline{b}(k)^T \; 0 \; ... \; 0]^T \end{aligned} \tag{6.36}$$

(of which a subset of the outputs $\overline{y}(.)$ is selected as CVs and) with

- the optimal control problem initial values for the state and disturbance vectors $\overline{x}(k)$ and $\overline{d}_{nm}(k)$ obtained as estimates from the state estimator
- the optimal control problem initial values for $\overline{d}_m(k)$ obtained as $d_m(k-1)$
- the optimal control problem initial values for $\overline{b}_k$ obtained as the first nb elements from

  $$\overline{b}(k-1) = \overline{y}(k-1) - \underline{h}(\overline{x}(k-1), \overline{u}(k-1), \overline{d}_m(k-1), \overline{d}_{nm}(k-1), \theta) \tag{6.37}$$

  with nb the column size of $\overline{b}(.)$ and where $\overline{x}(k-1)$ and $\overline{d}_{nm}(k-1)$ are obtained as well from the state estimator but one sample instant earlier.

Even more specific, good results have been obtained with this state estimator and optimal control problem model with (see also chapter 2)

- the vector $\overline{d}_{nm}(.)$ containing X_{inert} and X_{mois} as elements
- the CVS included to calculate the values for z (which otherwise also must have been contained in $\overline{d}_{nm}(.)$)
- the output measurements used for estimating the states *c.q.* included in $\overline{y}(.)$ being ϕ_{st}, $Y_{O_2,fg}$, $Y_{CO_2,fg}$, $Y_{H_2O,fg}$, ϕ_{fg} and T_g where ϕ_{st}, $Y_{O_2,fg}$, ϕ_{fg} and T_g are measurements that are commonly logged at MSWC plants and where the remaining two veriables $Y_{CO_2,fg}$, $Y_{H_2O,fg}$ are measured due to the usage of the CVS (which requires these two variables as inputs).

- the vector $\overline{b}(.)$ containing ϕ_{st} and $Y_{O_2,fg}$ as elements
- the state vector containing $M_{CH_yO_z}$, M_{mois}, T_s and ϕ_{st} as elements
- the measured disturbances, contained in $\overline{d}_m(k)$, being T_{air}, RH_{air} (both needed for the CVS), T_{prim}, T_{sec}, $T_{w,in}$ and T_{leak}, where the latter four temperatures are assumed equal to T_{air}.

It is noted that no full disturbance model optimization step has been performed yet for the MSWC plant MPC applications considered here and that, hence, there still may be room for improvement in the MPC disturbance modeling choices made here. It is also noted that it was observed that usage of the output disturbance $\overline{b}(k)$ as discussed above can largely improve the disturbance rejection properties of the MPC control strategy.

Removal of offsets towards dominating constraints

In case of a constrained maximization/constraint pushing type of MSWC plant operation by means of MPC, model errors and/or the presence of unmodeled disturbances may cause the operating point to lie further from the dominating constraint than is required from a constraint violation *c.q.* disturbance rejection perspective, leading to a possibly significant degradation in economic performance. These errors and disturbances may also cause the operating point to lie too close to the dominating constraint, leading to too many environmental or maintenance related constraint violations (the latter eventually also leading to a degradation in economic performance). A method for removing the offset due to model error or unmodeled disturbances is through the mechanism of manually (re)setting the corresponding back-off ($\delta_{\ldots}$: see section 6.3.2) or, in case no such back-off terms are used in the MPC optimal control problem, the corresponding constraint value. The offset is then removed simultaneously with setting a safe distant from the dominating constraint to take into account other possible sources for constraint violation (in particular, as will be seen in chapter 7, fastly varying disturbances may cause constraint violation in case of no model error or unmodeled disturbances and require setting a back-off). The methods discussed in the previous section for the setpoint tracking case might also be used to remove the offsets towards dominating constraints. In particular, the disturbance modeling approach might be an attractive alternative to the simple, manual approach discussed here because of its more systematic and *c.q.* more automated and, thereby, potentially more optimal handling of offsets. Integral action must then be incorporated in the state estimator through choosing a suitable disturbance model to ensure convergence of the CV corresponding to the dominating constraint to its measured counterpart.

For the MSWC plant combustion control NMPC strategy outlined here it is proposed to use the manual setting method discussed above for removing offsets towards dominating constraints, in particular for a constraint pushing type of optimal control problem, because of its simplicity. However, for future applications it may be worthwhile to investigate the opportunities for improving this offset-free control enforcement method through a more systematic method such as *e.g.* the disturbance modeling approach discussed above.

6.3.6 Enforcing stability

A main requirement for the proposed MSWC plant NMPC control strategy is that it should result in a stable controlled MSWC plant. Even though an MSWC plant is stable by itself, arbitrary implementation of an MSWC plant NMPC control strategy may lead to instability and specific measures need to be implemented in such a strategy to prevent that. In this section, typically employed such measures are discussed and a specific one is proposed to be used as part of the proposed MSWC plant NMPC strategy.

Two approaches can be distinsguished by means of which one enforces stability on NMPC controlled plants:

- the usage of *long (pseudo-infinite) prediction horizons* in the NMPC optimal control problem
- the usage of so called *stability constraints* in this problem

The first approach is commonly applied in commercially available MPC packages, probably because of its easy implementation. It is a consequence of a well known stability result for such control strategies that states that, under certain conditions of which some are discussed below, stability is obtained with an *infinite* prediction horizon. See *e.g.* [72] for this result for the NMPC case. This infinite horizon approach translates to MPC practice in the choice for a (very) long prediction horizon, thereby mimicking the infinite horizon NMPC controller and hoping that for some sufficiently high chosen value for the prediction horizon its stability properties are inherited.

Disadvantage of the pseudo-infinite horizon approach above can be the long computation time. Another disadvantage is that stability is obtained via tuning a parameter that also affects the controller performance, *i.e.* the prediction horizon. Hence, this approach does not allow for a complete separation of establishing stability and tuning for performance. Although this does not significantly break down the easy applicability of the pseudo-infinite horizon approach, arguments like these two have motivated the research for other ways of establishing stability for NMPC controlled systems. In fact, much of the literature on NMPC is about this topic and contains many different stabilizing NMPC formulations. This research on stability for NMPC controllers has culminated recently in a unifying stability theory for most of these, seemingly different, stabilizing NMPC formulations. This theory, which can be found in [68] and is based on Lyapunov function theory (with the value of the objective function being the Lyapunov function), consists of four axioms which, when fulfilled, establish stability under (additional) conditions to be discussed later on. These axioms, in return, are fulfilled when adding so called *stability constraints* to the optimal control problem formulation. These constraints are generally of the following two forms (see *e.g.* [68, 72]):

- a specific penalty function on the state at the end of the prediction horizon which is added to the objective function of the optimal control problem, *e.g.* as $l_{terminal}(x(N_p))$ in (6.2). Such a penalty function is also denoted as *terminal penalty function* or *terminal cost function*.
- a specific (hard) constraint on (again) the state at the end of the prediction horizon, enforcing this state to be contained in some region X_f: $x_N \in X_f$. The latter

constraint is also denoted as a *terminal constraint* with the corresponding region X_f denoted as *terminal region* or *terminal constraint set*. In terms of the optimal control problem formulation (6.2), the terminal constraint is implemented via constraints of the form $g_{terminal}(x(N_p)) \leq 0$ or in the form of its equality constrained counterpart $h_{terminal}(x(N_p)) = 0$.

It must be noted here that, sometimes, these constraints are accompanied by conditions on some (local) controller that never appears explicitly in the optimal control problem formulation. For that reason, these conditions are not mentioned here. An advantage of the stability constraints is that these do not depend on any tuning parameter (Tuning parameter independent stability is denoted in MPC literature by *guaranteed* stability). Also, by using stability constraints one can choose the prediction horizon much shorter than with the pseudo-infinite horizon approach, which may lead to a significant decrease in computation time for stability constrained NMPC control strategies. A disadvantage of the usage of stability constraint approaches seems to be that these are less easy to implement than the pseudo-infinite horizon approach.

An important note for the stability approaches discussed here is that these guarantee stability only under *nominal* conditions, *i.e.* under the assumptions of state feedback and absence of model error, while no such results are available yet for the practically more relevant cases of output feedback and presence of model error. Current research in the field of NMPC is largely focused on deriving stability conditions for these cases. In particular, *robust stability*, *i.e.* stability in the presence of model error, is subject of research while stability for the output feedback case is addressed by establishing some sort of *separation principle/theorem* (or *certainty equivalence principle*. See *e.g.* [93]) for nonlinear systems. Such a theorem says, roughly stated, (when considering stability) that a separately designed stable state estimator combined with a separately designed stable state feedback controller leads to a stable closed-loop system. Such a separation theorem is known to hold for LQG controllers but not for other types of controllers, in particular not for nonlinear ones. These non-nominal stability issues are not further adressed here. Instead, it is assumed that the basic rules set out by the stability approaches discussed here can also be applied to establish stability under non-nominal conditions.

For NMPC based MSWC plant combustion control it is proposed to employ a combination of the pseudo-infinite horizon and stability constraint approach to establish stability. More specific, it is proposed to first of all apply a terminal penalty function of the form

$$\begin{aligned} l_{terminal}(.) \quad = \quad & W_{M_{CH_yO_z}}(M^{sp}_{CH_yO_z} - M_{CH_yO_z}(N_p))^2 + \\ & W_{M_{mois}}(M^{sp}_{mois} - M_{mois}(N_p))^2 \end{aligned} \tag{6.38}$$

with light *c.q.* small weighting factors $W_{M_{CH_yO_z}}$ and $W_{M_{mois}}$ to prevent runaway behavior of the corresponding (state) variables $M_{CH_yO_z}$ resp. M_{mois} and desired (steady-state) values $M^{sp}_{CH_yO_z}$ resp. M^{sp}_{mois}. The latter values could be chosen on the basis of *e.g* maximum grate weight, hence lifetime, and/or fire extinction considerations. Additionally, to ensure stability of the remaining variables, it is proposed to use the pseudo-infinite horizon approach with N_p chosen equal to *e.g.* 60 (which was found to be sufficiently large in simulations) for a sampling period of $60\ s$. In terms of time scales, the terminal penalty functions are used to stabilize the relatively slowly varying variables $M_{CH_yO_z}$ resp. M_{mois} while the pseudo-infinite horizon approach is used to stabilize the faster - remaining - MSWC plant (state) variables: the latter approach was found not to allow stabilization of $M_{CH_yO_z}$ and M_{mois} due to, apparently, the too low value for N_p. The combined approach has been found to be easy to implement, despite the presence of controller performance tuning parameter dependency, and, as observed from simulations, to lead to a sufficiently low NMPC computation time (well below the sampling period), despite the required relatively high prediction horizon (N_p).

6.3.7 Summary of the strategy

The main characteristics of the MSWC plant NMPC strategy proposed here are:

- *Optimal control problem formulation*:
  A comprehensive optimal control problem formulation taking into account any kind of market situation, in particular the over- and undercapacity one, is prefered. However, one may also choose one of the following specific, non-comprehensive optimal control problem formulations for more ease of implementation:
  - a setpoint tracking type of optimal control problem formulation in case of an overcapacity market situation.
  - a constrained maximization / constraint pushing type of optimal control problem formulation in case of an undercapacity market situation
- The usage of the sequential method to numerically solve the optimal control problem, with the required derivatives computed by means of sensitivity equations. The corresponding NLP is solved through an SQP method.
- The usage of either an UKF or EnKF as the state estimator.
- *Offset-free control*
  - The usage of integrating disturbance models in the state estimator to remove offsets towards setpoints induced by model errors and/or unmodeled disturbances. A specific state estimator model including such disturbance models is proposed for MSWC plant NMPC combustion control applications.
  - Manual (re)setting the back-off term for the dominating constraint, or simply the corresponding constraint value itself, to remove offsets towards such constraints induced by model errors and/or unmodeled disturbances.
- Closed-loop stability is enforced by a combination of (i) choosing a sufficiently high prediction horizon in the NMPC optimal control problem formulation and (ii) incorporating terminal penalty functions.

6.4 A linear model predictive control strategy for MSWC plants

To obtain a suitable MSWC plant *L*MPC control strategy one only needs to make a few rather trivial adaptations to its *N*MPC counterpart proposed above. These adaptations are discussed here, which, together with the unadapted parts of the NMPC strategy proposed above, form the MSWC plant LMPC strategy proposed here.

The first adaptation is the translation of the optimal control problem into an LP (linear programming: linear objective function and constraints) or QP problem, depending on the (optimal) control problem formulation at hand, and the application of one of the widely available efficient solvers for such problems. In particular, a comprehensive or a setpoint tracking type of optimal control problem formulation as considered here translates to a QP problem and a constrained maximization type of optimal control problem formulation as considered here translates to an LP. In contrast to its NMPC counterpart, no (generic type of) NLP needs to be solved through *e.g.* an SQP method.

A minor adaptation is also the formulation of the optimal control problem in terms of deviation variables, *i.e.* deviations from a specific *a priori* chosen operating point.

A third adaptation is the usage of a KF as a state estimator, being the linear version of both an UKF and EnKF.

The remaining main characteristics of the NMPC strategy proposed above, *e.g.* with respect to obtaining offset-free control and closed-loop stability, require little to no main adaptations.

6.5 Conclusions

In this chapter, both a linear and nonlinear model predictive control based MSWC plant combustion control strategy have been proposed and discussed in detail. These strategies follow a standard moving horizon control strategy where the MVs are computed by repeatedly solving a finite-horizon open-loop optimal control problem for newly determined plant state and disturbance values. A state estimator is employed to predict the values for these states and disturbances for the next sampling instant from measurements available until and at the current sampling instant. The optimal control problem is subsequently solved for these values with all future disturbance values chosen equal to the ones predicted for the next sampling instant. At the latter sampling instant, the resulting optimal MV trajectory is then implemented on the plant to be controlled until at the sampling instant after a new MV trajectory is implemented that has been computed for newly predicted state and disturbance values. The proposed MSWC plant *nonlinear* model predictive control strategy is further characterized by:

- The usage of
  - a setpoint tracking type of optimal control problem formulation in case of an overcapacity market situation, or
  - a constrained maximization / constraint pushing type of optimal control problem formulation in case of an undercapacity market situation, or

  - a comprehensive optimal control problem formulation taking into account any kind of market situation, including the two ones just above

  of which the third formulation is prefered but the first two ones are more easy to implement.

- The usage of the sequential method to numerically solve the optimal control problem, with the required derivatives computed by means of sensitivity equations. The corresponding nonlinear programming problem is solved through a sequential quadratic programming method.

- The usage of either an Unscented Kalman filter or Ensemble Kalman filter as the state estimator.

- The usage of integrating disturbance models in the state estimator to remove offsets towards setpoints. A specific state estimator model including such disturbance models has been proposed for MSWC plant NMPC combustion control applications. To remove offsets towards dominating constraints, in particular in case operation is dominated by a constrained maximization type of control, manual (re)setting the back-off term for the dominating constraint, or simply the corresponding constraint value itself, is proposed.

- Closed-loop stability is enforced by a combination of (i) choosing a sufficiently high prediction horizon in the NMPC optimal control problem formulation and (ii) using terminal penalty functions.

The proposed MSWC plant *linear* model predictive control strategy is the same as the proposed NMPC strategy except for the following rather trivial adaptations:

- the translation of the optimal control problem to a linear or quadratic programming problem, depending on the considered optimal control problem formulation, and application of one of the widely available efficient solvers for such problems.

- the usage of a Kalman filter as a state estimator

A minor adaptation is also the formulation of the optimal control problem in terms of deviation variables.

Chapter 7

The added value of MPC for MSWC plant combustion control

7.1 Introduction

In this chapter, opportunities of model predictive control (MPC) for improving the combustion control and overall economic performance of MSWC plants are explored. This is done through assumed close-to-realistic simulations involving the linear and nonlinear MPC strategies presented in the previous chapter. In principle, MPC allows for an at least equally well to better combustion control and, thereby, overall economic performance for MSWC plants compared to conventional control strategies, *e.g.* of the PID type. This is due to its systematic handling of constraints and (due to the exploitation of model predictions) any characteristic of the plant to be controlled, *e.g.* nonlinearity, multi-variability, interaction, delays, inverse responses, etc. However, it remains to be answered still how this in principle better control performance *c.q.* these systematic handling capabilites, of MPC can actually be translated to a better MSWC plant combustion control and overall economic performance and what the extent of improvement actually can be. These questions are the main ones addressed in this chapter.

The contents of this chapter are as follows. First, in section 7.2, specifics of the considered simulations are given. In section 7.3, it is discussed how the systematic constraint handling capabilities of MPC can be translated to a better MSWC plant combustion control and overall economic performance, both for an under- and over-capacity market situation (see section 6.3.2). Then, in section 7.4, it is discussed how the systematic handling capabilities of MPC can be translated into improved process variation minimization properties for MSWC plant combustion control systems, in particular into improved setpoint deviation minimization properties in conditions where constraints do not play a role, and what the consequences are for the two considered

MSWC plant market situations. The main conclusions of this chapter are then given in section 7.5.

7.2 Setup of the simulations

Throughout this chapter, simulations are used to discuss the added value of MPC for MSWC plant combustion control. These simulations involve the linear and nonlinear MPC based MSWC plant combustion control strategies of the previous chapter and additionally, for comparison purposes, the new PID control based MSWC plant combustion control strategy proposed in chapter 5. In this section, main features of these simulations are discussed to provide the required background for the discussions to come. Further details on the simulations can be found in the remainder of this chapter.

First of all, in all simulations, the new first-principles model proposed in chapter 2 acts as the plant to be controlled. Secondly, the sample time employed during the simulations is 60 seconds, which is a common value for MSWC plant combustion control applications. Also, a zero-order-hold condition applies to the MVs and DVs during the simulations, *i.e.* the MVs and DVs are held constant during each sample interval.

The MVs of the plant to be controlled are the waste inlet flow $\phi_{w,in}$ $[kg/s]$, interfacial area a $m^2/kg]$, primary air flow ϕ_{prim} $[kg/s]$ and secondary air flow ϕ_{sec} $[kg/s]$ (see also chapter 2). Here, $\phi_{w,in}$ and a are different from the corresponding ram resp. grate speed MVs typically encountered in practice, which is motivated by the fact that the new first-principles model from chapter 2 acts as the plant to be controlled and not a real-life MSWC plant or empirical model derived thereof. It is noted that during first-principles model validation work (performed by the author of this thesis on the new model proposed in chapter 2 and using the MSWC plant system identification methodology proposed in chapter 3) a has been found to be well modeled as a linear function of the grate speed MV $[\%]$ employed at several large scale MSWC plants. The DVs of the plant to be controlled during the simulations are the secondary air flow/air leakage flow/waste inlet flow and ambient temperatures $T_{sec} = T_{leak} = T_{w,in} = T_{air}$ (all in $[K]$), the relative humidity of the air RH_{air} $[\%]$, the inert fraction of the waste X_{inert} $[-]$, the moisture fraction of the waste X_{mois} $[-]$ and the waste composition variable z $[-]$ (again, see also chapter 2). CVs *c.q.* variables that can be of relevance from a control point of view are the steam flow ϕ_{st} $[kg/s]$, the flue gas oxygen concentration O_2 $[\%]$, the solid waste layer temperature T_s $[K]$, the amount of combustible waste on the grate $M_{CH_yO_z}$ $[kg]$, the amount of moisture in the waste on the grate M_{mois} $[kg]$, the carbon dioxide content in the flue gas CO_2 $[\%]$, the water content of this gas H_2O $[\%]$, the flue gas flow ϕ_{fg} $[kg/s]$ and the gas phase temperature T_g $[K]$, where typically ϕ_{st} and O_2 are the main CVs of interest. Also, $M_{CH_yO_z}$, M_{mois}, T_s and ϕ_{st} are state variables in the considered models.

The DVs in the simulations are mostly represented by assumed close-to-realistic stochastic processes, to render close-to-realistic simulation results, but in some cases constant and step type of signals are used to enhance the illustrative value of the corresponding simulations. When represented by stochastic processes, the DVs T_{air} and RH_{air} are modeled as stochastic realizations derived from available measurements from a real-life MSWC plant. These, therefore, are considered to well represent real-life realizations of these DVs. See further on in this chapter for example realizations, *e.g.* in figure 7.15. Also when modeled as a stochastic process, the DV z is represented by a stochastic realization derived from calorific value sensor (CVS) estimates available from large scale MSWC plant data. Therefore, these realizations are also considered good representatives of their true counterpart. See, again, further on in this chapter for an example realization. Both X_{inert} and X_{mois} are chosen as a zero mean white noise (ZMWN) signal added to a non-zero nominal/operating point value, when represented by stochastic processes. The reason for this choice is that no measurements or estimates have been available for these DVs. The white noise assumption is thought to be valid because of the physical interpretation of X_{inert} and X_{mois} *c.q.* because of the (here) assumed white noise nature of the waste composition. The nominal values for X_{inert} and X_{mois} are both chosen equal to 0.2, which is in the range of values commonly used in MSWC plant calculations. The amplitudes for the ZMWN parts of these signals are chosen such that the standard deviations of ϕ_{st} and O_2 resulting from closed-loop runs with the new PID based MSWC plant combustion control system of chapter 5 and with all mentioned DV realizations are of the same order of magnitude as commonly encountered in practice (*i.e.* roughly between 0.1 and 1.5 for both CVs).

Stability of the considered MPC based MSWC plant combustion controllers is guaranteed by means of the measures mentioned in section 6.3.6 *c.q.* by means of using a prediction horizon N_p of 60 and the usage of a terminal penalty function (6.38), involving $M_{CH_yO_z}$ and M_{mois}.

Move blocking (see *e.g.* [11]) is applied to the MVs in the considered MPC optimal control problems to limit the number of corresponding optimization variables and, thereby, the overall computation time. With move blocking, consecutive (future) MVs are forced to be equal. Care must be taken not to perform too many such move blocking enforcements as that may lead to a significant degradation of the controller performance. Here, each MV is allowed to switch only at the instants $i = 1, 2, 3$, ..., $16, 18, 20, 22$, ... , $30, 35, 40$, ..., 55, leaving a total number of $4 \times 28 = 112$ variables to be optimized by the MPC optimal control problem solver.

The considered NMPC control strategies employ the same model for predictions as the one used as the plant-to-be-controlled. The considered LMPC control strategies employ a linearized version of this model with the linearization operating point within or close to the operating range eventually encountered during closed-loop simulations with the considered control strategies. Both linear and nonlinear control strategies are considered to assess the influence of the presence of nonlinear MSWC plant dynamics on the MPC based combustion control performance.

LMPC control strategies both with and without integral action are considered. Integral action is used to, first of all, avoid offsets in the CVs introduced by plant-model mismatch. Secondly, it can be used as a means for positively influencing the closed-loop (L)MPC performance, either through additional filtering or through disturbance modeling. In fact, as has been observed during the simulations (see also further on in this chapter), the addition of integral action can play an important role in obtaining a good process variation minimization performance.

Both state- and output feedback MPC control strategies are considered. The considered state feedback MPC strategies refer to strategies that have the availability of perfect one-sample-ahead predictions of the (initial) state and disturbance values used by the corresponding optimal control problems. These conditions allow for an assessment of the performance of MPC based combustion control strategies in an idealized setting, in particular for removing the effect of state and disturbance estimator errors using, additionally, noisy measurements. In contrast, output feedback MPC combustion control strategies are used here to more realistically represent the performance of real-life application of MPC on MSWC plant combustion control problems. These are the same as their state feedback counterparts but do use a state and disturbance estimator and noisy measurements. More specific, a state and disturbance estimator of the EnKF (NMPC) or KF (LMPC) type is employed of the form outlined in section 6.3.5 (at the end of the setpoint offset removal part of this section; eqn. (6.35) and further below this equation). This estimator employs

- ϕ_{st}, O_2, CO_2, H_2O, ϕ_{fg} and T_g as output measurements
- T_{prim}, T_{sec}, $T_{w,in}$ and T_{leak} as input measurements *c.q.* measured DVs, which all are assumed to be equal to and measured through T_{air}
- an estimate for z as input, using the calorific value sensor to deliver this estimate from measurements of O_2, CO_2, H_2O, T_{air} and RH_{air}
- $M_{CH_yO_z}$, M_{mois}, T_s and ϕ_{st} as states to be estimated/predicted
- X_{inert} and X_{mois} as disturbances to be estimated/predicted and included in the state vector to be estimated/predicted

Furthermore, the model employed in the optimal control problem to be solved by the MPC controller (either fully or a linearized version) is also as outlined there. Finally, to represent measurement noise, the measurements employed by the state and disturbance estimator (T_{air}, RH_{air}, O_2, H_2O and CO_2, ϕ_{st}, ϕ_{fg} and T_g) contain additive zero mean white noise (ZMWN) realizations with standard deviations equal to $0.5\ \%$ of the average values of the considered measurements. These values are chosen arbitrarely and assume the measurement noise to be not that large.

The difference between the considered state and output feedback MPC strategies is schematically depicted in figure 7.1.

The length of the simulations typically is $600 - 800$ samples (minutes) to have a good compromise between simulation time and obtaining a good assessment of the performances of the considered control strategies.

7.3 The added value of MPC in constraint handling

The added value of systematic constraint handling by MPC for MSWC plant combustion control applications is discussed here first for the overcapacity market situation, then for the undercapacity market situation.

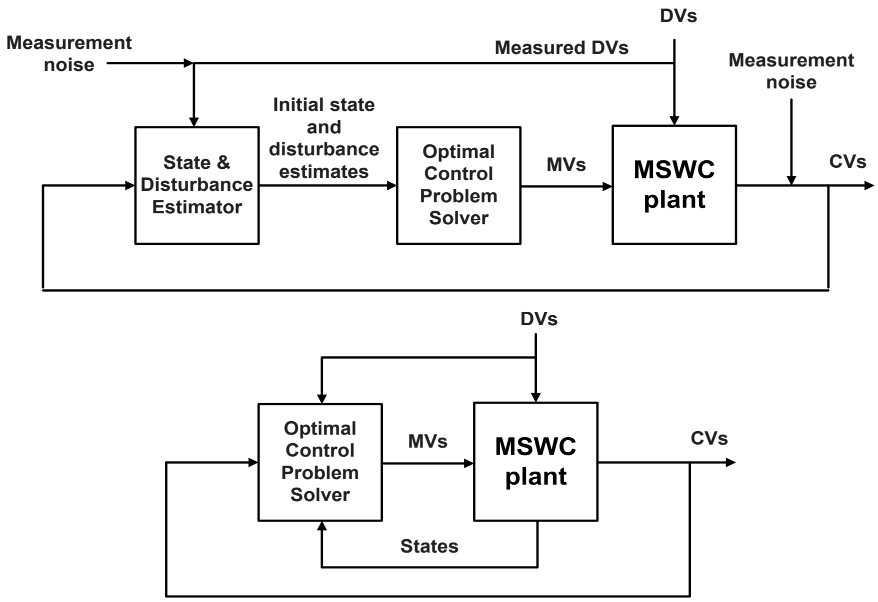

Figure 7.1: Output feedback (top) versus state feedback MPC based MSWC plant combustion control. Note: the connection between CVs and optimal control problem solver in the state feedback controller is used for integral action.

7.3.1 Overcapacity market situation

The added value of systematic constraint handling by MPC for MSWC plant combustion control applications in an overcapacity market situation is discussed here predominantly through two simulation cases. The first one is under idealized and less realistic conditions of a perfect state estimator *c.q.* a state feedback MPC based MSWC plant combustion control strategy and, additionally, with no stochastic DVs but, instead, constant and step(s) type of DVs. These conditions are imposed for the purpose of (easier) illustration. The second simulation case aims to demonstrate the added value of constraint handling by MPC for MSWC plant combustion control applications under (assumed) close-to-realistic conditions, with an output feedback combustion control strategy and stochastic DVs as defined in section 7.2). To accomodate (only and fully) the overcapacity market situation, all considered MPC controllers have a setpoint tracking type of objective incorporated in their optimal control problem formulation as defined by eqn. (6.3) and no income maximization type of objective (6.8). The setpoint values to be maintained here are $\phi_{st}^{sp} = 16\ [kg/s]$ and $O_2^{sp} = 6.5\ [Vol.\%]$, which are chosen arbitrarely but within the range of values commonly encountered in practice. Also, all control strategies considered here are of the *L*MPC type with the only motivation being to add a discussion on the effect of model error on constraint handling by MPC. It is noted that hard constraints on *MVs* are applied as well in the simulations to be discussed but that these do not play a role in the corresponding discussions as these are not violated or touched upon and, additionally and arguably, are considered to be of lower importance when evaluating MPC as a potential control strategy for MSWC plant combustion control applications.

Consider, then, the results obtained for the first - state feedback and constant/step(s) type of DVs - simulation case depicted in figures 7.2 - 7.4. These results have

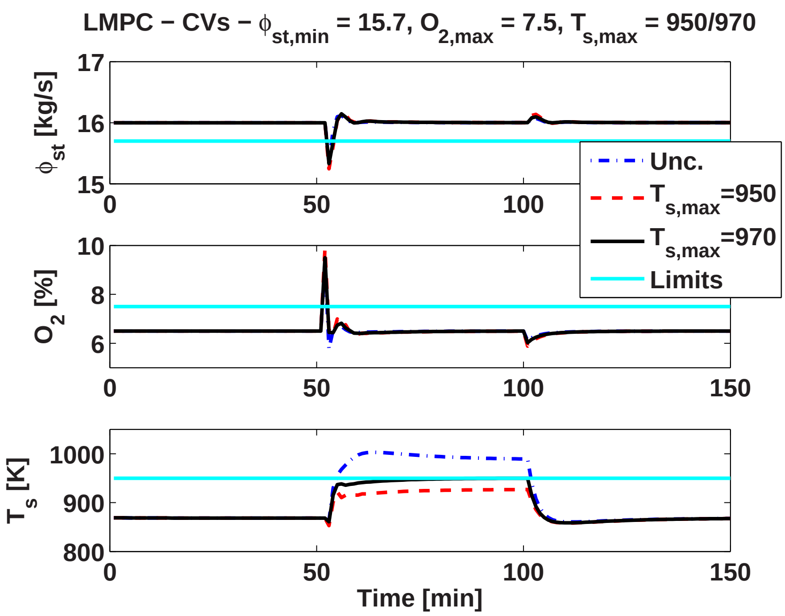

Figure 7.2: CVs obtained with closed-loop simulation with state feedback LMPC based combustion control strategy aimed at an overcapacity market situation: constraint handling properties of the control strategy. See also figures 7.3 and 7.4.

been obtained with three state feedback LMPC based MSWC plant combustion control strategies under the conditions discussed above and subject to constant DVs except for a temporarily large upset in z (see figure 7.4)[1]. Furthermore, each of the LMPC strategies has a different set of CV constraints to fulfill:

- one without any constraint imposed on the CVs (refered to as $'Unc.'$)
- one having to fulfill a lower limit $\phi_{st,min} = 15.7$ $[kg/s]$, an upper limit $O_{2,max} = 7.5$ $[\%]$ and another upper limit $T_{s,max} = 950$ $[K]$ (refered to as $'T_{s,max} = 950'$)
- one with the same constraints on ϕst and O_2 as the one just above but with a different upper limit $T_{s,max} = 970$ $[K]$ (refered to as $'T_{s,max} = 970'$)

[1]These variations in z here correspond to a change in calorific value of the burning waste on the grate $\Delta H_{CH_yO_z}$ of 12 to 27 $[MJ/kg]$ and back, which values are notably not compensated for the amount of water yet.

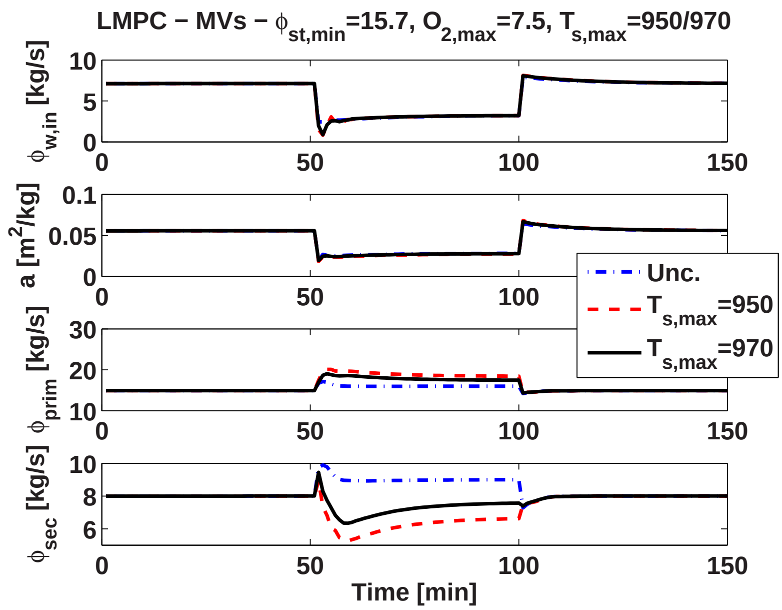

Figure 7.3: MVs obtained with closed-loop simulation with state feedback LMPC based combustion control strategy aimed at an overcapacity market situation, used to determine the constraint handling properties of the control strategy. See also figures 7.2 and 7.4.

It is noted that the actual constraint on T_s to be fulfilled by the considered LMPC control strategies is the upper limit $T_{s,max} = 950\ [K]$. The motivation for considering two LMPC controllers with each a different constraint on this temperature lies in the discussion to come on the effect of model error on constraint handling by MPC. The constraints on ϕ_{st} and O_2 are chosen for illustration purposes only and not with some real-life meaning. In contrast, the constraint on T_s could well represent a real-life limit. More specific, the ability to maintain T_s below a user-defined level can be used to significantly enhance the lifetime of furnace components like *e.g.* the grate bars and furnace walls and, thereby, to significantly reduce the overall costs.

The responses in the figures provide, first of all, an illustration of the property of MPC that it allows for improved constraint handling compared to unconstrained MSWC plant combustion control strategies. More specific, the violation of the limit on T_s obtained with the unconstrained LMPC controller is removed effectively with the other two LMPC controllers incorporating constraints on this temperature. On the other hand, considering the violations of the constraints on ϕ_{st} and O_2, the responses in the figures also indicate that not every constraint to be fulfilled by an MSWC plant combustion control strategy may be fulfilled by an MPC based such strategy, at least not by the MPC implementation proposed in chapter 6. More specific, as results from this and also other simulations have indicated (including simulations with *N*MPC based control strategies), CV constraints cannot be fulfilled by the proposed MPC based MSWC

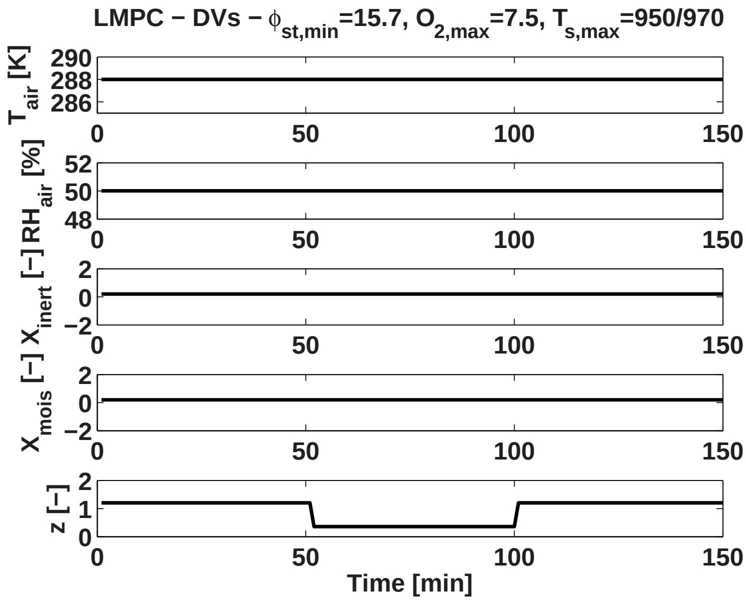

Figure 7.4: DVs used for closed-loop simulation with state feedback LMPC based combustion control strategy aimed at an overcapacity market situation to determine its constraint handling properties. See also figures 7.2 and 7.3.

plant combustion control strategy when the changes in the corresponding CVs are very fast. This is subscribed to the feedback nature of this strategy, *re*-acting to changes rather than *pro*-acting. Slower CV changes can be handled, on the other hand, *c.q.* do not lead to constraint violations as the MPC controller has time to take proper control action.

Another observation from simulations is that incorporating (in the MPC control strategy) constraints on CVs that have to fulfill setpoint tracking objectives, *c.q.* constraints on ϕ_{st} and O_2, have a negligible effect on the fulfillment of these constraints by the MPC controller if the setpoint deviation minimization performance of this controller is good. Consequently, if the latter is true (as in the simulation case discussed here), these constraints may well be removed from the MPC optimal control problem formulation without significantly affecting the handling of these CV constraints. This can already be observed from the simulation results above from comparing the handling of the constraints on ϕ_{st} and O_2 by the unconstrained MPC control strategy with that by the constrained MPC control strategies. It has also been confirmed through other similar simulation based comparisons between constrained and unconstrained MPC strategies not discussed here. From the observation here it may be concluded that the added value of MPC in fulfilling constraints in MSWC plant combustion control applications in an overcapacity market situation lies particularly in the prevention of constraint violations of CVs that are not subject to setpoint tracking.

From the figures 7.2 - 7.4 it can also be observed that imposing a constraint of $T_{s,max} = 950$ $[K]$ leads to T_s to remain well below this upper limit. Likewise, imposing a constraint of $T_{s,max} = 970$ $[K]$ leads also to T_s to remain well below this upper limit (of 970 $[K]$). In other words, these responses show a persistent distance to be present between the imposed limit and actual CV value. This is subscribed to the model bias between the linear model used in the LMPC control strategy and the nonlinear plant to be controlled. If it is desired to operate eventually/on the longer term *at* rather than from a distance to the constraint one could switch to using a nonlinear MPC strategy (although, in practice, *any* model, linear or nonlinear, is likely to be biased) or, alternatively, one can manually adapt the constraint value (which is equivalent to adding some back-off value to the corresponding limit). In the simulation case discussed here, if it is for some *e.g.* economic reason desired to operate on the longer term exactly at $T_{s,max} = 950$ $[K]$, one could employ an actual constraint of $T_{s,max} = 970$ $[K]$ within the LMPC control strategy to account for offset due to model error as this leads to exact fulfillment of the constraint $T_{s,max} = 950$ $[K]$ as can be observed from the simulations.

Through additional simulations, the (positive) constraint handling results discussed above, in an idealized setting, for the considered state feedback LMPC based combustion control strategies are now re-evaluated for the output feedback case, *i.e.* with application of both the more realistic stochastic disturbances and state and disturbance estimator defined in section 7.2, to evaluate its performance in a close(r)-to-realistic setting. More specific, an output feedback LMPC based control strategy is considered with the same objectives and constraints as its constrained state feedback counterparts used in the state feedback simulation case discussed earlier except for a different limit to be fulfilled on T_s. More specific, the actual limit T_s to be fulfilled is an upper one of 930 $[K]$ (rather than 950 $[K]$) while the constraint implemented in the LMPC controller is set on 924 $[K]$ to remove offset. The corresponding simulation results are depicted in figure 7.5. It clearly can be seen from this simulation that even under the defined close-to-realistic conditions, the limit on T_s is not violated by the output feedback LMPC based MSWC plant combustion control strategy. The figure also depicts simulation results obtained with the new PID combustion control strategy of chapter 5 and with the unconstrained counterpart of the considered output feedback LMPC control strategy and under the same experimental conditions (disturbances, plant to be controlled) as applied to the latter control strategy. Note that this PID and unconstrained LMPC control strategy are not able to maintain T_s below the required upper limit and must accept a large violation, indicating that MPC allows for a large improvement in constraint handling performance compared to conventional control strategies in MSWC plant combustion control applications.

A final note here is that, whereas the simulation results here involve mainly only linear MPC based control strategies, nonlinear ones will only improve upon the positive constraint handling performances observed for these strategies, although not sufficient with respect to handling very fast CV changes as discussed above.

To briefly summarize the results above, constraints on MSWC plant CVs can be fulfilled by means of MPC on at least a longer term (violation may temporarily occur for

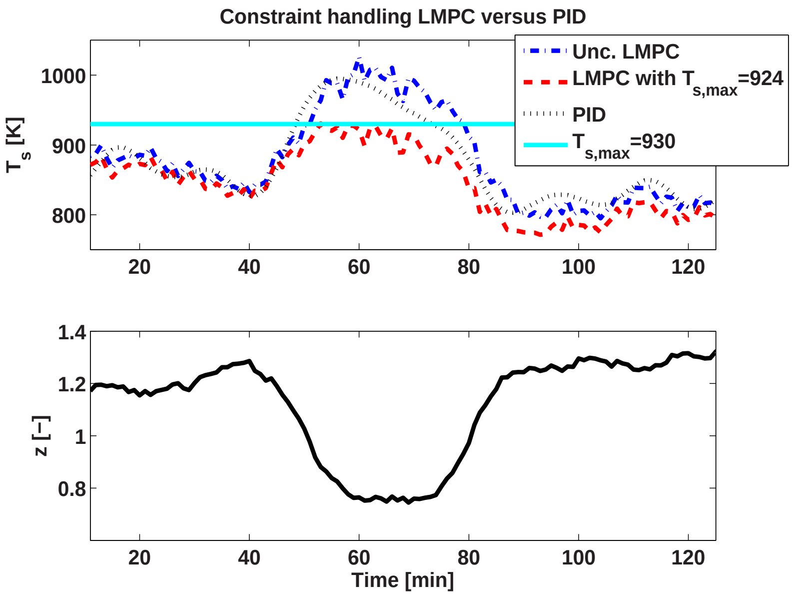

Figure 7.5: CV (T_s) obtained with closed-loop simulation with output feedback LMPC based combustion control strategy aimed at an overcapacity market situation: constraint handling for given DV (z) realization.

very fastly varying CVs/disturbances) for particularly those CVs that are not subject to setpoint tracking. This has been illustrated above through the simulation examples employing (in particular) an upper limit on the solid waste layer temperature T_s, which fulfillment can lead to a significant reduction in maintenance costs. However, this T_s based opportunity for improving the overall economic performance in an overcapacity market situation is not the only one. Particular other opportunities are *e.g.*

- the (complete or at least eventual) fulfillment of upper limits on the masses on the grate, $M_{CH_yO_z}$, M_{mois}, M_{inert}, to prevent overweight (for too long time) on this grate and to, thereby, enhance its lifetime, or to prevent extinguishing of the fire in case of too much M_{mois} and/or M_{inert}.

- the fulfillment of an upper and lower limit on T_g to main a good efficiency of selective non-catalytic reduction (SNCR) equipment used for NO_x removal from the flue gas as the temperature range in which this equipment works well is relatively small (approx. 850 - 1000 $[C]$).

Requirements for any such opportunity to be exploited by an MPC based MSWC plant combustion control strategy are (i) the presence of the corresponding CV in the MPC model as an output and (ii) this CV should not be subject to a setpoint tracking objective that can be tackled well by the MPC controller.

7.3.2 Undercapacity market situation

To discuss the added value of the constraint handling properties of MPC for MSWC plant combustion control applications in an undercapacity market situation, consider the simulation results depicted in figures 7.6 - 7.9. These simulation results have been

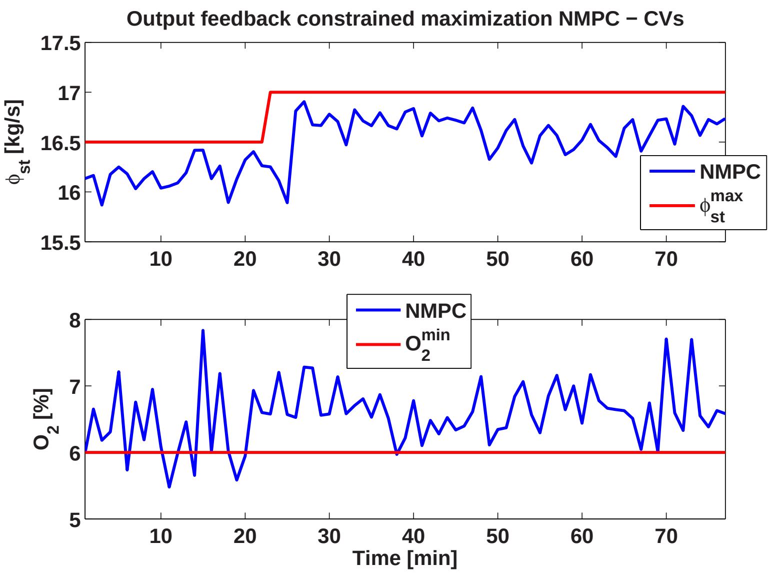

Figure 7.6: CVs obtained with closed-loop simulation with output feedback NMPC aimed at an undercapacity market situation: constraint handling properties of the control strategy. See also figures 7.7 - 7.9.

obtained with an output feedback NMPC based MSWC plant combustion control strategy accomodating the undercapacity market situation with an optimal control problem formulation containing income (and not setpoint tracking) type of stage cost terms of the form of eqn. (6.8) and, additionally, constraints on ϕ_{st} and O_2. These simulation results have been obtained under similar close-to-realistic experimental conditions as employed for the simulations discussed above: see sections 7.2 and 7.3.1. More specific, stochastic disturbances (see figure 7.9), 0.5 $[\%]$ measurement noise and the state and disturbance estimator proposed in section 6.3.5 are employed. Main difference with the previously considered (L)MPC strategies is that the nonlinear model acting as the plant to be controlled also acts as the prediction model in the MPC controller.

The implemented NMPC strategy aims to maximize the steam production and waste inlet flow until a dominating constraint is met to which the operating point of the MSWC plant is pushed as closely as possible by the MPC controller while violating the dominating constraint a minimum number of times. This constraint pushing behavior is particularly visible when the upper limit on ϕ_{st}, which is the dominating constraint here, is increased from 16.5 to 17 $[kg/s]$ at $t = 22$ $[min]$, leading to an immediate jump of the actual value for ϕ_{st} to this constraint: see figure 7.6 (the motivation

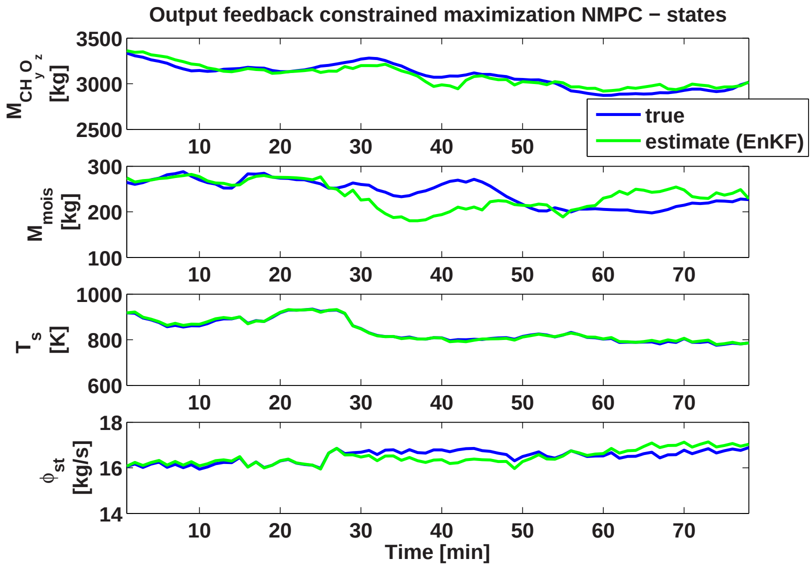

Figure 7.7: States obtained with closed-loop simulation with output feedback NMPC aimed at an undercapacity market situation, used to determine the constraint handling properties of the control strategy. 'true' = actual (state) values resulting from application of NMPC control strategy. 'estimate' = EnKF estimates of these values employed by this strategy. See also figures 7.6, 7.8 and 7.9.

for applying this change in constraint value for ϕ_{st} is actually and only to make this jump / constraint pushing behavior visible, although it could well represent a real-life event).

The simulation results demonstrate the main added value of the systematic constraint handling capacity of MPC for MSWC plant combustion control applications in an undercapacity market situation: that it (simply) allows, together with the ability to include income maximization type of objectives in the optimal control problem formulation, for handling constrained maximization type of optimal control problems. As discussed in section 6.3.2, this type of problem formulation is most suited to handle an undercapacity market situation and, also, control strategies incorporating this formulation, by definition, provide an economically better operation of MSWC plants than other combustion control strategies, including PID type and (other) setpoint tracking type of such strategies. To give a simple motivating example of this potential for an improved economic operation, note that a setpoint tracking type of MSWC plant combustion control strategy would not respond to the change in constraint value for ϕ_{st} in the simulation discussed here but rather would maintain the operating point of the MSWC plant close to the given setpoint values, thereby maintaining the MSWC plant at an economically non-optimal operating point.

Some notes are in order here. First of all, even though having no direct influence

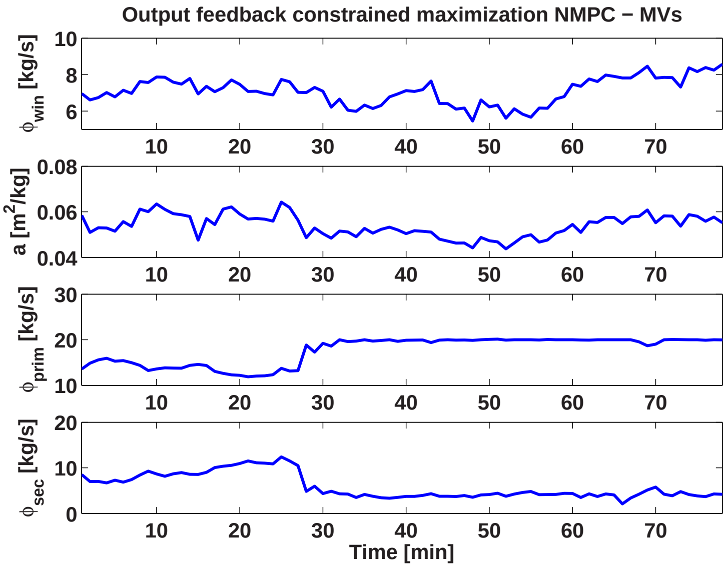

Figure 7.8: MVs obtained with closed-loop simulation with output feedback NMPC aimed at an undercapacity market situation to determine the constraint handling properties of the control strategy. See also figures 7.6, 7.7 and 7.9.

on the overall MSWC plant economics, it was found to be necessary to introduce the lower constraint on O_2, with the reason being that, without it, the furnace would be depleted from oxygen (resulting in bad or no combustion) due to the high amount of waste the corresponding controlled MSWC plant will aim to combust. Notably, it was found that this constraint is not limiting the amount of produced steam and processed waste in normal operation of the MSWC plant as it then can be maintained through the primary and secondary air flows. If, however, these air flows are both at their constraints, the constraint on O_2 can become limiting and prevent the waste inlet flow and steam production, and thereby the overall economic income, to be further maximized. It is also noted that back-off terms were introduced in the considered NMPC control strategy to guarantee a sufficiently low level of constraint violations. These back-off terms may have been chosen somewhat sub-optimally with the back-off towards ϕ_{st}^{max} chosen somewhat too conservatively (wide) here and the back-off towards O_2^{min} chosen somewhat too low (close). A final note here is that the primary air flow MV ϕ_{prim} encounters its upper limit with the jump at $t = 22\ [min]$ and remains there and, also, without (apparent) effect on the overall control performance. This demonstrates the ability of MPC to optimally maintain the *MVs* within imposed bounds as well.

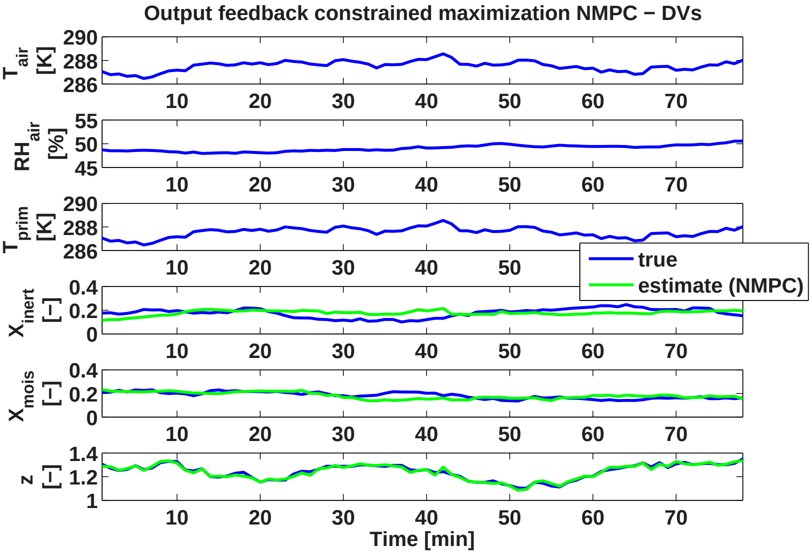

Figure 7.9: DVs for closed-loop simulation with output feedback NMPC aimed at an undercapacity market situation to determine the constraint handling properties of the control strategy. 'true' = actual (DV) values resulting from application of NMPC control strategy. 'estimate' = EnKF estimates of these values employed by this strategy. See also figures 7.6 - 7.8.

7.4 The added value of MPC in unconstrained process variation minimization

MPC may not only add value to the MSWC plant combustion control and overall economic performance in constraint handling sense (see section 7.3 for results), decreasing process variability by minimizing constraint violations, it may also add value in decreasing the process variability when constraints do not come into play. The latter may occur when the MSWC plant disturbances are of a stationary stochastic nature, with relatively low amplitudes, and not of a temporarily large upset type (see section 3.6.2). Reduction of the resulting process variability, which is often expressed in terms of standard deviation or variance and corresponds to the classical disturbance rejection problem, by means of MPC based MSWC plant combustion control may be expected due to its inherent systematic handling of all process characteristics such as interactions, inverse responses, delays, etc. This can be subscribed to the exploitation of predictions from a dynamic model of the plant to be controlled that incorporates all these characteristics. Reduction of the considered type of process variability can lead to a significantly improved overall economic MSWC plant performance, more specific in the overcapacity market situation due to reducing variability induced operational and maintenance costs and in an undercapacity market situation by allowing for closer operation to the dominating constraint. The question addressed here is what the extent is with which MPC can minimize this process variability in MSWC plant combustion control applications, and thereby improve the overall economic performance of these plants. This is mainly done through simulations with MPC based MSWC plant combustion control strategies aimed at handling an overcapacity market situation through employment of a setpoint tracking type of objective in their optimal control problem formulation (and not an income maximization type of objective). Main objective is to determine the deviations from the corresponding ϕ_{st} and O_2 setpoints, in standard deviation sense, in response to MSWC plant disturbances of a stationary stochastic type. The observed performances are assumed to be representative for the overall process variability reduction performance of the considered control strategies when constraints do not come into play. Consequences of the results from the overcapacity market situation case for the undercapacity market situation case are briefly discussed at the end of this section as well.

To have a first investigation of the added value of MPC for reducing the considered type of process variability in MSWC plant combustion control applications in an overcapacity market situation, consider the simulation results depicted in figures 7.10 and 7.11 and quantified in table 7.1. These results have been obtained with state feedback

	open-loop	*PID*	*LMPC (without I)*	*LMPC with I*	*NMPC*
STD(ϕ_{st}) $[kg/s]$	3.0	0.16	0.09 (44%)	0.02 (88%)	0.002 (99%)
STD(O_2) $[\%]$	2.9	0.29	0.29 (0%)	0.06 (79%)	0.002 (99%)
STD($\phi_{w,in}$) $[kg/s]$	0	2.2	1.71	1.74	2.4
STD(a) $[m^2/kg]$	0	0.003	0.01	0.01	0.014
STD(ϕ_{prim}) $[kg/s]$	0	0.57	0.43	0.37	0.97
STD(ϕ_{sec}) $[kg/s]$	0	0.11	0.42	0.38	0.83

Table 7.1: Standard deviations of the MVs and CVs obtained with open-loop run and with state feedback MPC based and PID based MSWC plant combustion control strategies, all aimed at an overcapacity market situation, and under stationary stochastic type of DVs. "with(out) I" = with(out) integral action. Values between parentheses for ϕ_{st} and O_2 (LMPC and NMPC): improvement compared to PID.

setpoint tracking type of MPC controllers, both LMPC and NMPC ones, with setpoints $\phi_{st}^{sp} = 16\ [kg/s]$ and $O_2^{sp} = 6\ [Vol.\%]$, which have been chosen arbitrarily but which are within the range of values commonly encountered in practice, and with no constraints on CVs. All considered control strategies are subject to the same - stationary stochastic type - disturbances, which are of the type defined in section 7.2. Also, both LMPC based MSWC plant combustion control strategies with and without integral (I) action have been evaluated with the integral action incorporated through the first (setpoint adaptation based) method of section 6.3.5.

From the simulation results it can be seen that there is potential for reducing process variability in MSWC plants by means of MPC even when constraints do not come into play. Under the considered idealized conditions, application of an NMPC based

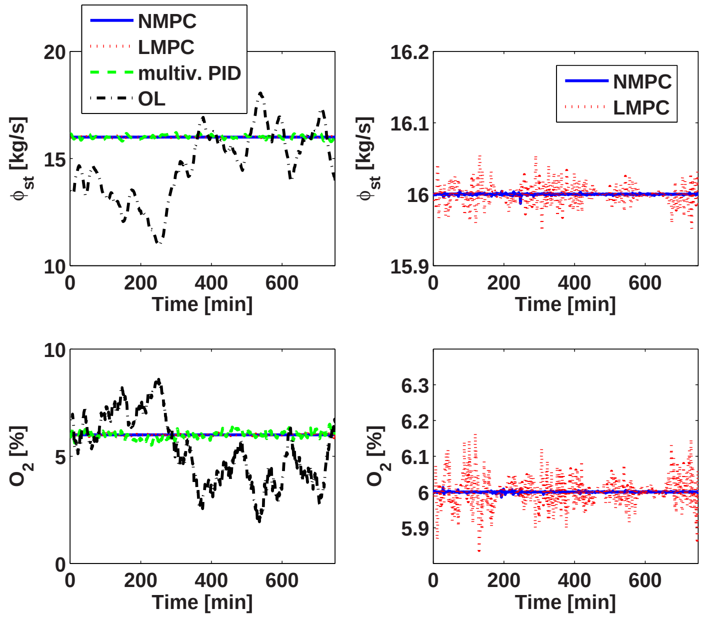

Figure 7.10: CVs obtained with closed-loop simulations under stationary stochastic type of disturbances with a state feedback NMPC based, state feedback LMPC based and a PID based MSWC plant combustion control strategy, all aimed at an overcapacity market situation, and an open-loop (OL; uncontrolled) run. The LMPC based controller here contains integral action. Right figures are details of the left figures and focus only on the MPC performances.

MSWC plant combustion control strategy even leads to almost zero deviations from the considered setpoints. An additional observation is that LMPC based strategies also may allow for a significantly reduced process variability, in particular when integral action is incorporated (which provides an extra degree-of-freedom in tuning the MPC controller and which may be exploited not only for offset removal but also for improving the overall variability minimization properties of the controller). Also note that the NMPC based control strategy performs better than the LMPC based ones, which can be subscribed to the improved handling of MSWC plant nonlinearities. Additionally, note the significantly improved performances of the MPC based combustion control strategies compared to the PID based combustion control strategy, which is the new one proposed in chapter 5. A final note here is that all controllers except for the LMPC one without integral action exhibit offset-free setpoint tracking of both ϕ_{st} and O_2.

The simulation discussed above has been repeated many times for different disturbance realizations, all leading to the same main results as discussed above.

The setpoint deviation minimization performances of the state feedback MPC based combustion control strategies discussed above have also been evaluated under close-

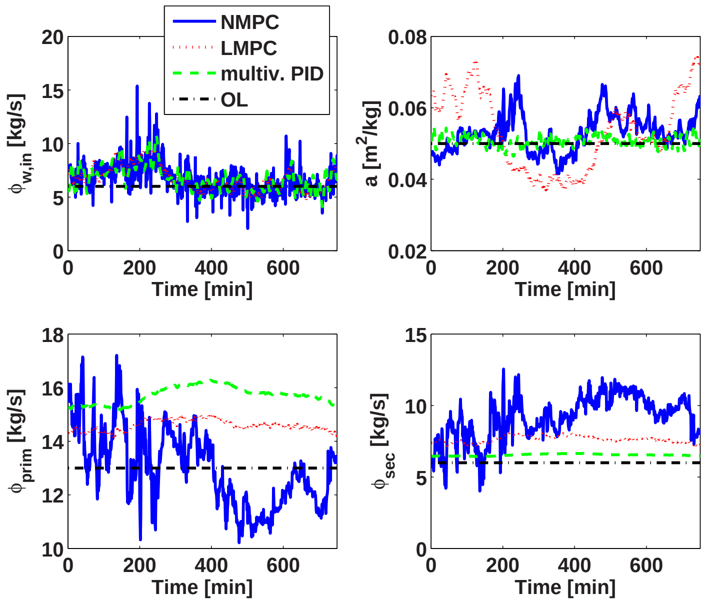

Figure 7.11: MVs obtained with closed-loop simulations under stationary stochastic type of disturbances with a state feedback NMPC based, state feedback LMPC based and a PID based MSWC plant combustion control strategy, all aimed at an overcapacity market situation, and an open-loop (OL; uncontrolled) run. The LMPC based controller here contains integral action. See also figure 7.10.

to-realistic output feedback conditions as specified in section 7.2, more specific with the stochastic disturbances, measurement noise and state and disturbance estimator outlined in this section. Typical simulation results are depicted in figures 7.12 - 7.15 and quantified in table 7.2 which, notably, have been obtained with the aims of (i) at least obtaining a lower variation in ϕst compared to the PID based combustion control strategy while (ii) maintaining the variation in O_2 for the NMPC control strategy at the same level as that of its PID control strategy counter part. It can be seen from these

		open-loop	*PID*	*NMPC*
STD(ϕ_{st})	$[kg/s]$	3.0	0.16	0.13
STD(O_2)	$[\%]$	2.9	0.30	0.30

Table 7.2: Variation in CVs obtained with open-loop simulation and with closed-loop simulations with output feedback NMPC based and PID based MSWC plant combustion control strategies aimed at an overcapacity market situation for stationary stochastic type of disturbances.

results that, for at least this case, the gain in deviation minimization performance is not large, though structural and approximately $19\ [\%]$ in percentage (for ϕ_{st}). In fact, nu-

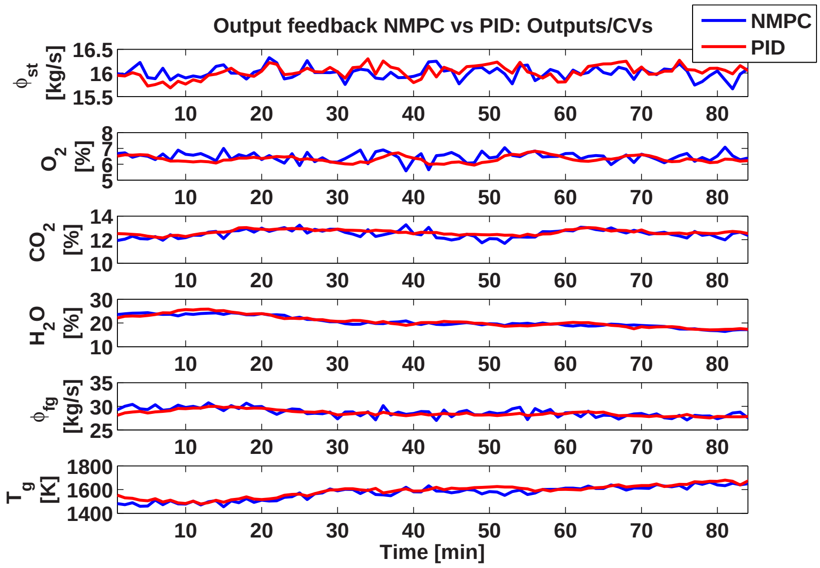

Figure 7.12: CVs obtained with closed-loop simulation with output feedback NMPC based combustion control strategy aimed at an overcapacity market situation: setpoint deviation minimization performance of the control strategy. See also figures 7.13 - 7.15.

merous other simulations (also including LMPC controllers) indicated a similar result leading to the main conclusions that

- MPC allows for a considerable reduction in process variability in MSWC plant combustion control applications (even) under unconstrained conditions and for stationary stochastic type of disturbances
- (However:) a well performing state and disturbance estimator is crucial for actually obtaining this reduction in real-life applications

Hence, if one aims to have a higher reduction in unconstrained process variability minimization performance than encountered with the simulations here, one should focus on further optimization of the state and disturbance estimator employed in the considered MPC strategies, *i.e.* the one proposed in section 6.3.5. A note in this respect is that it particularly was observed that a reduction in variation in O_2 is difficult to obtain with this state and disturbance estimator.

A constraint pushing implementation of an MPC based MSWC plant combustion control strategy, aimed at handling the undercapacity market situation, does not allow an evaluation of its setpoint deviation minimization performance. However, it may be assumed that the setpoint deviation minimization performance disclosed above for the setpoint tracking implementation of an MPC based MSWC plant combustion control

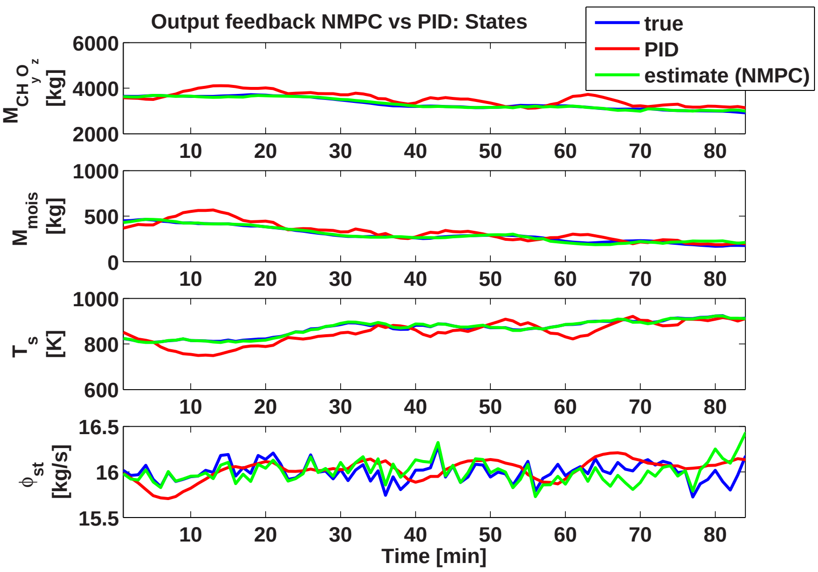

Figure 7.13: States obtained with closed-loop simulation with output feedback NMPC based combustion control strategy aimed at an overcapacity market situation to determine the setpoint deviation minimization performance of the control strategy. 'true' = actual (state) values resulting from application of NMPC control strategy. 'estimate' = EnKF estimates of these values employed by this strategy. See also figures 7.12, 7.14 and 7.15.

strategy carries over to the constraint pushing implementation in the sense that similar process variability minimization properties may be expected. Hence, the conclusions mentioned just above are expected to hold for this implementation as well.

7.5 Conclusions

The main conclusion from this chapter is that MPC allows for a significant improvement in MSWC plant combustion control and overall economic performance, both for an over- and undercapacity market situation. In particular, MPC allows for an improved constraint handling of CVs not subject to setpoint tracking when employing a setpoint tracking type of optimal control problem formulation to optimally deal with an overcapacity market situation. This ability can be used to significantly reduce MSWC plant downtime and maintenance costs, *e.g.* by maintaining furnace temperatures below a certain maximum level to increase the lifetime of furnace components. Additionally, the ability to include constraints in an MPC based MSWC plant combustion control strategy allows, together with its flexibility in manageable optimal control problem formulations, for application of constraint pushing/constrained maximization type of optimal control problem formulations which, thereby, allows for economically opti-

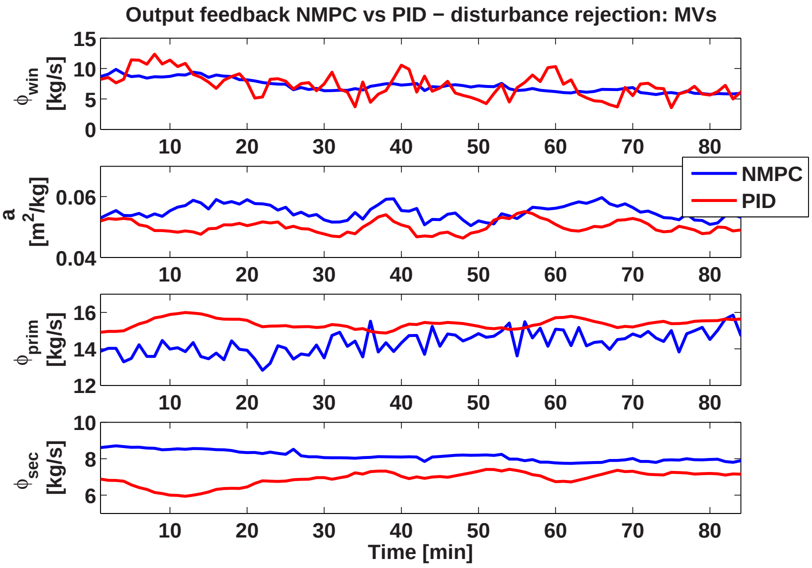

Figure 7.14: MVs obtained with closed-loop simulation with output feedback NMPC based combustion control strategy aimed at an overcapacity market situation to determine the setpoint deviation minimization performance of the control strategy. See also figures 7.12, 7.13 and 7.15.

mally handling by MSWC plant operators of an undercapacity market situation. MPC also allows for a significant reduction in process variability in MSWC plant combustion control applications when constraints do not come into play and, thereby, also for an improved MSWC plant combustion control and overall economic performance, both for an over- and undercapacity market situation. This is due to the resulting lower process variability induced operation and maintenance costs and, in an undercapacity market situation, due to the ability to operate more closely to the economically optimal operating point *c.q.* the dominating constraint. A well performing state and disturbance estimator is crucial for actually obtaining a large reduction in process variability in real-life applications.

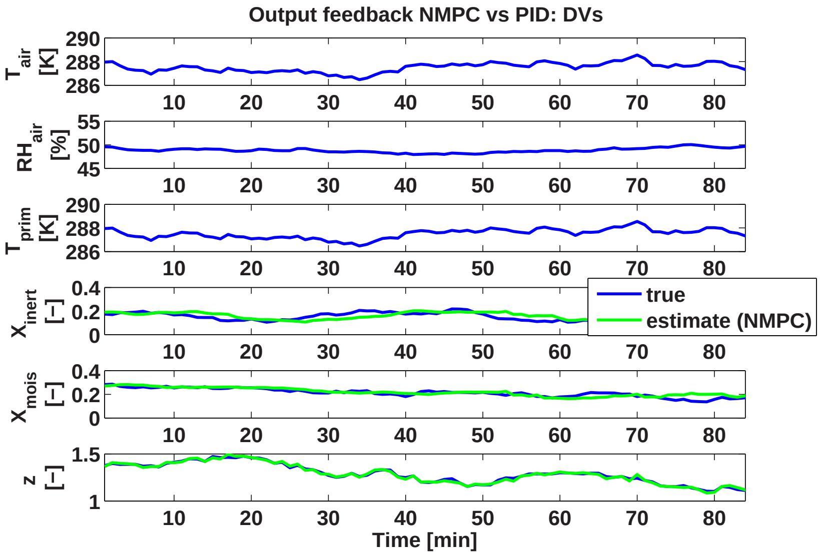

Figure 7.15: DVs for closed-loop simulation with output feedback NMPC based combustion control strategy aimed at an overcapacity market situation to determine the setpoint deviation minimization performance of the control strategy. 'true' = actual DV values. 'estimate' = EnKF estimates of these values employed by the NMPC strategy. See also figures 7.13 - 7.15.

Chapter 8

Conclusions and recommendations

8.1 Conclusions

The main conclusion from this thesis is that the combustion control and overall economic performance of MSWC plants, both for an under- and overcapacity market situation, can significantly be improved by means of model based control, in particular by means of model predictive control (MPC). This is due to the systematic handling of this control strategy of constraints and characteristics of the plant to be controlled such as *e.g.* nonlinearity, multi-variability, interaction, delays, inverse responses, etc. The absence of this property at conventional MSWC plant combustion control strategies has been identified as a major cause for preventing a better combustion control and overall economic performance of these plants. In an overcapacity market situation, MPC based MSWC plant combustion control particularly allows for an improved constraint handling of controlled variables not subject to setpoint tracking. This ability can be used to significantly reduce MSWC plant downtime and maintenance costs, *e.g.* by maintaining furnace temperatures below a certain maximum level to increase the lifetime of furnace components. Additionally, the ability to include constraints in an MPC based MSWC plant combustion control strategy allows, together with its flexibility in manageable optimal control problem formulations, for application of constraint pushing/constrained maximization type of optimal control problem formulations which allows for an economically optimal handling by MSWC plant operators of an undercapacity market situation. MPC also allows for a significant reduction in process variability in MSWC plant combustion control applications when constraints do not come into play and, thereby, also for an improved MSWC plant combustion control and overall economic performance, both for an over- and undercapacity market situation. This is due to the resulting lower process variability induced operation and maintenance costs and, in an undercapacity market situation, due to the ability to operate more closely to the economically optimal operating point. A well performing state and disturbance estimator is crucial for actually obtaining a large reduction in process variability in real-life applications.

Additionally, it has been found that by means of a new model based PID-type of combustion control strategy developed in this thesis the setpoint tracking properties of commonly used PID combustion control strategies can be improved substantially, and thereby the overall economic performance of MSWC plants in an undercapacity market situation. However, to improve upon the process variability minimization properties of the currently used PID combustion control strategies, and thereby making a further and large improvement in economic performance, non-PID type of combustion control strategies are required.

Also, in order to obtain models suitable for model based MSWC plant combustion control, two approaches have been considered and found to be valid: first-principles modeling and (linear) system identification. In particular, a new first-principles model suitable for model based MSWC plant combustion control has been developed that is an extended version of an available literature model. The extension involves the incorporation of the equations underlying the so called calorific value sensor (CVS). The latter is an on-line estimator of the MSWC plant waste composition and calorific value. This incorporation leads to a more detailed description of the waste composition in the model which, combined with the ability to estimate the main parameter of this description from large scale MSWC plant plant data through the CVS, allows for an improved simulation, validation and model based combustion control performance compared to existing models. Additionally, a specific system identification methodology has been developed to derive models suitable for MSWC plant combustion control. This methodology has been derived via identification of opportunities in the literature for overcoming all potential obstacles for arriving at a model suitable for model based MSWC plant combustion control through system identification. Through an application of the proposed first-principles modeling and system identification approaches on data experimentally obtained from a large scale Dutch MSWC plant, it was found that with both these modeling approaches it is possible to derive a model suitable for MSWC plant combustion control.

Another specific contribution of this thesis is an analysis of and solution approach to the occurrence of bias in case of partial closed-loop identification data, which type of data can be experimentally obtained at MSWC plants. It has been shown that, generally and certainly in the case of system identification of MSWC plants, partial closed-loop identification data should be treated in the same way as the commonly considered (completely) closed-loop identification data.

Also, validation of first-principles MSWC plant models directly on the basis of comparing simulated model outputs with their measured counterparts is generally impossible due to the presence of large nonmeasured disturbances on MSWC plant data. With other validation methods not being directly available, a new system identification based method of validating first-principles MSWC plant models has been developed in this thesis to overcome this validation problem.

Other main contributions of this thesis are both a linear and nonlinear MPC strategy for properly tackling MSWC plant combustion control problems, both for an under- and overcapacity market situation. These strategies follow a standard moving horizon control strategy where the manipulated variables are computed by repeatedly solving a finite-horizon open-loop optimal control problem for newly determined plant state and disturbance values. A specific state (and disturbance) estimator has been developed to predict the values for these states and disturbances for the next sampling instant from measurements available until and at the current sampling instant.

8.2 Recommendations for future research

From the work presented in this thesis, the following interesting future research directions have been identified:

- Determination of the true potential of the new PID based MSWC plant combustion control strategy proposed in this thesis by applying it on a real-life MSWC plant.
- Exploration of the opportunities of using the CVS, or other on-line disturbance estimator, in combination with PID based MSWC plant combustion control to enhance the performance of the latter.
- Further optimization of the state and disturbance estimator used in the developed MPC based MSWC plant combustion control strategies to exploit more the potential of MPC for minimizing MSWC plant process variations.
- Industrial verification of the true performance improvement of MPC based combustion control by applying it on a real-life MSWC plant, *i.e.* under truly realistic conditions.
- Application of nonlinear system identification techniques as an alternative to or validation tool for first-principles MSWC plant models. A particularly interesting research direction here may be the usage of linear parameter varying (LPV) models with the varying parameter being the composition parameter estimated with the calorific value sensor. This direction is particularly motivated by the recent occurrence of practically feasible estimation methods for such models, see *e.g.* [124], and the availability of a parameter that to a large extent determines the MSWC plant dynamics over its operating range *c.q.* the waste composition parameter estimated by the CVS.
- Experiment design for system identification of MSWC plants to maximize the ratio between model accuracy and experiment length while fulfilling operational constraints. (See *e.g.* [9] for recent work on this subject).

148

APPENDIXES

Appendix A

Approximate realization via step response data

A.1 Introduction

The MSWC plant system identification procedure proposed in chapter 3 employs a step responses based *approximate realization* algorithm due to [104] for model reduction. Here, 'realization' refers to the result of the algorithm being a state space realization of the dynamics to be modeled and 'approximate' refers to the fact that this realization is an approximation to these dynamics. In this appendix, for completeness, the considered approximate realization algorithm is outlined. Also, as a novelty, a link is established between between this algorithm and a major group of subspace methods. More specific, conditions are presented under which a member of this major group of subspace methods is equivalent to this approximate realization algorithm, implying that the latter algorithm can be regarded a subspace method.

The approximate realization algorithm of [104] is a modified version of the well-known pulse response data based approximate realization algorithm of Kung [49]. The latter algorithm, in return, is based on another well-known realization algorithm, namely the one by Ho and Kalman [36]. Therefore, in order to give a good idea of the relations between these methods, first the basic ideas behind Ho-Kalman's and Kung's algorithm are explained here before outlining the realization algorithm of [104]. Finally, the link between this algorithm and subspace methods is provided.

A.2 The Ho-Kalman algorithm

The problem the Ho-Kalman algorithm [36] aims to solve is the following one: given the pulse response coefficients $\{G_t\}_{t=0,1,\ldots,\infty}$ of some finite-dimensional, LTI, discrete-time (multivariable) dynamical system $G(z)$, *i.e.* the coefficients of the transfer function matrix

$$G(z) = \sum_{t=0}^{\infty} G_t z^{-t} \tag{A.1}$$

which are also referred to as *Markov parameters*, construct a minimal state space model

$$\begin{array}{rcl} x(t+1) & = & Ax(t) + Bu(t) \\ y(t) & = & Cx(t) + Du(t) \end{array} \tag{A.2}$$

i.e. a minimal *realization*, of this system. Minimal means that no state space model of lower state dimension can be found to describe the system (see *e.g.* [93] for a formal definition). The state dimension of any minimal realization of a system $G(z)$ is also called the *McMillan degree*. Note that the system is assumed disturbance free; the availability of disturbance-free Markov parameters is assumed.

The key to the solution to this problem is the following relation between the Markov parameters and the system matrices of the state space model (A.2):

$$G_t = \begin{cases} D & (t = 0) \\ CA^{t-1}B & (t > 0) \end{cases} \tag{A.3}$$

A first consequence of this relation is that the system matrix D is simply found as G_0. As a result, we only have to focus yet on the question of how to obtain the system matrices A, B and C. These matrices can be derived using another consequence of the relation (A.3) which is that the following matrix

$$H_{n_r,n_c} := \begin{pmatrix} G_1 & G_2 & \dots & G_{n_c} \\ G_2 & G_3 & \dots & G_{n_c+1} \\ \vdots & \vdots & \ddots & \vdots \\ G_{n_r} & G_{n_r+1} & \dots & G_{n_c+n_r-1} \end{pmatrix} \tag{A.4}$$

can be written as the product

$$H_{n_r,n_c} = \Gamma_{o,n_r} \cdot \Gamma_{c,n_c} \tag{A.5}$$

with

$$\Gamma_{o,n_r} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n_r-1} \end{pmatrix} \tag{A.6}$$

and

$$\Gamma_{c,n_c} = \begin{pmatrix} B & AB & \dots & A^{n_c-1}B \end{pmatrix} \tag{A.7}$$

The matrix (A.4) is also referred to as the (block) Hankel matrix of Markov parameters. A Hankel matrix is a matrix of which all (scalar or block) entries on a skew diagonal are the same. The matrix Γ_{o,n_r} is also known as the *observability matrix* and Γ_{c,n_c} is known as the *controllability matrix* (or sometimes *reachability matrix*).

One important characteristic of the Hankel matrix (A.4) is that for sufficiently large n_r and n_c the rank of this matrix is equal to the McMillan degree n of $G(z)$. In addition, if n_r and n_c are this large, it can be shown that for any full rank matrix decomposition

$$H_{n_r,n_c} = H_1 H_2 \tag{A.8}$$

with $H_1 \in \mathbb{R}^{n_r p \times n}$ and $H_2 \in \mathbb{R}^{n \times n_c m}$ satisfying

$$rank\ H_1 \ = \ rank\ H_2 \ = \ rank\ H_{n_r,n_c} \ = \ n \tag{A.9}$$

there exist matrices A, B and C from an n-dimensional state space model such that

$$H_1 = \Gamma_{o,n_r} \tag{A.10}$$

and

$$H_2 = \Gamma_{c,n_c} \tag{A.11}$$

A consequence of this result is that, when we are able to decompose somehow the Hankel matrix (A.4) as in (A.8), we already have solved the minimal realization problem due to the fact that the system matrices A, B and C can then be obtained from the matrices H_1 and H_2 as follows:

- The system matrix C can be extracted from H_1 by taking the first p rows of this matrix (recall that p is the number of outputs). This follows immediately from eqn. (A.6).

- The system matrix B can be extracted from H_2 by taking the first m (= number of inputs) columns of this matrix. This follows immediately from eqn. (A.7).

- The system matrix A can be extracted in different ways. With the Ho-Kalman algorithm this matrix is obtained as

  $$A = H_1^{\dagger} \cdot \vec{H}_{n_r,n_c} \cdot H_2^{\dagger} \tag{A.12}$$

  with $H_1^{\dagger} = (H_1^T H_1)^{-1} H_1^T$ being the left pseudo-inverse of H_1, $H_2^{\dagger} = H_2^T (H_2 H_2^T)^{-1}$ being the right pseudo-inverse of H_2 and

  $$\vec{H}_{n_r,n_c} := \begin{pmatrix} G_2 & G_3 & \dots & G_{n_c+1} \\ G_3 & G_4 & \dots & G_{n_c+2} \\ \vdots & \vdots & \ddots & \vdots \\ G_{n_r+1} & G_{n_r+2} & \dots & G_{n_c+n_r} \end{pmatrix} \tag{A.13}$$

  being the so called *shifted* Hankel matrix of Markov parameters. This derivation of A is based on the fact that

  $$\vec{H}_{n_r,n_c} = \Gamma_{o,n_r} A \Gamma_{c,n_c} \tag{A.14}$$

The answer to the minimal realization problem then boils down to the question of how to obtain a full rank decomposition of the Hankel matrix of Markov parameters as described above. The answer to this question is given as a *singular value decomposition (SVD)* of this matrix. If the size of the Hankel matrix of Markov parameters is chosen such that its rank equals the McMillan degree n of the underlying system $G(z)$, an SVD of this matrix delivers a decomposition of the form

$$H_{n_r,n_c} = U_n \Sigma_n V_n^T \tag{A.15}$$

where U_n and V_n are so called *unitary matrices, i.e.* they fulfill the properties $U_n^T U_n = I_n$ and $V_n^T V_n = I_n$, and Σ_n is a diagonal matrix with positive entries $\sigma_1 \geq ... \geq \sigma_n$, also known as *singular values*. Having obtained the matrices U_n, Σ_n and V_n, one can then choose the matrices H_1 and H_2 equal to

$$H_1 = U_n \Sigma_n^{\frac{1}{2}} \tag{A.16}$$

and

$$H_2 = \Sigma_n^{\frac{1}{2}} V_n^T \tag{A.17}$$

The use of an SVD to obtain the decomposition of the Hankel matrix as discussed here was introduced in [121].

Summarizing, the Ho-Kalman algorithm starts by constructing a sufficiently large Hankel matrix of Markov parameters, *i.e.* one that is assumed to have a rank equal to the McMillan degree n of the underlying system $G(z)$. After that, an SVD is applied to this Hankel matrix to deliver this McMillan degree and the matrices H_1 and H_2. Then, the system matrices A, B and C are obtained from these latter two matrices according to the three-step procedure decribed above. Finally, D is obtained as G_0.

Remark A.2.1 *A short-cut to obtain A with the matrices obtained with the SVD is*

$$A = \Sigma_n^{-\frac{1}{2}} U_n^T \vec{H}_{n_r,n_c} V_n \Sigma_n^{-\frac{1}{2}} \tag{A.18}$$

Remark A.2.2 *When $n_r,\ n_c \rightarrow \infty$, the realization that is obtained with the Ho-Kalman algorithm can be shown to converge to one that is "balanced". A balanced realization is one where each state is "as controllable as it is observable". See e.g. [18] for a formal definition of this property.*

A.3 Kung's modification of Ho-Kalman's algorithm to deal with noisy pulse response data

When the Markov parameters are noisy and thereby nonexact representations of the true Markov parameters, the rank of the Hankel matrix of these parameters will generally not be equal to the McMillan degree n of the underlying system. Instead, it will be equal to either the row or column dimension of this matrix; to be more specific, to the smallest of these two dimensions. The consequence is that, when applying Ho-Kalman's algorithm to noisy Markov parameters, the number of nonzero singular values obtained with the SVD of the Hankel matrix of Markov parameters will, generally, keep on growing with growing dimensions n_r and n_c, even when the rank of the Hankel matrix has already passed the value of the true McMillan degree n. This distorts a proper identification of this McMillan degree. The solution to this problem proposed by Kung [49] is simple: the user has to decide for himself, on the basis of the computed singular values, what the McMillan degree for the underlying system $G(z)$ is. With this chosen value for n, lower-dimensional versions are then constructed of the matrices obtained with the SVD. These are then used to construct H_1 and H_2 and, subsequently, the system matrices. The idea behind this rather small adaptation to Ho-Kalman's algorithm is, in fact, based on a basic fact in linear algebra that any matrix can be approximated by a matrix of lower rank using the SVD.

In more detail, the approximate realization algorithm of Kung is as follows:

1. Construct, as before with the Ho-Kalman algorithm, a Hankel matrix of Markov parameters H_{n_r,n_c} of sufficiently large size.

2. Apply an SVD to this matrix

$$H_{n_r,n_c} = U\Sigma V^T \tag{A.19}$$

   and choose, on the basis of values found for the singular values, an appropriate value for the McMillan degree n.

3. Construct the matrices

$$\Sigma_n = [I_n \; 0]\Sigma \begin{bmatrix} I_n \\ 0 \end{bmatrix}, \quad U_n = U \begin{bmatrix} I_n \\ 0 \end{bmatrix}, \quad V_n = V \begin{bmatrix} I_n \\ 0 \end{bmatrix} \tag{A.20}$$

   Note that such a construction amounts to throwing away all singular values smaller than σ_n; these are regarded as insignificant *i.e.* the result of noise.

4. Use these matrices together with the shifted Hankel matrix of Markov parameters to construct the system matrices A, B, C and D in the same way as with the Ho-Kalman algorithm.

In step 2, the singular values are typically evaluated in a plot. Often that value for n is chosen that corresponds to a 'knee' in the shape of such a singular value plot. Another way of choosing the model order is simply via comparing the original pulse response with that of the resulting model. In that case that model order is chosen that corresponds to a desired level of resemblance between these two pulse responses.

The algorithm described here is actually a small variation on Kung's algorithm; the difference is another way of calculating the A-matrix. The practical differences are, however, moderate.

Approximate realization with step response data

If one doesn't have pulse response data available but, instead, (noisy) step response data, say $\{S_t\}_{t=0,1,\ldots}$, one could still use the approximate realization method of Kung to obtain a minimal realization. The way that is proceeded then is by, first, transforming the step response data to pulse response data by computing

$$G_t = S_t - S_{t-1} \tag{A.21}$$

for all available step response coefficients and by, subsequently, applying Kung's method to these estimated pulse response coefficients. The differencing operation (A.21), however, introduces an amplification of high frequency noise present on the step response coefficients, which may significantly reduce the quality of the resulting model. As an alternative, it is possible to use the step response coefficients directly in an approximate realization method that is a slightly modified version of the one by Kung. This modified version is, as mentioned before, due to [104] and follows the same procedure as Kung's method except for the following:

- The modified version uses the step response matrix

$$R_{n_r,n_c} := \begin{pmatrix} S_1 - S_0 & S_2 - S_0 & \dots & S_{n_c} - S_0 \\ S_2 - S_1 & S_3 - S_1 & \dots & S_{n_c+1} - S_1 \\ \vdots & \vdots & \ddots & \vdots \\ S_{n_r} - S_{n_r-1} & S_{n_r+1} - S_{n_r-1} & \dots & S_{n_r+n_c-1} - S_{n_r-1} \end{pmatrix} \tag{A.22}$$

  in place of the Hankel matrix of Markov parameters H_{n_r,n_c}

- It uses a shifted matrix $R^{\uparrow}_{n_r,n_c}$, which is obtained by shifting R_{n_r,n_c} one block row upwards, in place of the shifted Hankel matrix $\vec{H}_{n_r,n_c}$

- The system matrix D is obtained as $D = G_0 = S_0 - S_{-1} = S_0 - 0 = S_0$

The algorithm is based on the fact that the basic properties of H_{n_r,n_c} for the construction of a realization on the basis of Markov parameters are also owned by the matrix R_{n_r,n_c} for the construction of a realization on the basis of step response data. This is due to the fact that these matrices are related according to

$$R_{n_r,n_c} = H_{n_r,n_c} T_{n_c} \tag{A.23}$$

with T_{n_c} being a non-singular matrix

$$T_{n_c} = \begin{pmatrix} I_m & \dots & \dots & I_m \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & I_m \end{pmatrix} \quad (dim = n_c m \times n_c m) \tag{A.24}$$

(Note, however, that one property is not retained by this transformation: R_{n_r,n_c} is *not* a Hankel matrix). The difference between this modified version of Kung's method and the original method of Kung is that the former results in models which better retain the slow dynamics of the system to be identified than the latter, which retain better the fast dynamics of this system.

Remark A.3.1 *It is noted that the method of [104] assumes* **unit** *step responses. When no unit step responses are experimentally obtained, these can be easily transformed to equivalent unit step responses by dividing the outputs by the sizes of the applied steps.*

A.4 A link with subspace methods

The last 15 to 20 years a number of state space *c.q.* realization based identification methods have been developed which have in common that they estimate a realization on the basis of *arbitrary* in- and output data, instead of data in a specific form as pulse or step response data, as is the case with the approximate realization algorithms discussed above. These methods have become known as *subspace* methods (see *e.g.* [109]), the name of which may be attributed to the fact that they estimate somewhere during the process a subspace of some quantity. The aim of this appendix is to present conditions for which a subspace method belonging to a major group of these methods is equivalent to the step response based approximate realization algorithm of [104]. A similar link has been made between this major group of subspace methods and the algorithms of Ho-Kalman and Kung in (chapter 8 of) [122].

For ease of explanation, we will consider only the case of a disturbance-free data generating system. Subspace methods assume this system to be given by the state space description (A.2). Starting point of all subspace methods is an equation that can readily be derived from this state space description and which is given as

$$Y_{0,s,N} = \Gamma_{o,s} X_{0,N-s+1} + H_s U_{0,s,N} \tag{A.25}$$

with

$$U_{0,s,N} := \begin{pmatrix} u(0) & u(1) & \dots & u(N-s+1) \\ u(1) & u(2) & \dots & u(N-s+2) \\ \vdots & \vdots & \ddots & \vdots \\ u(s-1) & u(s) & \dots & u(N) \end{pmatrix} \tag{A.26}$$

and $Y_{0,s,N}$ of the same shape (replace $u(.)$ by $y(.)$ in eqn. (A.26)) and with $\Gamma_{o,s}$ defined by eqn. (A.6) with $n_r = s$ and further with

$$H_s := \begin{pmatrix} D & 0 & \dots & 0 & 0 \\ CB & D & \dots & 0 & 0 \\ \vdots & & & \ddots & \\ CA^{s-2}B & \dots & \dots & CB & D \end{pmatrix} \tag{A.27}$$

and

$$X_{0,N-s+1} = [x(0)\ x(1)\ ...\ x(N-s+1)] \tag{A.28}$$

Here, we adopted largely the notation of (chapter 8 of) [122]. Like the state space description (A.2), eqn. (A.25) establishes a connection between the measured data and the system matrices of the assumed data generating system. Given this equation, subspace methods proceed in different directions. One group of subspace methods, for instance, tries, roughly stated, to estimate the state matrix $X_{0,N-s+1}$ and a shifted version of this matrix $X_{1,...} = [x(1)\ x(2)\ ...]$ and then, using these state matrices and matrices containing in- or output data, estimate the system matrices in a least squares sense [53]. An example of such a method is the so called N4SID (Numerical algorithms for Subspace State Space System IDentification) method of [108]. Another group of subspace methods, one for which a link is established here with the step response data based approximate realization algorithm of [104], tries to estimate the system matrices by first deriving a matrix from the data matrix $Y_{0,s,N}$ that has the same column space as that of the (so called *extended*) observability matrix $\Gamma_{o,s}$ and then derive the system matrices and order from the matrices obtained from an SVD applied to this derived matrix. The way that is proceeded is by post-multiplying $Y_{0,s,N}$ with some (projection) matrix Π and then perform an SVD on the obtained product $Y_{0,s,N}\Pi$. The function of the matrix Π is to separate the first, relevant, term of the sum (A.25), *i.e.* $\Gamma_{o,s}X_{0,N-s+1}$, from the second, irrelevant, term of this sum, *i.e.* $H_sU_{0,s,N}$. The first term is relevant, indeed, because of the presence of $\Gamma_{o,s}$. In mathematical terms, Π should establish that

$$Y_{0,s,N}\Pi = \Gamma_{o,s}X_{0,N-s+1}\Pi + H_sU_{0,s,N}\Pi = \Gamma_{o,s}X_{0,N-s+1}\Pi + 0 = \Gamma_sX_{0,N-s+1}\Pi \tag{A.29}$$

Different projection matrices Π can be found in the literature. See *e.g.* [122] where for in- and output data of arbitrary form an RQ factorization of the matrix $[U_{0,s,N}^T\ Y_{0,s,N}^T]^T$ is proposed to deliver Π. The system matrices of the realization are then typically obtained (i) all directly from the matrices obtained with the SVD or (ii) partly from these matrices and partly from solving a least squares problem. Subspace methods that belong to this second category are *e.g.* the MOESP (Multivariable Output Error State sPace)-algorithm of [112] and the ORSE (Observability Range Space Extraction)-algorithm of [62].

Returning now to the link to be made, it can be shown that a subspace method belonging to the second category discussed above reduces to the approximate realization algorithm of [104] if this subspace method fulfills certain conditions. In particular, this subspace method needs to be applied in a particular SIMO way, *i.e.* the input data used for estimation should be given in the following form:

$$u(t) = \begin{cases} 1 & for\ t \geq 0 \\ 0 & for\ t < 0 \end{cases} \tag{A.30}$$

Secondly, the matrix Π should be chosen equal to

$$\Pi = \begin{pmatrix} -1 & -1 & \dots & -1 & -1 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & & \vdots & \vdots \\ \vdots & 0 & \ddots & 0 & \vdots \\ \vdots & \vdots & & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \tag{A.31}$$

(dimension: $(N-s+2)\times(N-s+1)$). A first consequence of choosing the projection matrix Π according to (A.31) is that

$$H_sU_{0,s,N}\Pi = H_s0 = 0 \tag{A.32}$$

(Note that $u(t_1) - u(t_2) = 1 - 1 = 0\ \forall\ t_1,\ t_2$) and that, hence, this projection matrix establishes the desired situation depicted in eqn. (A.29). Moreover, the following relation holds:

$$
\begin{aligned}
Y_{0,s,N}\Pi &= \begin{pmatrix} y(1)-y(0) & y(2)-y(0) & \dots & y(N-s+1)-y(0) \\ y(2)-y(1) & y(3)-y(1) & \dots & y(N-s+2)-y(1) \\ \vdots & \vdots & \ddots & \vdots \\ y(s)-y(s-1) & y(s+1)-y(s-1) & \dots & y(N)-y(s-1) \end{pmatrix} \\
&= R_{s,N-s+1} \qquad \text{(A.33)}
\end{aligned}
$$

with $R_{s,N-s+1}$ being equal to R_{n_r,n_c} for $n_r = s$ and $n_c = N-s+1$. Hence,

$$
Y_{0,s,N}\Pi = \Gamma_{o,s}X_{0,N-s+1}\Pi + 0 = \Gamma_{o,s}X_{0,N-s+1}\Pi = R_{s,N-s+1} \qquad \text{(A.34)}
$$

which shows that, for the particular choice for the inputs (eqn. (A.30)) and projection matrix Π (eqn. (A.31)), a subspace method of the second category discussed above employs an SVD to exactly the same matrix $R_{s,N-s+1}$ as the step response based approximate realization algorithm of [104] does. From this latter observation it follows immediately that such a subspace method reduces to the approximate realization algorithm of [104] if, in addition to the two conditions *c.q.* choices already mentioned above, the system matrices are obtained from an SVD applied to $Y_{0,s,N}\Pi$ in the same way as is done with the algorithm of [104]. It is noted that the latter condition is typically not fulfilled: in [122] the system matrices are *e.g.* obtained in a different way from the SVD results than is done with the algorithm of [104].

Remark A.4.1 *In [122] a subspace algorithm is described that also estimates a realization on the basis of step response data. This algorithm differs from the approximate realization algorithm of [104] in the choice for the matrix* Π *and the way the system matrices are obtained from the results of the SVD.*

Appendix B

Constrained multiple data set estimation of linear regression models

B.1 Introduction

The prediction error method employed by the MSWC plant system identification procedure proposed in this thesis allows for estimating linear regression models, *i.e.* models which are linear in the parameters to be estimated such as ARX or FIR models, on the basis of multiple data sets and with the static gains enforced. For completeness, the techniques for estimating such models, which are fairly standard, are provided here.

B.2 Estimation of linear regression models

When the model structure is chosen to be one of the linear regression types, the system to be modeled is assumed to be of the form

$$y(t) \quad = \quad \varphi^T(t)\theta_o + e(t) \tag{B.1}$$

with θ_o the system parameter (column) vector and $\varphi(t)$ the so called *regression vector* or, in the multivariable case, regression *matrix*. Note that the outputs of the system to be modeled are assumed to be a linear function of the parameters in θ_o. The regression vector or matrix contains present and past values of the inputs and/or (depending on the specific model structure) outputs. The optimal one-step-ahead predictor for the outputs of a system of the form (B.1) is given as

$$\hat{y}(t,\theta) \quad = \quad \varphi^T(t)\theta \tag{B.2}$$

which results in the following prediction error:

$$\varepsilon(t,\theta) \quad = \quad y(t) - \varphi^T(t)\theta \tag{B.3}$$

The estimation of the parameter vector θ, *i.e.* through minimizing the LS criterion (3.7), can be done very efficiently in two ways. The first, most common way is to use the (straightforwardly) analytically derived expression for its solution

$$\hat{\theta}_N = [\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\Lambda^{-1}\varphi^T(t)]^{-1}[\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\Lambda^{-1}y(t)] \tag{B.4}$$

A second way is obtain the parameter vector estimate through solving a *quadratic program* (QP) (see also appendix D), *i.e.* by using a numerical solver for the optimization problem

$$\hat{\theta}_N = \arg\min_{\theta} 0.5 \cdot \theta^T H\theta + g^T\theta \tag{B.5}$$

with

$$H = \frac{1}{N}\sum_{t=1}^{N}\varphi(t)\Lambda^{-1}\varphi^T(t) \tag{B.6}$$

(which, in optimization nomenclature, is called the *Hessian*) and

$$g = -\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\Lambda^{-1}y(t) \tag{B.7}$$

(the *gradient*). Note that the first-order necessary conditions for (B.5) lead to the solution

$$\hat{\theta}_N = -H^{-1}g \tag{B.8}$$

which is equal to the analytical solution (B.4).

Very efficient and, thereby, fast numerical solvers exist presently for QP problems, thereby rendering it a useful alternative to the more common way of computing the parameter vector through (B.4). Here, the QP way of computing the parameter vector is considered as it allows for a straightforward inclusion of static gains of the system to be modeled as constraints in the parameter estimation problem, as explained now below.

B.3 Enforcement of static gains on linear regression models

One way to enforce the static gains of the system to be modeled on the linear regression model to be estimated is by (i) transforming the corresponding conditions on the (steady-state form of the) model equations into a set of equalities that are linear in the parameter vector θ

$$A\theta = b \tag{B.9}$$

with A and b a matrix resp. vector of appropriate dimensions consisting of the steady state gain values, and (ii) adding this set to the QP (B.5) to form the equality constrained

QP

$$\begin{aligned} \hat{\theta}_N &= \arg\min_{\theta} 0.5 \cdot \theta^T H \theta + g^T \theta \\ s.t. &\quad A\theta = b \end{aligned} \tag{B.10}$$

with the Hessian H and gradient g as defined before by eqns. (B.6) resp. (B.7), to subsequently (iii) solve this QP for θ.

For an ARX model structure, with G_{ss} denoting the $p \times m$ matrix containing the static gains to be enforced, the A matrix and b vector are derived by rewriting the condition (see also eqns. (3.8) - (3.9))

$$\begin{aligned} G_{ss} &= G(q,\theta)|_{q=1} = A(1,\theta)^{-1}B(1,\theta) \\ &= [I_p + A_1 + ... + A_{na}]^{-1}[B_0 + B_1 + ... + B_{nb}] \end{aligned} \tag{B.11}$$

as

$$col[A(1,\theta)G_{ss}] = col[B(1,\theta)] \tag{B.12}$$

and reordering the resulting $p \times m$ equations such that the set of equalities (B.9) is obtained. Here, $col[.]$ represents the *column operator* which stacks the columns of a matrix on top of each other.

B.4 Multiple data set estimation of linear regression models with enforced static gains

The key to proper estimation of a PEM model on the basis of multiple data sets is to avoid that the output predictors (3.4), or (B.2) in case of linear regression models, are a function of more than one data set. This happens *e.g.* when one simply stacks the separately obtained data sets on top of each other and provides the resulting data set to the estimation algorithm if it were one and the same data set. Such a simple multiple data set estimation strategy results in errors in the model to be estimated, even though these may be small when the lengths of the data sets are large compared to their number.

In case of employing linear regression models, the following proposition provides a method for properly estimating a model on the basis of multiple data sets while also taking into account constraints of the type of (B.9).

Proposition B.4.1 *Consider the PEM setting with linear regression models and least squares criterion as described in section 3.2.3. Also, assume that linear equality constraints of the form of eqn. (B.9) are imposed on the model to be estimated. Then, with q the number of data sets used for estimation, a proper multiple data set identification method, i.e. one that computes each of the output predictions on the basis of the correct single data set, is given by QP (B.10) with*

$$H = H_1 + ... + H_q \tag{B.13}$$

and

$$g = g_1 + ... + g_q \tag{B.14}$$

where H_j and g_j, $j = 1 \dots q$, are the Hessian resp. gradient computed on the basis of data set j only according to eqns. (B.6) resp. (B.7).

Proof B.4.2 *A correct way of multiple data set identification within the considered setting is by minimizing a criterion*

$$V(\theta) = V_{N_1}(\theta) + ... + V_{N_q}(\theta) \tag{B.15}$$

where

$$V_{N_j}(\theta) = \frac{1}{N_j} \sum_{t=1}^{N_j} \varepsilon_j^T(t,\theta)\Lambda^{-1}\varepsilon_j(t,\theta) \qquad j = 1...q \tag{B.16}$$

and

$$\varepsilon_j(t,\theta) = y_j(t) - \hat{y}_j(t,\theta) = y_j(t) - \varphi_j^T(t)\theta \qquad t = 1...N_j \tag{B.17}$$

Note that the output predictors for data set j are constructed from values from this data set only. Proposition B.4.1 follows now from the latter fact and from the fact that the criterion defined by eqns. (B.15) - (B.17) is equivalent to that defined in proposition B.4.1 by eqns. (B.10), (B.13) and (B.14).

Multiple data set estimation is also discussed in [103] though for the unconstrained case and without the explicit statement that the key issue in multiple data set estimation is the avoidance of output predictors being a function of more than one data set. Also, consequences of the multiple data set estimation approach are discussed for model bias and variance and for the identification experiment(s).

Remark B.4.3 *For the unconstrained case, it can be easily shown (see e.g. [103]) that the parameter vector $\hat{\theta}$ obtained with the proposed multiple data set identification method is equal to a weighted sum of the parameter vectors that would be obtained from each data set individually. This sum is given by*

$$\hat{\theta} = [H_1 + ... + H_q]^{-1}[H_1\hat{\theta}_1 + ... + H_q\hat{\theta}_q] \tag{B.18}$$

with $\hat{\theta}_j$, $j = 1 \dots q$, the parameter vector that would be estimated on the basis of data set j only. Hence, the weighting 'factors' here are the Hessians corresponding to each of the data sets individually.

Appendix C

A simple MIMO ARX model order selection strategy

In this appendix a simple order selection strategy is presented for MIMO ARX models of which experience with identification of several large scale MSWC plants has proven it to be a suitable one for, at least, this type of plants. This order selection strategy concerns the choice of the structure indices na and nb of the matrix polynomials $A(q, \theta)$ and $B(q, \theta)$, as given by eqns. (3.9), of the ARX model structure (3.8). In fact, for MIMO ARX model structures these structure indices are matrices, and not scalars, with each entry of these matrices representing the structure index for one of the SISO transfers. This leaves a high number of structure indices to be chosen which makes the order selection rather cumbersome. Therefore, in order to considerably simplify the order selection for MIMO ARX models, the order selection strategy presented here assumes only ARX model structures of which the matrix polynomials are fully parameterized, *i.e.* of which all matrices A_i, $i = 1...na$, and B_j, $j = 0...nb$ contain a parameter to be estimated at each of their entries. As a result of this choice, only one scalar structure index na and one scalar structure index nb need to be chosen. Even with only these two indices to be chosen the model order selection procedure may be cumbersome. Therefore, in order to facilitate the order selection even further, the selection procedure proposed here chooses these values equal to each other, thereby reducing this procedure to the choice of only one index $n = na = nb$. It is noted that order selection for MIMO ARX model structures also requires values to be chosen for the entries of a matrix nk which represent the delays, expressed in number of samples, between the in- and outputs of the plant to be modeled. The problem of the (also) cumbersome choice of these values is avoided by a priori choosing these values equal to either 0 or 1, depending on whether one expects a direct throughput from the corresponding in- to outputs or not. The choice for the latter values generally does not pose any problem. With the entries of nk chosen to some fixed values in this way, the MIMO ARX order selection procedure proposed here is the following one:

1. Compute for each model order n in some a priori chosen range (say 5 to 60) and

for each output $i = 1, ..., p$ the value of the criterion

$$V_{OE}(i,n) = \frac{1}{N}\sum_{t=1}^{N}(\hat{y}_{i,n}(t) - y_i(t))^2 \tag{C.1}$$

with OE standing for *output error*. Here, $y_i(t)$ is the real c.q. measured output i at time instant t and $\hat{y}_{i,n}(t)$ is the *simulated* c.q. noise free output i at time t (for model order n) (and, hence, not the *predicted* output i at this time instant).

2. Compute for each output i the lowest value for $V_{OE}(i,n)$ as a function of n, *i.e.* compute for each output i the value

$$V_{OE}^{min}(i) = \min_{n} V_{OE}(i,n) \tag{C.2}$$

3. Divide for each output i all the values $V_{OE}(i,n)$ corresponding to this output by the minimum value $V_{OE}^{min}(i)$ found for this output, *i.e.* compute for each output i for all n the value

$$V_{OE}^{scaled}(i,n) = \frac{V_{OE}(i,n)}{V_{OE}^{min}(i)} \tag{C.3}$$

Note that, for each output i, the minimum value for $V_{OE}^{scaled}(i,n)$ taken over n will be 1 whereas the rest will be larger than 1.

4. Average for each order n the values found for $V_{OE}^{scaled}(i,n)$ over all outputs, *i.e.* compute

$$\overline{V}_{OE}^{scaled}(n) = \frac{1}{p}\sum_{i=1}^{p} V_{OE}^{scaled}(i,n) \tag{C.4}$$

with p representing the number of outputs. Typically, the value of $\overline{V}_{OE}^{scaled}(n)$ is either equal to or (a little) larger than 1. Equality is only obtained when all outputs correspond to the same optimal model order.

5. Finally, determine the minimum value, over n, of the values found for $\overline{V}_{OE}^{scaled}(n)$, *i.e.* compute

$$n_{opt} = \arg\min_{n} \overline{V}_{OE}^{scaled}(n) \tag{C.5}$$

It is recommended to use a data set for computing the criteria here that has not been used for estimation as this avoids a possible monotone decrease of the criterion $\overline{V}_{OE}^{scaled}(n)$ as a function of n and, thereby, an impossible choice for the latter order. The scaling in the order selection procedure above (step 2 and 3) is applied to take care of the differences in dimension between the outputs, which else could result in an unfair contribution of one output over the other in the final order selection step. Averaging over the outputs, step 4, is applied to arrive at one unique criterion, thereby preventing the problem of having to choose between different orders which are optimal for different outputs. It must be noted here, however, that experience also has shown that the choice for the model order has not turned out to be that very critical in the sense that the differences between the finally obtained models (after model reduction: see chapters 3 and

4) were small if the model orders were not chosen too far away from the optimal one obtained with the order selection procedure presented here.

An example of the resulting criterion values $\overline{V}_{OE}^{scaled}(n)$ as a function of n is given in figure C.1 which is a result from an application of the identification procedure proposed in chapter 3 to the large scale MSWC plant AZN situated at Moerdijk, The Netherlands. The values for n_{opt} encountered in practice were within the range of

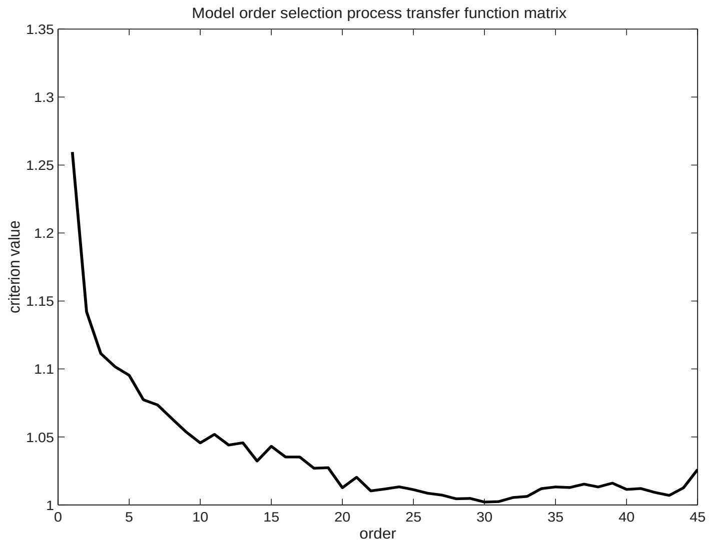

Figure C.1: Model order selection for high order MIMO ARX models: $\overline{V}_{OE}^{scaled}(n)$ *as a function of* n *for deriving a model order for the Dutch MSWC plant AZN. Note that* $n_{opt} = 30$.

$n_{opt} = 10$ to 50, usually around 20 to 30, which are values that are typically found in the literature.

Many researchers have investigated the model order selection problem for system identification. See *e.g.* [63] and [122].

168

Appendix D

Sequential Quadratic Programming

D.1 Introduction

In this appendix an introduction is given to a specific method, known as *sequential* or *succesive quadratic programming (SQP)*, that is able to solve generic nonlinear optimization problems of the form

$$\begin{aligned} \min_x \;\; & f(x) \\ s.t. \;\; & g(x) \leq 0 \\ & h(x) = 0 \end{aligned} \tag{D.1}$$

This type of optimization problem is referred to as a *nonlinear programming (NLP)* problem. The SQP method solves this NLP, as the name suggests, by sequentially / successively (iteratively) solving a so called *quadratic program (QP)*, which is an optimization problem of the form

$$\begin{aligned} \min_x \;\; & \tfrac{1}{2} x^T H x \; + \; f^T x \\ s.t. \;\; & A_{ineq} x \leq b_{ineq} \\ & A_{eq} x = b_{eq} \end{aligned} \tag{D.2}$$

A QP is itself an NLP due to the presence of a quadratic objective function, but for this type of optimization problem efficient solvers are widely available.

Apart from providing an introduction to SQP, a specific such method is presented that has been found to work well on (at least) relatively small sized NLPs, including ones resulting from small scale nonlinear model predictive control (NMPC) applications (such as discussed in this thesis). This SQP algorithm is an extension towards inequality constraints of an SQP method given in [8] that deals only with NLPs with equality constraints. The properties of this SQP method are discussed both by investigating its local and global convergence properties (the definition of which is given later on in this appendix) and by means of an application to an alkylation plant steady-state optimization problem obtained from the literature. It must be noted that, because the resulting SQP algorithm seems to be quite commonly used, the results presented here with respect to this algorithm (in particular with respect to its local and global convergence properties) may also be found elsewhere in the literature. Nevertheless, so far, the author of this thesis has not encountered them yet.

SQP is a popular technique for solving NLPs and, additionally, one of the most successful ones [71]. The earliest reference to SQP-type algorithms seems to have been in the 1963 PhD thesis of Wilson [115]. Real impetus was given to SQP methods, however, by the papers of Han and Powell, in the late seventies of the twentieth century [33, 34, 80, 81, 82]. SQP is presently the number one choice for solving dynamic optimization problems (via the *direct approach*: see chapter 6), in particular for solving the optimal control problems underlying NMPC problems.

The outline of this appendix is as follows. First, in section D.2 some preliminaries are given including (i) a concise statement of the NLP, (ii) the necessary and sufficient conditions that apply to each local minimum of such an NLP, (iii) a discussion of Newton's method and (iv) a statement of the assumptions on the QP solver used as part of the SQP methods considered here. Then, in section D.3, the basic idea and main issues regarding SQP methods are discussed, such as (i) formulation of the QP subproblem, (ii) approximation of the Hessian, (iii) establishing global convergence and (iv) measures to provide robustness to the SQP algorithm. The section ends with an outline of the specific SQP algorithm referred to above. After that, in section D.4, the local and global convergence properties of this SQP algorithm are discussed. Finally, in section D.5, the application of the proposed SQP algorithm to a moderately sized NLP is discussed, *i.e.* the steady state optimization of an alkylation plant referred to above.

A final note here is that the results presented in this appendix rely to a large extent on results presented in [8], which provides an excellent and extensive introduction into SQP.

D.2 Preliminaries

D.2.1 The nonlinear programming problem

The NLP considered in this appendix is of the form (D.1) with the functions $f : R^n \rightarrow R$, $g : R^n \rightarrow R^m$ and $h : R^n \rightarrow R^p$ all assumed to be three times continuously differentiable. It may be sufficient for the problem functions to be less times continuously differentiable, *e.g.* two, but because the results presented here rely heavily on those given in [8], it has been chosen here to assume these functions to be as many times continuously differentiable as assumed in this reference.

Note that no limiting assumptions are made on the linearity of the NLP problem functions $f(.)$, $g(.)$ and $h(.)$. Hence, these are all allowed to be nonlinear. Actually, an NLP is particularly characterized by at least one of the its problem functions being nonlinear.

D.2.2 Necessary and sufficient conditions for a local minimizer

Much of the discussion in this appendix involves the first-order necessary conditions for a local minimizer x^* of the NLP (D.1). These are, therefore, explicitly stated here. In fact, for reference and for completeness, the so called *strong second order sufficient conditions* for such a local minimizer are given here, as stated in [8]. These include not only the first-order necessary conditions but also, as the name suggests, second order *sufficient* conditions. Before giving these conditions, however, first the so called *Lagrangian (function)* needs to be defined for the NLP (D.1):

$$\mathcal{L}(x,u,v) = f(x) + u^T g(x) + v^T h(x) \tag{D.3}$$

Here, (the column vectors) u and v represent the so called *Lagrange multipliers*. Then, the strong second order sufficient conditions for any local minimum x^* of the NLP (D.1) are given as [8]:

Assumption D.2.1 *(***Strong second order sufficient conditions for a local minimum of NLP (D.1))**

1. *There exist optimal multipliers v^* and $u^* \geq 0$ such that*

$$\nabla_x \mathcal{L}(x^*, u^*, v^*) \quad = \quad \nabla_x f(x^*) + \nabla_x g(x^*) u^* + \nabla_x h(x^*) v^* = 0 \tag{D.4}$$

   (Here, ∇_x... means the derivative / gradient / Jacobian towards x).

2. *With $\mathcal{I}(x) = \{i: \ g_i(x)\}$ denoting the set of indexes of inequality constraints that are active (also called binding) at x, the columns of the matrix $G(x^*)$ made up of the matrix $\nabla_x h(x^*)$ along with the columns $\nabla_x g_i(x^*)$, $i \in \mathcal{I}(x^*)$ are linearly independent.*

   (This is also called a regularity condition, *i.e.* a condition that ensures that x^* is a so called regular point).

3. *Strict complementary slackness holds, i.e. $g_i(x^*)u_i^* = 0$ for $i = 1...p$ and, if $g_i(x^*) = 0$, then $u_i^* > 0$.*

4. *The Hessian of the Lagrangian function with respect to x is positive definite on the null space of $G(x^*)^T$, i.e. $\Delta x^T \nabla^2_{xx}[\mathcal{L}(x^*, u^*, v^*)]\Delta x > 0$ for all $\Delta x \neq 0$ such that $G(x^*)^T \Delta x = 0$.*

Of particular importance in this appendix are the so called *stationary point*s (also sometimes referred to as *critical point*s: see *e.g.* [8]), which are points x that satisfy both the first-order necessary conditions such as (D.4) and, if present in the NLP, the constraints.

D.2.3 Newton's method

Of major importance when discussing SQP methods is Newton's method, which is a method to find the zeros of a generic set of nonlinear equations;

$$f(x) = 0 \tag{D.5}$$

Newton's method computes the solution x^* to (D.5) iteratively by solving for each iterate x_k the linear set of equations

$$[\nabla_x f(x_k)]^T \Delta x_k = -f(x_k) \tag{D.6}$$

to obtain a new search direction Δx_k and subsequently computing a new estimate for the solution x^* as

$$x_{k+1} = x_k + \Delta x_k \tag{D.7}$$

This iteration is repeated until some termination criterion is fulfilled. The motivation for using these equations to find the solution to (D.5) reads as follows. The aim is to find, for a current estimate x_k, a change Δx_k that ensures that

$$f(x_k + \Delta x_k) = f(x^*) = 0 \tag{D.8}$$

Approximating the first term on the left here with a first-order Taylor series expansion then leads to

$$f(x_k + \Delta x_k) \approx f(x_k) + [\nabla_x f(x_k)]^T \Delta x_k = 0 \tag{D.9}$$

which, when rewritten, leads to the set of equations (D.6) and, when this set is subsequently solved for Δx_k, to a new estimate x_{k+1}, as given by eqn. (D.7).

Newton's method exhibits good *local convergence* properties which refers to its behavior when the solution estimate x_k is close to a solution x^*. More specifc, Newton's method is known to converge very rapidly to a local solution when sufficiently close to this solution: it then typically converges at a *quadratic* convergence rate, *i.e.* according to

$$\|x_{k+1} - x^*\| \leq \nu \|x_k - x^*\|^2 \tag{D.10}$$

for some positive constant ν. On the other hand, Newton's method is also known to have poor *global convergence* properties, meaning its behavior when the solution estimate is far from a possible solution. More specific, it may then *not* converge to such a solution. In order to enforce convergence to a point that is sufficiently close to some local optimum, a set of measures denoted as a *globalization strategy* is used.

Newton's method is often applied with an approximation to $\nabla_x f(x_k)$, in particular with the aim to reduce the computation time. In that case, the local convergence rate typically reduces to a *superlinear* or even *linear* convergence rate, with the first defined as

$$\|x_{k+1} - x^*\| \leq \nu_k \|x_k - x^*\| \tag{D.11}$$

for some sequence of positive constants $\nu_k \to 0$ and the latter defined in the same way except that ν_k is now replaced by a positive constant ν.

The convergence rate of Newton's method is often used in local convergence proofs of other optimization methods by exploiting the equivalence between the considered optimization method and Newton's method when close to a solution. The quadratic convergence rate is then considered the maximum achievable convergence rate, which the considered optimization method then ideally also exhibits. The equivalence with Newton's method has also been used to establish local convergence results for SQP methods, including the one presented in this appendix.

D.2.4 Assumptions on the QP solver

To facilitate both the implementation of the SQP algorithm proposed in this appendix and the derivation of theoretical results for this SQP algorithm, a number of assumptions are made on the solver used to obtain a solution for its QP subproblem. These assumptions are typically fulfilled for the commercial QP solvers that are currently available (such as *e.g.* quadprog.m in the optimization toolbox of MATLAB [100] and (d)qprog.f in FORTRAN). First of all, it is assumed that the QP solver efficiently and reliably delivers the global solution to the QP subproblem. It should also deliver the Lagrange multipliers corresponding to this solution, including the zero ones that correspond to the inequality constraints that are inactive at this solution. It is also assumed that the solution of the QP subproblem fulfills the strong second order sufficient conditions given above under assumption D.2.1. This implies, amongst others, that the solution to the QP subproblem fulfills the strict complementarity slackness condition, which is used in the derivation of the global converge properties of the proposed SQP algorithm. Finally, it is assumed that the QP solver stops and delivers a notifying signal (a *flag*) when infeasibility of the QP subproblem has been detected.

The current available QP solvers can be divided in *active set* and *interior point* QP solvers. It is outside of the scope of this thesis to provide a detailed discussion of both these two solution approaches (for an introduction one is referred to *e.g.* [71]). It suffices here to say that interior point QP solvers are more efficient than active set QP solvers when the number of inequality constraints is large. This is a consequence of the fact that active set QP solvers try to find the solution by searching for the set of inequality constraints that are active at the solution and then solve the set of equalities that correspond to this solution. This is a combinatorial problem requiring the consideration of many active set possibilities. In contrast, rather than looking for the solution over the border of the feasible region, interior point methods look for the solution from the inside of this region.

D.3 SQP: an introduction and an algorithm

D.3.1 The main idea

Many optimization problems are solved by applying some form of Newton's method to the set of equations consisting of the first-order necessary conditions and the constraints that correspond to the considered optimization problem. For example, for unconstrained problems

$$\min_x \; f(x) \tag{D.12}$$

with continously differentiable function $f : R^n \rightarrow R$, the first-order necessary conditions for a(ny) local minimizer x^* are given as

$$\nabla_x f(x^*) = 0 \tag{D.13}$$

Applying Newton's method to this set of equations leads to an iteration (D.7) where the search direction is computed according to

$$[\nabla^2_{xx} f(x_k)]\Delta x_k = -\nabla_x f(x_k) \tag{D.14}$$

Because Newton's method is used, applying this iteration results in a quadratically (locally) convergent algorithm or, when using an approximation to the Hessian $\nabla^2_{xx} f(x_k)$, in an algorithm with a slower convergence rate such as a linear or superlinear superlinear convergence rate.

Another example is the equality constrained NLP

$$\begin{aligned} \min_x \; & f(x) \\ s.t. \quad & h(x) = 0 \end{aligned} \tag{D.15}$$

For this case, the first-order necessary conditions for a(ny) local solution x^* are given as

$$\nabla_x \mathcal{L}(x^*, v^*) \quad = \quad \nabla_x f(x^*) + \nabla_x h(x^*) v^* = 0 \tag{D.16}$$

while the constraint equations that need to be satisfied at this solution are given as

$$h(x^*) = 0 \tag{D.17}$$

Applying Newton's method to the set of equations consisting of (D.16) and (D.17) leads to an iteration

$$\begin{bmatrix} x_{k+1} \\ v_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ v_k \end{bmatrix} + \begin{bmatrix} \Delta x_k \\ \Delta v_k \end{bmatrix} \tag{D.18}$$

with the changes Δx_k and Δv_k computed by solving the following set of linear equations:

$$\begin{bmatrix} \nabla^2_{xx} \mathcal{L}(x_k, v_k) & \nabla_x h(x_k) \\ \nabla_x h(x_k)^T & 0 \end{bmatrix} \begin{bmatrix} \Delta x_k \\ \Delta v_k \end{bmatrix} = \begin{bmatrix} -\nabla_x \mathcal{L}(x_k, v_k) \\ -h(x_k) \end{bmatrix} \tag{D.19}$$

A difficulty arises when the optimization problem contains inequality constraints. In that case, the conventional form of Newton's method discussed here so far does not work anymore due to the presence of such constraints. Consider, for instance, the NLP (D.1):

$$\begin{aligned} \min_x \; & f(x) \\ s.t. \quad & g(x) \leq 0 \\ & h(x) = 0 \end{aligned} \tag{D.20}$$

The first-order necessary conditions and constraints for this optimization problem consist of the following equations:

$$\begin{aligned} \nabla_x \mathcal{L}(x^*, u^*, v^*) \quad &= \quad 0 \\ g(x^*) \quad &\leq \quad 0 \\ h(x^*) \quad &= \quad 0 \end{aligned} \tag{D.21}$$

Newton's method does not provide a solution to this set of equations unless one is willing to make an *ad hoc* adaptation to this method. More specific, one can simply treat the inequality constraints in the same way as the equalities are treated by Newton's method ($g(x_k+\Delta x_k) \leq 0$ + first-order Taylor series expansion) to arrive at the iteration

$$\begin{bmatrix} x_{k+1} \\ u_{k+1} \\ v_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ u_k \\ v_k \end{bmatrix} + \begin{bmatrix} \Delta x_k \\ \Delta u_k \\ \Delta v_k \end{bmatrix} \tag{D.22}$$

with the search directions Δx_k, Δu_k and Δv_k obtained by solving the set of equations

$$\begin{bmatrix} \nabla^2_{xx}\mathcal{L}(x_k,u_k,v_k) & \nabla_x g(x_k) & \nabla_x h(x_k) \\ \nabla_x g(x_k)^T & 0 & 0 \\ \nabla_x h(x_k)^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x_k \\ \Delta u_k \\ \Delta v_k \end{bmatrix} \begin{matrix} = \\ \leq \\ = \end{matrix} \begin{bmatrix} -\nabla_x \mathcal{L}(x_k,u_k,v_k) \\ -g(x_k) \\ -h(x_k) \end{bmatrix} \tag{D.23}$$

One is left now, however, with the problem of how to solve this set of equations. The solution for that chosen by SQP methods is (indeed) to use a QP. More specific, for the set of equations (D.23), the QP

$$\begin{aligned} \min_{\Delta x_k} \quad & \frac{1}{2}\Delta x_k^T[\nabla^2_{xx}\mathcal{L}(x_k,u_k,v_k)]\Delta x_k + [\nabla_x \mathcal{L}(x_k,u_k,v_k)]^T\Delta x_k \\ s.t. \quad & [\nabla_x g(x_k)]^T \Delta x_k \leq -g(x_k) \\ & [\nabla_x h(x_k)]^T \Delta x_k = -h(x_k) \end{aligned} \tag{D.24}$$

is solved for Δx_k, Δu_k and Δv_k with the latter two chosen as the optimal QP Lagrange multipliers corresponding to the inequality resp. equality constraints. One can easily show that the first-order necessay conditions of this QP combined with the its constraints form the set of equations given by (D.23) and that, hence, their solutions are the same.

The latter idea of solving Newton's method by means of a QP when this method is applied to the first-order necessary conditions and constraints of the NLP to be solved is the main idea behind SQP methods. It is noted that also the unconstrained and equality constrained problems discussed above in this section allow for QP solutions (This can be easily shown but has been left out here for the purpose of space. The equality constrained case is discussed in [71]). However, for these two NLP types there is no real need for that as the corresponding equations that follow from applying Newton's method to the first-order necessary conditions and, if present, constraints (eqns. (D.14) and (D.19)) can also be solved via the usual linear algebra techniques for solving linear sets of equations.

Even though solving the QP (D.24) is an excellent way of computing the search directions for Newton's method when applied to compute a stationary point of NLP (D.1), SQP methods often employ a different, though related, form of QP. In the next two sections, these other QP forms are discussed and motivated. To limit the discussion, only NLPs of the (most generic) form (D.1) are considered. Extensions to unconstrained or equality constrained versions of this NLP are straightforward.

D.3.2 Reformulations of the QP subproblem

A first alternative QP subproblem formulation that one may encounter with SQP methods (to that of (D.24)) is one that follows from employing the seemingly logical approach of approximating the objective function of the NLP by a second-order Taylor series expansion and the constraints by a first-order Taylor series expansion. When applying this approach to NLP (D.1), the QP subproblem becomes, at each iteration k for the current estimate of the solution x_k, equal to

$$\begin{aligned} \min_{\Delta x_k} \quad & \frac{1}{2}\Delta x_k^T[\nabla_{xx}^2 f(x_k)]\Delta x_k + [\nabla_x f(x_k)]^T \Delta x_k \\ s.t. \quad & [\nabla_x g(x_k)]^T \Delta x_k \leq -g(x_k) \\ & [\nabla_x h(x_k)]^T \Delta x_k = -h(x_k) \end{aligned} \tag{D.25}$$

This QP form is a reasonable choice when only linear constraints are present in the NLP but does not always properly deal with nonlinear constraints, as shown by an example in [8]. The main problem with this QP formulation is that it contains insufficient information about the (nonlinear) constraints of the NLP. In contrast, the QP (D.24) *does* contain sufficient such information, *i.e.* via the Lagrangian information contained in the objective function of this QP.

Another alternative QP form encountered in the literature, with no apparently clear justification, is

$$\begin{aligned} \min_{\Delta x_k} \quad & \frac{1}{2}\Delta x_k^T[\nabla_{xx}^2 \mathcal{L}(x_k, u_k, v_k)]\Delta x_k + [\nabla_x f(x_k)]^T \Delta x_k \\ s.t. \quad & [\nabla_x g(x_k)]^T \Delta x_k \leq -g(x_k) \\ & [\nabla_x h(x_k)]^T \Delta x_k = -h(x_k) \end{aligned} \tag{D.26}$$

i.e. with the the gradient term in the QP objective function now given by the gradient of the NLP objective function instead of by the gradient of the NLP Lagrangian: information about the Lagrangian is only contained in the QP Hessian. This QP form can be shown to deliver the same solution for Δx_k as QP (D.24) if the QP solver delivers zero values for the Lagrange multipliers that correspond to inactive linear inequality constraints (which actually is assumed in section D.2.4) (see also [8]). The fact that both QPs deliver the same solution Δx_k can be easily shown for the case that only equality constraints are present (while this is more difficult when inequality constraints are present). In that case the equality constraints of the QP subproblem guarantee that

$$[\nabla_x h(x_k)]^T \Delta x_k = -h(x_k) = constant\ (vector) \tag{D.27}$$

i.e. that he term $[\nabla_x h(x_k)]^T \Delta x_k$ effectively becomes independent from Δx_k. This implies that the gradient part of the objective function of QP subproblem (D.24) (with no inequality constraints present) effectively is equal to

$$\begin{aligned} [\nabla_x \mathcal{L}(x_k, v_k)]^T \Delta x_k &= [\nabla_x f(x_k)]^T \Delta x_k + v_k^T [\nabla_x h(x_k)]^T \Delta x_k \\ &= [\nabla_x f(x_k)]^T \Delta x_k - v_k^T h(x_k) \\ &= [\nabla_x f(x_k)]^T \Delta x_k - \{term\ independent\ from\ \Delta x_k\} \end{aligned} \tag{D.28}$$

and, thus, can be reduced to $[\nabla_x f(x_k)]^T \Delta x_k$ without changing the outcome of the QP for Δx_k.

Equivalence of the two QPs (D.24) and (D.26) does not hold with respect to the solution for the Lagrange multipliers: the interpretation of the Lagrange multipliers u_k^{qp} and v_k^{qp} corresponding to QP (D.26) differs from the interpretation associated with the Lagrange multipliers corresponding to QP (D.24), *i.e.* Δu_k resp. Δv_k. The connection and difference between these two Lagrange multiplier interpretations can be easily shown for the equality constrained case by rewriting the corresponding first-order necessary conditions of QP (D.24)

$$\nabla_{xx}^2 \mathcal{L}(x_k, v_k)\Delta x_k + \nabla_x \mathcal{L}(x_k, v_k) + \nabla_x h(x_k)\Delta v_k = 0 \tag{D.29}$$

as

$$\nabla_{xx}^2 \mathcal{L}(x_k, v_k)\Delta x_k + \nabla_x f(x_k) + \nabla_x h(x_k)(v_k + \Delta v_k) = 0 \tag{D.30}$$

and subsequently comparing them to the corresponding first-order necessary conditions of QP (D.26):

$$\nabla_{xx}^2 \mathcal{L}(x_k, v_k)\Delta x_k + \nabla_x f(x_k) + \nabla_x h(x_k) v_k^{qp} = 0 \tag{D.31}$$

From comparison of the last terms of (D.30) and (D.31), it follows that the Lagrange multipliers for the equality constrained versions of the two QPs (D.24) and (D.26) are related according to the equality $v_k^{qp} = v_k + \Delta v_k$. The same link can be followed for the case that also inequality constraints are present: $u_k^{qp} = u_k + \Delta u_k$.

In order to reduce the computational load, many SQP algorithms do not use the QP forms considered here directly but replace the corresponding Hessians for an approximation B_k. This will be discussed now in more detail. This will be done for the QP (D.26) with the reason being that it is the one that has been chosen for the SQP algorithm presented here. The reason for the latter is that it is the one that is most often found in the literature.

D.3.3 Approximation of the Hessian

Often, computation of the exact Hessian $\nabla_{xx}^2 \mathcal{L}(x_k, u_k, v_k)$ takes too much time. A suitable alternative then is to approximate the Hessian, in particular with (amongst others) first order *c.q.* gradient information. Approximation of the Hessian will generally lead to a lower speed of convergence, *i.e.* a larger number of iterations for the SQP algorithm in comparison to the same SQP algorithm using the exact Hessian. This lower convergence speed is, however, compensated for by a significant decrease in time spent on computing the Hessian, thereby effectively resulting in a lower computation time for the SQP algorithm.

Approximating the Hessian is common practice in optimization applications. Unconstrained optimization methods that employ an approximation of the Hessian are referred to as *quasi*-Newton methods (see *e.g.* [71]) as the approximation of the Hessian corresponds to modifying the Newton equations that are used to compute a new search direction. The quasi-Newton updating techniques that are used with these methods are also used within constrained optimization methods.

SQP methods either use *full* or *reduced* Hessian approximations [8]. Reduced Hessian SQP methods approximate only the portion of the Hessian matrix relevant to a particular subspace; these methods will not be discussed here. Instead, one is referred to [8]. Full Hessian SQP methods typically use a so called *secant* approximation of the Hessian, *i.e.* a Hessian approximation that fulfills the so called *secant equation* or *condition* to be given below. This condition is based on the fact that the Hessian $\nabla^2_{xx}\mathcal{L}(x_k, u_k, v_k)$ fulfills the equation

$$[\nabla^2_{xx}\mathcal{L}(x_k, u_k, v_k)](x_k - x_{k-1}) \approx \nabla_x\mathcal{L}(x_k, u_k, v_k) - \nabla_x\mathcal{L}(x_{k-1}, u_k, v_k) \quad \text{(D.32)}$$

with equality if the Lagrangian function is quadratic. Considering this equation, it makes sense to require an approximation of the Hessian $\nabla^2_{xx}\mathcal{L}(x_k, u_k, v_k)$, denoted as B_k, to fulfill the equation

$$B_k(x_k - x_{k-1}) = \nabla_x\mathcal{L}(x_k, u_k, v_k) - \nabla_x\mathcal{L}(x_{k-1}, u_k, v_k) \quad \text{(D.33)}$$

This equation is the mentioned secant equation or condition. A common procedure for generating the Hessian approximations that fulfill this condition is by using an *update formula* of the form

$$B_{k+1} = B_k + U_k \quad \text{(D.34)}$$

where the matrix U_k represents an update to the approximation of the Hessian obtained at iteration k. A variety of such quasi-Newton update formulas are available, each one enforcing particular characteristics on the Hessian approximation. See [8] for an overview of available quasi-Newton update formulas (one of them associated with the beautiful name of SALSA-SQP). The most widely used update formula, and the one that is considered to be most effective, is the so called *BFGS* update formula which derives its name from the four people who developed it: Broyden, Fletcher, Goldfarb and Shanno [71]. This update formula is also used within the SQP method presented in this appendix and will therefore be discussed in more detail. It is given as

$$B_{k+1} = B_k + \frac{y_k y_k^T}{s_k^T y_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} \quad \text{(D.35)}$$

with

$$s_k = x_{k+1} - x_k \quad \text{(D.36)}$$

and

$$y_k = \nabla_x\mathcal{L}(x_{k+1}, u_{k+1}, v_{k+1}) - \nabla_x\mathcal{L}(x_k, u_{k+1}, v_{k+1}) \quad \text{(D.37)}$$

Note that with the Hessian of the Lagrangian function approximated this way, the QP subproblem that is solved each iteration k for the SQP method presented in this appendix is given as

$$\begin{array}{ll} \min_{\Delta x_k} & \frac{1}{2}\Delta x_k^T B_k \Delta x_k + [\nabla_x f(x_k)]^T \Delta x_k \\ s.t. & [\nabla_x g(x_k)]^T \Delta x_k \leq -g(x_k) \\ & [\nabla_x h(x_k)]^T \Delta x_k = -h(x_k) \end{array} \quad \text{(D.38)}$$

One important requirement for the approximation of the Hessian B_k is that it needs to be positive definite: else the QP subproblem (D.38) may not have a solution. It is known for the BFGS updating formula that B_{k+1} is positive definite if B_k is positive definite and

$$y_k^T s_k > 0 \tag{D.39}$$

Positive definiteness of B_{k+1} can therefore be guaranteed by choosing an initial B_0 that is positive definite and by making sure that for all subsequent iterations condition (D.39) is fulfilled. Unfortunately, fulfillment of the latter condition cannot be guaranteed. An alternative way of ensuring positive definiteness of the Hessian approximations is by replacing y_k in (D.35) and (D.39) by

$$\hat{y}_k = \theta y_k + (1-\theta) B_k s_k \tag{D.40}$$

where the parameter $\theta \in (0, 1]$ is chosen each iteration k such that condition (D.39) is fulfilled. One way to choose θ is to use an iterative procedure called *backtracking* where for subsequent $i = 0, 1, 2, ...$ this parameter is chosen as $\theta = 2^{-i}$ until condition (D.39) is fulfilled (with y_k replaced by $\hat{y}_k$, of course). A parameter θ fulfilling this condition can always be found this way. Maintaining positive definiteness of the Hessian approximations via eqn. (D.40), etc. is used in the so called Powell-SQP method [8] and is also used in the SQP algorithm proposed in this appendix with θ chosen according to the backtracking procedure discussed above.

D.3.4 Line search

Newton's method prescribes an update of the current estimates for the solution and Lagrange multipliers of the form

$$\begin{bmatrix} x_{k+1} \\ u_{k+1} \\ v_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ u_k \\ v_k \end{bmatrix} + \begin{bmatrix} \Delta x_k \\ \Delta u_k \\ \Delta v_k \end{bmatrix} \tag{D.41}$$

However, an SQP algorithm generally does not proceed to its solution according to this iteration as this may easily cause this algorithm to be non-*globally convergent*: an optimization algorithm is said to be globally convergent if it converges, under suitable conditions, to some *local* (so not the global) solution from any remote starting point [8]. Non-convergence may occur with SQP, and also with Newton's method, if the initial point is chosen too far away from a local minimum. In order to ensure global convergence, SQP methods employ a so called *globalization strategy* [71]. In particular, most SQP methods employ a so called *line search*. In this case, the iteration becomes of the form

$$\begin{bmatrix} x_{k+1} \\ u_{k+1} \\ v_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ u_k \\ v_k \end{bmatrix} + \alpha \begin{bmatrix} \Delta x_k \\ \Delta u_k \\ \Delta v_k \end{bmatrix} \tag{D.42}$$

with $\alpha \in (0, 1]$ being a so called *step length* parameter that is chosen such that it fulfills a condition on a so called *merit function* $\phi(x_k, u_k, v_k)$ *e.g.* such that

$$\phi(x_{k+1}, u_{k+1}, v_{k+1}) < \phi(x_k, u_k, v_k) \tag{D.43}$$

The merit function reflects progress towards a solution: a condition on this merit function in the context of ensuring global convergence of an SQP method, as *e.g.* given by eqn. (D.43), aims at ensuring that $\{x_{k+1}, u_{k+1}, v_{k+1}\}$ be a better "estimate" of the solution than $\{x_k, u_k, v_k\}$. A natural merit function for the unconstrained case is the NLP objective function. In the constrained case, the merit function generally is a sum of the NLP objective function and a term that reflects the amount of infeasibility of the constraints. The reason for adding this second term is that progress in the objective function may come at the expense of feasibility, and vice versa. Well known merit functions are the l_1 merit function, the augmented Lagrangian function (see [8] for both these merit functions) and the quadratic penalty function (see [71]). The l_1 merit function is used within the SQP algorithm proposed in this appendix and is, therefore, discussed in more detail below. A standard requirement for an SQP method to ensure it is globally convergent is that each of the solutions of its QP subproblem Δx_k is a descent direction for the merit function. A somewhat stronger condition that often is found is that each of these solutions be *sufficiently* descent ([71]). Another standard requirement that one ideally would like to see to be fulfilled for a merit function to ensure global convergence of the SQP method is that the solutions of the NLP coincide with the unconstrained minimizers of this merit function. However, this can rarely be guaranteed. See [71] for more details.

From the point of view of *local convergence* behaviour of an SQP method, *i.e.* the behaviour of the iterates close to a local minimum, α is ideally chosen equal to 1 as this leads to the fastest local convergence rate. Being an implementation of Newton's method, and thereby inheriting its local convergence behaviour, this convergence rate is then (under some additional, seemingly reasonable, assumptions for which one is referred to [8]) quadratic when the true Hessian is used and linear or superlinear when the Hessian is approximated. However, it may happen that the globalization strategy prevents the usage of the step size $\alpha = 1$ (which is also denoted as the Newton step) close to a solution. In particular, this feature is absent when using certain merit functions. In these cases it is possible to give examples where a step of $\alpha = 1$ is unacceptable (for the globalization strategy) at every iteration, no matter how close the current point is to the solution [71].

Several methods of adapting α, in the context of a globalization strategy, can be found in the literature. These methods differ with respect to the merit function that is used and the way that α is computed with this merit function. Given a merit function, α is computed such that it fulfills certain conditions on the merit function and on itself. In particular, α should be computed such that at each iteration (i) the merit function is sufficiently decreased in value and (ii) this step lenght is not too small. A typical condition that is used to ensure the sufficient decrease condition is the so called *Armijo* condition. This condition is also used in the SQP algorithm proposed here (eqn. (D.44) is an Armijo condition). Methods of ensuring a not too small step length parameter α are (i) fulfilling the so called *Wolfe* condition, see [71], and (ii) backtracking (see section D.3.3). (N.B. somewhat confusingly, in [8] the Armijo and Wolfe condition are collectively denoted as *the Wolfe conditions*). In the SQP algorithm presented here backtracking is used to ensure a not too small α. (More specific, this step length parameter is chosen as $\alpha = 2^{-i}$ with $i = 0, 1, 2, ...$ until condition (D.44) is fulfilled.)

An alternative framework for guaranteeing convergence is offered by the so called *trust region* methods. These methods will not be discussed here. Instead, one is referred to [71] and, for a discussion of trust region methods within the context of SQP, [8].

As already discussed above, in the SQP algorithm presented in this appendix a globalization strategy is used where α is chosen each iteration according to a backtracking procedure such that it fulfills the Armijo condition

$$\phi(x_k + \alpha \Delta x_k) \leq \phi(x_k) + \zeta \alpha \overline{D}(\phi(x_k); \Delta x_k) \tag{D.44}$$

with $\zeta \in (0, \frac{1}{2})$ some a priori chosen constant. The merit function that is used in this algorithm is the so called l_1 *merit function*

$$\phi(x) = f(x) + \rho[\|g^+(x)\|_1 + \|h(x)\|_1] \tag{D.45}$$

(indeed, l_1 refers to the usage of the l_1-norm $\|...\|_1$, see eqn. (D.49), to measure infeasibilities) where each element $g_i^+(x)$ of $g^+(x)$ is computed as

$$g_i^+(x) = \{ \begin{array}{ccc} 0 & if & g_i(x) \leq 0 \\ g_i(x) & if & g_i(x) > 0 \end{array} \tag{D.46}$$

and with

$$\begin{aligned} \overline{D}(\phi(x_k); \Delta x_k) &= [\nabla_x f(x_k)]^T \Delta x_k - \rho[\|g^+(x_k)\|_1 + \|h(x_k)\|_1] \\ &= [\nabla_x f(x_k)]^T \Delta x_k - (\phi(x_k) - f(x_k)) \end{aligned} \tag{D.47}$$

representing an *upper bound* on the so called *directional derivative of* ϕ *along* Δx_k *at* x_k (see section D.4.3 for a definition of directional derivative and a derivation of this upper bound. See also *e.g.* [71]) and

$$\rho = \| \left[\begin{array}{c} u_k^{qp} \\ v_k^{qp} \end{array} \right] \|_\infty + \overline{\rho} \tag{D.48}$$

with $\overline{\rho} > 0$ also some *a priori* chosen constant. The norms $\|..\|_1$ and $\|..\|_\infty$ are defined as

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \tag{D.49}$$

resp.

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i| \tag{D.50}$$

The motivation for using an l_1 merit function is its simplicity. Nevertheless, the usage of this merit function is accompanied by one disadvantage denoted as the *Maratos effect* which is the phenomenon that the globalization strategy may not allow a step length of $\alpha = 1$ near the solution [8]. This may slow the rate of convergence of the SQP method considerably. If interested, several, *ad hoc*, procedures to overcome the Maratos effect can be found in the literature. See *e.g.* [8]. It must be noted that problems like the Maratos effect were effectively solved with the SPQ method proposed here by setting limits on the number of iterations involved in this method. See section D.3.5. Nevertheless, future research could be aimed at the usage of other merit functions that do allow the Newton step $\alpha = 1$ to be taken near the solution.

The globalization strategy used in the SQP algorithm presented in this appendix, as discussed above, is based on the one given in section 4.2 of [8] for the case of only equality constraints being present in the NLP. It differs (slightly) from the one given in this and other references in the usage of an *upper bound* on the directional derivative of the l_1 merit function instead of the directional derivative itself (compare the SQP method proposed here with *e.g.* the one in [34] (where also an expression for the directional derivative of the l_1 merit function can be found)). More background on and motivation for the globalization strategy chosen here can be found in section D.4.3 when discussing the global convergence properties of the proposed SQP algorithm. Local convergence properties of this algorithm are discussed in section D.4.2. First now, measures are discussed to improve the robustness of SQP methods. After that, having discussed all of its main ingredients, a global outline of the proposed SQP algorithm is given.

D.3.5 Robustness measures

An SQP algorithm often must be equipped with measures to enhance its robustness against failure of the algorithm to deliver a solution within some pre-specified time or to deliver a solution at all. This is particularly the case for NMPC applications. Here, briefly, some SQP failure types are discussed. Also, for some of these failure types possible solutions are given to prevent them.

Types of failure that may occur with SQP methods are *e.g.*:

- The Hessian does not fulfill the requirements needed for proper solution of the QP. Such requirements are *e.g.* positive definiteness and well conditioning.
- The SQP method requires too many iterations to reach a local optimum within a prespecified time.
- Infeasibility

Hessian positive-definiteness and well conditioning issues can be relatively easily dealt with and are not further discussed here (See also section D.3.3, where a specific measure is discussed for ensuring a positive definite Hessian).

With respect to the second type of failure it is important to note that no guarantee exists for SQP methods to converge to a local minimizer within an *a priori* defined number of iterations. In fact, several iteration loops are often performed within an SQP method and each of them must have reached its optimum before the prespecified computation time (*i.e.* sample time in NMPC applications) is reached. For instance, in case of the SQP method presented in this appendix, the following two types of iterations are of concern:

- the number of QP (Newton) iterations, with a QP iteration defined as one where, through solving a QP, new search directions for the optimization variables and Lagrange multipliers are computed

- the number of iterations used in the line search backtracking procedure to determine the step length α

Experience with the applications discussed in chapter 7 has shown that the simple robustness measure of setting an upper limit on the numbers of these iterations works well. In particular, setting a limit on the number of QP iterations prevents the SQP method to keep on searching for an improved solution while already being close to one. Setting an upper limit on the number of line search iterations proved to be particularly helpful in preventing the SQP method to get stuck in a Maratos effect-like routine where the SQP method continuously choses a very small value for α. Note that these upper limits may not be equal to the optimal numbers of iterations and that the suboptimal solutions obtained at these limits have to be acceptable. The latter is generally the case when the limits are chosen carefully, as experience has shown. For instance, experience showed that the upper limit on the number of line search iterations should not be chosen too small, *e.g.* not below 3, as then the jump in solution obtained after violation of the Armijo condition may sometimes be too high and convergence of the algorithm may not be obtained within the imposed time limit.

The SQP method may not be able to find a feasible solution, especially when the constraints are chosen very tight. This may cause such problems as the algorithm to break down. Infeasibility problems can be solved by introducing *slack variables* in either the NLP to be solved or in the QP to be solved. It is not clear at the moment which of these two solutions is best. Here, the second one is discussed, which has proven useful in applications such as discussed in chapter 7.

One generally applied solution to the problem of avoiding an infeasible solution to the QP subproblem is the usage of an alternative QP subproblem, related to the original one, that allows the violation of the hard constraints of the original QP subproblem, though as little as possible. This alternative QP subproblem is solved with some or all of these hard constraints replaced by soft constraint counterparts. This is done by introducing additional variables called slack variables into both the constraint equations and the objective function of the QP subproblem. These slack variables represent the allowed violation of the constraints. One such slack-variable extended QP subproblem formulation, one that takes into account possible violations of inequality constraints only, is *e.g.* (compare with QP (D.38))

$$\begin{aligned} \min_{\Delta x_k,\,\theta} \quad & \frac{1}{2}\Delta x_k^T B_k \Delta x_k + [\nabla_x f(x_k)]^T \Delta x_k + [M_1\ ...\ M_{n\theta}]\theta \\ s.t. \quad & g(x_k) + [\nabla_x g(x_k)]^T \Delta x_k \leq \theta \\ & h(x_k) + [\nabla_x h(x_k)]^T \Delta x_k = 0 \\ & \theta \geq 0 \end{aligned} \tag{D.51}$$

where θ is an $n\theta$-dimensional vector (with $n\theta$ equal to the number of inequality constraints) containing the slack variables, which have to computed by the QP algorithm now in addition to Δx_k, and M_1, ..., $M_{n\theta}$ are positive constants that should be chosen sufficiently large (*e.g.* as 10^8) to penalize large values for the corresponding elements of θ. This QP subproblem formulation is closely related to the "big M"-method of [8].

An important issue when implementing a slack-variable extended QP subproblem is the minimization of additional computation time. A first aspect related to this issue concerns the choice of the number of slack variables: the higher this number, the more one can localize the constraint violation but also the higher the computation time. Hence, a compromise must be sought between flexibility towards constraint violation and computation time. A second aspect related to the issue of computation time concerns the implementation of the QP subproblem. If one simply replaces the original QP for the slack-variable QP one simultaneously solves for the directions and slack-variables at every QP iteration. As infeasibility generally occurs only rarely, this implies that most of the times a too large QP problem is solved, *i.e.* a problem that could as well have been solved without solving also for the slack-variables. An alternative implementation that prevents this is one where in a preliminary stage it is determined whether the solution will be infeasible or not and then is decided to solve the original or slack-variable extended QP. This approach ensures that the slack-variable QP is only solved when there is infeasibility while for the rest the original QP is solved. This may effectively lead to a lower overall computation time in comparison to the first approach of always solving the slack-variable extended QP. The second approach has therefore been implemented in the SQP method presented in this appendix, taking advantage of a built-in feature of the QP solver that consists of determining the feasibility of the QP and of breaking down the QP solver and delivering a notifying signal to the outside world when infeasibility has been detected. In addition, it is assumed that a slack-variable extended QP subproblem of the form of (D.51) is used.

D.3.6 Outline of the SQP algorithm

Having discussed all essential elements in the previous sections, the SQP algorithm proposed in this appendix can be outlined as follows:

1. Choose initial values for x_0, u_0, v_0 and B_0, the latter *e.g.* as the identity matrix I. In addition, choose appropriate values for $\overline{\rho}$ and ζ. Also, choose values for the parameters that determine the termination criterion.

2. Using the chosen initial values for x_0, u_0, v_0 and B_0, compute $f(x_0)$, $h(x_0)$, $g(x_0)$, $g^+(x_0)$, $\nabla_x f(x_0)$, $\nabla_x h(x_0)$, $\nabla_x g(x_0)$.

3. Using these computed quantities, check whether the termination criterion is fulfilled. As long as this is not the case, perform the steps of the following loop:

   (a) Solve QP (D.38) to obtain values for Δx_0, u_0^{qp} and v_0^{qp}. If infeasibility is detected during solution of this QP, obtain values for these quantities by solving a slack-variable extended QP subproblem of the form of (D.51).

   (b) Compute ρ, $\phi(x_0)$ and $\overline{D}(\phi(x_0);\Delta x_0)$ using, amongst others the values obtained at the previous step.

   (c) Set $nr_{iter,\alpha} = 0$ (start iteration to determine step length).

   (d) Compute $\alpha = 2^{-nr_{iter,\alpha}}$ and $x_1 = x_0 + \alpha\Delta x_0$.

   (e) Compute $f(x_1)$, $h(x_1)$, $g(x_1)$, $g^+(x_1)$, $\nabla_x f(x_1)$, $\nabla_x h(x_1)$ and $\nabla_x g(x_1)$.

   (f) Compute $\phi(x_1)$ from the values computed at step (e) together with the computed value for ρ.

(g) Check whether $\phi(x_1) \leq \phi(x_0) + \zeta\alpha\overline{D}(\phi(x_0); \Delta x_0)$ (see eqn. (D.44)). If so, then proceed with step (h), else set $nr_{iter,\alpha} = nr_{iter,\alpha} + 1$ and return to step (d).

(h) Set $u_1 = u_0 + \alpha(u_0^{qp} - u_0)$ and $v_1 = v_0 + \alpha(v_0^{qp} - v_0)$.

(i) Compute s_0 according to eqn. (D.36). (Start computation of B_1 according to BFGS update formula).

(j) Compute y_0 according to eqn. (D.37).

(k) Set $nr_{iter,\theta} = 0$ (start iteration to determine positive definite Hessian approximation).

(l) Compute $\theta = 2^{-nr_{iter,\theta}}$ and subsequently $\hat{y}_0^T s_0$ (see eqn. (D.40)).

(m) Check whether $\hat{y}_0^T s_0 > 0$; if this is true then proceed with step (n); else set $nr_{iter,\theta} = nr_{iter,\theta} + 1$ and return to step (l).

(n) Compute B_1 according to eqn. (D.35) with $y_0 = \hat{y}_0$.

(o) Check whether the Hessian B_1 is sufficiently well conditioned and, if not, adapt this Hessian until it is.

(p) Set $x_0 = x_1$, $u_0 = u_1$, $v_0 = v_1$, $B_0 = B_1$, $f(x_0) = f(x_1)$, $h(x_0) = h(x_1)$, $g(x_0) = g(x_1)$, $g^+(x_0) = g^+(x_1)$, $\nabla_x f(x_0) = \nabla_x f(x_1)$, $\nabla_x h(x_0) = \nabla_x h(x_1)$ and $\nabla_x g(x_0) = \nabla_x g(x_1)$. Return to the very start of this loop (check termination criterion, etc.).

4. End the algorithm by setting the desired optimal values for x_k, u_k and v_k (if the latter two quantities are desired) equal to the obtained values for x_0, u_0 and v_0.

D.4 Convergence properties of the proposed SQP algorithm

D.4.1 Introduction

In this section the convergence properties of the SQP algorithm presented in section D.3 are discussed. As may have become clear from the discussion in this section, convergence properties of an SQP algorithm consist of *local* and *global* convergence properties. Local convergence properties are related to the behavior of the iterates close to a local minimum. More specific, local convergence analysis establishes conditions under which the iterates converge to a solution and at what rate, given that the starting data (x_0, u_0, v_0 and B_0) are sufficiently close to the corresponding data at a solution [8]. Global convergence analysis establishes conditions under which the SQP algorithm converges to some local minimum from any remote starting point, *i.e.* from any starting point for which the data are *not* suffciently close to the corresponding data of a local minimum.

Local convergence properties of the proposed SQP algorithm are briefly discussed in section D.4.2. Its global convergence properties are discussed in section D.4.3. For the purpose of explanation, only part of the conditions underlying these convergence properties is given here. The remaining part of these conditions can be found in [8], on which the results to be presented here rely heavily.

In addition to investigating the local and global convergence properties of an SQP algorithm, it is also of importance to investigate the problems this SQP algorithm might have with respect to the *transition* from global to local convergence. Apart from the Maratos effect, see sections D.3.4, D.4.2 and D.3.5, such problems will not be discussed in this appendix. For that, one is referred to [8].

D.4.2 Local convergence behavior

Briefly summarized, the iterates obtained with the SQP algorithm proposed above, x_k, u_k and v_k converge *superlinearly* to a local minimum, x^*, u^* and v^*, if these iterates are sufficiently close to such a local minimum in some sense and some additional conditions are fulfilled. (See eqn. (D.11) for a definition of superlinear convergence rate). The complete result can be found in [8] in the form of theorem 3.7. Two important assumptions that underly this local convergence theorem are the following:

- the active inequality constraints for the NLP (D.1) at x^* are known
- the merit function allows a steplength of $\alpha = 1$ to be used near x^*

The motivation for the first assumption is that it allows the local convergence behavior of the SQP algorithm to be investigated on the basis of a QP that only has equality constraints, including the active inequality constraints. This assumption can be jusitified for many of the SQP algorithms because the QP subproblem will have the same active constraints as the NLP possesses at a local minimum when the iterates are near this minimum [8]. The motivation for the second assumption is that the local convergence result for the SQP algorithm, as for many other SQP and non-SQP optimization methods, is based on Newton's method [8] for which the step length α is equal to 1. As may have become clear from the discussion in section D.3.4, this assumption may be violated by the Maratos effect, which is the effect that the l_1 merit function (D.45) may not allow a steplength of $\alpha = 1$ near the solution. Hence, the Maratos effect can prohibit the superlinear convergence rate of the proposed SQP algorithm.

D.4.3 Global convergence behavior

The global convergence properties of the proposed SQP method can be summarized by the following results:

- under certain conditions, the SQP method converges from any starting point to a stationary point of NLP (D.1)
- convergence to a local minimum of the NLP can*not* be guaranteed
- any local minimum of the NLP to which the SQP method converges, if it converges to such a minimum, is also a local minimum for the l_1 merit function

The first of these results is derived later on in this section. The second result follows from the fact that a stationary point may also be a saddle point or a maximum. The third result follows from a well-known result in optimization theory that states that a local minimizer of the NLP is also local minimizer of the l_1 merit function if the parameter ρ (see eqns. (D.45) - (D.48)) in this latter function is chosen sufficiently large, in particular larger than or equal to the largest, in absolute value, of the Lagrange multipliers at the local minimum of the NLP (see *e.g.* lemma 16.2 of [71], in particular also its proof, or the discussion on exact penalty functions in [20].). This condition on ρ, obviously, cannot be fulfilled *a priori* for the SQP algorithm due to the fact that at the start of this algorithm the Lagrange multipliers at any local minimum of the NLP are still unknown. However, if the SQP algorithm *überhaupt* converges to a local minimizer of the NLP, the Lagrange multipliers of its QP subproblem would approach the Lagrange multipliers of this local minimizer of the NLP. As a result, condition (D.48) then ensures that the mentioned condition on ρ would be fulfilled; this follows from substituting in (D.48) the Lagrange multipliers corresponding to the local minimum of the NLP for those corresponding to the QP subproblem.

The rest of this section is now devoted to deriving the first of the three results given above for the SQP algorithm presented in this appendix. This derivation closely follows the derivation of the corresponding theorem of section 4.2 of [8] for the case that only equality constraints are present in the NLP. The first step in the derivation is to show that the direction Δx_k computed with the QP subproblem of the proposed SQP method is a descent direction for the l_1 merit function $\phi(x)$. For that purpose, the following proposition is first needed:

Proposition D.4.1 *The directional derivative $D(\|g^+(x_k)\|_1; \Delta x_k)$ of $\|g^+(x)\|_1$ at x_k along Δx_k, with the latter being obtained as the solution of QP (D.38), fulfills the following inequality:*

$$D(\|g^+(x_k)\|_1; \Delta x_k) \leq -\|g^+(x_k)\|_1 \tag{D.52}$$

Proof D.4.2 *The directional derivative of a function $f(x)$ at x_k along the direction Δx_k is defined as:*

$$\lim_{\alpha \to 0} \frac{f(x_k + \alpha \Delta x_k) - f(x_k)}{\alpha} \tag{D.53}$$

Using this definition it can be derived that the directional derivative $D(g_i^+(x_k); \Delta x_k)$ of any of the inequality constraints $g_i^+(x)$ of the NLP at a point x_k along the direction Δx_k is given as follows:

- *if $g_i^+(x_k) = 0$ and $g_i(x_k) < 0$ then*

$$D(g_i^+(x_k); \Delta x_k) = 0 = -|g_i^+(x_k)| \tag{D.54}$$

- *if $g_i^+(x_k) = 0$ and $g_i(x_k) = 0$ then either*

$$D(g_i^+(x_k); \Delta x_k) = 0 = -|g_i^+(x_k)| \tag{D.55}$$

or

$$\begin{aligned} D(g_i^+(x_k); \Delta x_k) &= D(g_i(x_k); \Delta x_k) \\ &= [\nabla_x g_i(x_k)]^T \Delta x_k \leq -g_i(x_k) \\ &= 0 = -|g_i^+(x_k)| \end{aligned} \tag{D.56}$$

where the inequality follows from fulfillment of the inequality constraints in the QP subproblem (D.38).

- *if* $g_i^+(x_k) > 0$ *then (*$g_i^+(x_k) = g_i(x_k)$ *and:)*

$$\begin{aligned} D(g_i^+(x_k); \Delta x_k) &= D(g_i(x_k); \Delta x_k) \\ &= [\nabla_x g_i(x_k)]^T \Delta x_k \leq -g_i(x_k) \\ &= -g_i^+(x_k) = -|g_i^+(x_k)| \end{aligned} \tag{D.57}$$

where the inequality again follows from fulfillment of the inequality constraints in the QP subproblem (D.38).

From these expressions it follows that the directional derivative $D(g_i^+(x_k); \Delta x_k)$ *can be characterized by the inequality*

$$D(g_i^+(x_k); \Delta x_k) \leq -|g_i^+(x_k)| \tag{D.58}$$

This inequality implies that

$$\begin{aligned} D(\|g^+(x_k)\|_1; \Delta x_k) &= D(\sum_i |g_i^+(x_k)|; \Delta x_k) \\ &= D(\sum_i g_i^+(x_k); \Delta x_k) \\ &= \sum_i D(g_i^+(x_k); \Delta x_k) \\ &\leq -\sum_i |g_i^+(x_k)| \\ &= -\|g^+(x_k)\|_1 \end{aligned} \tag{D.59}$$

which is the inequality to be proved (D.52) of proposition D.4.1.

Using inequality (D.52), the directional derivative $D(\phi(x_k); \Delta x_k)$ of the l_1 merit function ϕ at x_k along Δx_k can be written as

$$\begin{aligned} D(\phi(x_k); \Delta x_k) &= D(f(x_k) + \rho[\|g^+(x_k)\|_1 + \|h(x_k)\|_1]; \Delta x_k) \\ &= D(f(x_k) + \rho\|g^+(x_k)\|_1; \Delta x_k) + D(\rho\|h(x_k)\|_1; \Delta x_k) \\ &= D(f(x_k) + \rho\|g^+(x_k)\|_1; \Delta x_k) + \rho D(\|h(x_k)\|_1; \Delta x_k) \\ &\leq D(f(x_k) + \rho\|g^+(x_k)\|_1; \Delta x_k) - \rho\|h(x_k)\|_1 \end{aligned} \tag{D.60}$$

if $\rho \geq 0$. Using the fact that

$$D(f(x_k) + \rho\|h(x_k)\|_1; \Delta x_k) = [\nabla_x f(x_k)]^T \Delta x_k - \rho\|h(x_k)\|_1 \tag{D.61}$$

(see eqn. (4.19) of [8]), inequality (D.60) can be written as

$$\begin{aligned} D(\phi(x_k);\Delta x_k) &\leq [\nabla_x f(x_k)]^T \Delta x_k - \rho[\|h(x_k)\|_1 + \|g^+(x_k)\|_1] \\ &:= \overline{D}(\phi(x_k);\Delta x_k) \end{aligned} \tag{D.62}$$

It can be seen that $\overline{D}(\phi(x_k);\Delta x_k)$ represents an upper bound on the directional derivative $D(\phi(x_k);\Delta x_k)$ of ϕ along Δx_k at x_k, as stated in section D.3.4. The idea now is to derive conditions that ensure that this overbound on the directional derivative is smaller than zero for every $\Delta x_k \neq 0$: if these conditions are fulfilled, the directional derivative itself is also smaller than zero for every such Δx_k. Hence, these conditions guarantee that Δx_k is a descent direction for the l_1 merit function at every iteration x_k until the SQP algorithm has converged. In order to arrive at these conditions, first, the expression for $\nabla_x f(x_k)$ obtained from the first-order necessary conditions for the QP subproblem (D.38)

$$B_k \Delta x_k + \nabla_x f(x_k) + \nabla_x g(x_k) u_k^{qp} + \nabla_x h(x_k) v_k^{qp} = 0 \tag{D.63}$$

is substituted in the expression for $\overline{D}(\phi(x_k);\Delta x_k)$ given by eqn. (D.62). This results in the following expression for $\overline{D}(\phi(x_k);\Delta x_k)$:

$$\begin{aligned} \overline{D}(\phi(x_k);\Delta x_k) &= -(\Delta x_k)^T B_k \Delta x_k - (\Delta x_k)^T \nabla g(x_k) u_k^{qp} - \\ &\quad (\Delta x_k)^T \nabla h(x_k) v_k^{qp} - \rho[\|h(x_k)\|_1 + \|g^+(x_k)\|_1] \end{aligned}$$

As (see [8])

$$-(\Delta x_k)^T \nabla h(x_k) v_k^{qp} = h(x_k)^T v_k^{qp} \leq \|v_k^{qp}\|_\infty \|h(x_k)\|_1 \tag{D.64}$$

and (exploiting the fulfillment of the strict complementary slackness condition at the solution of the QP subproblem),

$$-(\Delta x_k)^T \nabla g(x_k) u_k^{qp} = g(x_k)^T u_k^{qp} \leq g^+(x_k)^T u_k^{qp} \leq \|u_k^{qp}\|_\infty \|g^+(x_k)\|_1 \tag{D.65}$$

it follows that

$$\begin{aligned} \overline{D}(\phi(x_k);\Delta x_k) &\leq -(\Delta x_k)^T B_k \Delta x_k - \\ &\quad (\rho - \|u_k^{qp}\|_\infty)\|g^+(x_k)\|_1 - (\rho - \|v_k^{qp}\|_\infty)\|h(x_k)\|_1 \end{aligned} \tag{D.66}$$

If B_k now fulfills the uniform positive definiteness condition

$$\mu_1 \|\Delta x_k\|^2 \leq (\Delta x_k)^T B_k \Delta x_k \qquad \forall\ k \tag{D.67}$$

for some positive constant μ_1 and with $\|.\|$ denoting the 2-norm, the first term on the right of (D.66) is negative for nonzero Δx_k. As a consequence, $\overline{D}(\phi(x_k);\Delta x_k)$ is guaranteed to be strict negative for nonzero Δx_k if ρ is chosen such that both $\rho > \|u_k^{qp}\|_\infty$ and $\rho > \|v_k^{qp}\|_\infty$. These latter two conditions are fulfilled if condition (D.48) is fulfilled. Hence, if B_k fulfills the uniform positive definiteness condition (D.67), Δx_k is a guaranteed descent direction for $\phi(x_k)$ for the proposed SQP algorithm. This result is summarized in the next proposition.

Proposition D.4.3 *Consider the SQP algorithm presented in this appendix with QP subproblem (D.38) and with a QP solver fulfilling the conditions of section D.2.4. Assume the uniform positive boundedness condition (D.67) to hold. Then, every vector Δx_k computed at the QP subproblem of this SQP algorithm is a descent direction for the l_1 merit function $\phi(x_k)$ until this algorithm has converged.*

In obtaining, now, the result that the proposed SQP algorithm converges from any starting point to a stationary point of the NLP (D.1), *i.e.* proposition D.4.4, a similar line will be followed here as in [8] for the derivation of a corresponding result for the case that only equality constraints are present in the NLP[1] (see section 4.2 of that reference). The idea applied there, and also here, is to show that Δx_k, $h(x_k)$ and $g^+(x_k)$ converge to 0 for $k \to \infty$. The fact that $\Delta x_k = 0$ for $k \to \infty$ implies that the first-order necessary conditions of the QP subproblem (D.38) then fulfill

$$\begin{aligned} B_k \Delta x_k + \nabla_x f(x_k) + \nabla_x g(x_k) u_k^{qp} + \nabla_x h(x_k) v_k^{qp} = \\ 0 + \nabla_x f(x_k) + \nabla_x g(x_k) u_k^{qp} + \nabla_x h(x_k) v_k^{qp} = 0 \end{aligned} \tag{D.68}$$

and that, hence, the first-order necessary conditions of the NLP (D.1) (see eqn. (D.4)) are fulfilled for the limiting point x_k. As also $h(x_k) = 0$ and $g^+(x_k) = 0$ for this very same limiting point, it is, by definition (see section D.2.2), a stationary point of the NLP (D.1).

Using assumption (D.67) and eqn. (D.66), the Armijo condition (D.44) can be written as

$$\begin{aligned} \phi(x_k + \alpha \Delta x_k) - \phi(x_k) &\leq \zeta \alpha \overline{D}(\phi(x_k); \Delta x_k) && \text{(D.69)} \\ &\leq -\zeta \alpha [(\Delta x_k)^T B_k \Delta x_k + \\ &\quad (\rho - \|u_k^{qp}\|_\infty) \|g^+(x_k)\|_1 + (\rho - \|v_k^{qp}\|_\infty) \|h(x_k)\|_1] \\ &\leq -\zeta \alpha [\mu_1 \|\Delta x_k\|^2 + \mu_2 \|g^+(x_k)\|_1 + \mu_3 \|h(x_k)\|_1] \end{aligned}$$

with $\mu_2 = \rho - \|u_k^{qp}\|_\infty$ and $\mu_3 = \rho - \|v_k^{qp}\|_\infty$ being positive constants determined by condition (D.48) on ρ (one of these two constants is equal to $\bar{\rho}$; the other one is larger). Note that $\alpha > 0$ due to the backtracking way of selection of this parameter with the SQP algorithm proposed here and that also ζ is chosen positive with this SQP algorithm. Assuming now the starting point x_0 and all succeeding iterates x_k of the SQP algorithm to lie in some compact set $\mathcal{C}$ (see *e.g.* [47] for a definition) and that, hence, the l_1 merit function is bounded below for all these iterates (which follows from *e.g.* the continuous mapping theorem 2.5-6 of [47]), it follows from (D.69) that

$$\sum_{k=0}^{\infty} [\|\Delta x_k\|^2 + \|g^+(x_k)\|_1 + \|h(x_k)\|_1] < \infty \tag{D.70}$$

and therefore

$$[\|\Delta x_k\|^2 + \|g^+(x_k)\|_1 + \|h(x_k)\|_1] \to 0 \tag{D.71}$$

[1] In this corresponding result (theorem 4.3 or 11, depending on which version of [8] is employed: the published or preliminary one), as confirmed from personal communication with the authors of [8], ”stationary point of ϕ_1” (*i.e.* the l_1 merit function) can be replaced by ”stationary” or ”critical point of NLP”.

as $k \to \infty$. This implies that Δx_k, $g^+(x_k)$ and $h(x_k)$ converge to 0 for $k \to \infty$. From the discussion above, the following proposition then immediately follows:

Proposition D.4.4 *Consider the SQP algorithm proposed in this appendix with QP subproblem (D.38) and with a QP solver fulfilling the conditions of section D.2.4. Assume the uniform positive boundedness condition (D.67) to hold. Also assume that the starting point x_0 and all succeeding iterates x_k of the SQP algorithm lie in some compact set $\mathcal{C}$. Then, this SQP algorithm converges to a stationary point of the NLP (D.1) from any starting point.*

More on global convergence of SQP methods can be found in *e.g.* [8, 34, 80]. A particularly interesting global convergence result on an SQP algorithm that is closely related to the one proposed here can be found in [34]. In this reference, NLPs with only inequality constraints are considered and the directional derivative of the l_1 merit function itself is used instead of its upper bound. A more thorough discussion of l_1 merit functions and its usage in optimization methods can be found in [26, 27].

D.5 Application: steady state optimization of an alkylation plant

In this section, an application is described of the proposed SQP algorithm to an NLP encountered in [15] (problem 13), which involves the determination of the optimal operating conditions for a (simplified) alkylation process. This application serves to show that, from a solution point of view, the SQP algorithm behaves similarly as some literature NLP solvers. A comparison from a computational point of view has been left out because of lack of reliable comparison data.

Details about the optimization problem (equations etc.) can be found in [15]. It involves an optimization over 10 variables with 3 equality constraints and 28 inequality constraints. The nonlinearity of this NLP exhibits itself in a nonlinear objective function, 2 nonlinear equality constraints and 4 nonlinear inequality constraints. The SQP algorithm proposed in this appendix was applied to this NLP and the resulting optimal values were compared to values reported in [15]. All these values were obtained under the same conditions, *i.e.* for the same NLP to be solved with the same starting estimates (including the same starting estimates for the Lagrange multipliers). The results are given in table D.1 (see at the end of this appendix). The values in this table corresponding to ”Reported values I” are obtained with the so called penalty/modified barrier method used by the authors of [15] (which, notably, uses exact Hessians). The values in table D.1 corresponding to ”Reported values II” are obtained with another NLP solver referred to in [15] which is of an unknown nature. The values corresponding to ”Values SQP I” were obtained with the SQP algorithm proposed in this appendix for a relatively high value for the termination criterion and, as a result, for a relatively low number of (18) QP iterations (which equals the number of Hessian evaluations mentioned in table D.1). Hence, for this termination criterion the SQP algorithm is not pushed to deliver a very optimal solution. This explains the relatively high objective function value (-1.758) in comparison with its counterparts reported in [15] (-1.771

and -1.765) and also the relatively large (though arguably still acceptable) constraint violations for this case, as can be seen from the values $\|h(x^*)\|_\infty = 0.0029871$ and $\|g^+(x^*)\|_\infty = 0.0006054$. In order to obtain a better solution, a new optimization was performed with the proposed SQP algorithm for a considerably lower value for the termination criterion. This, first of all, resulted in an increased number of (31) QP iterations. The corresponding newly obtained optimal values are given in table D.1 under ”Values SQP II”. One can see that the corresponding objective function value has decreased siginificantly in comparison to the value found for the SQP I case: -1.769 versus -1.758. It is even of about the same magnitude as for the ”Reported values I” case and the ”Reported values II” case, even better *c.q.* lower than for the latter case. In addition, the magnitude of the constraint violations also has decreased considerably, as evident by the values $\|h(x^*)\|_\infty = 8.9017 \times 10^{-11}$ and $\|g^+(x^*)\|_\infty = 0$ (!).

Summarizing, the SQP algorithm presented in this appendix is capable of delivering the same (good) solutions as other NLP solvers found in the literature. From a computational point of view, it is only noted that the number of Hessian evaluations for the penalty/modified barrier method is significantly higher than the numbers of Hessian evaluations found for the proposed SQP algorithm: 40 versus 18 and 31. This *may* indicate that the SQP algorithm proposed here is more efficiently.

A final note here is that the proposed SQP algorithm has also been tested, with successful outcome(s), on several other NLP problems, varying from relatively simple example problems given in [20] and [71] to moderately sized NMPC problems. It also has been used to solve a moderately sized Gibbs energy minimization problem for finding values for state variables corresponding to thermodynamic equilibrium.

Variable	Lower bound	Upper bound	Reported values I	Reported values II	Values SQP I	Values SQP II
x_1	0.0	2.00	1.698	1.704	1.693	1.698
x_2	0.0	16.0	15.80	15.85	14.67	15.82
x_3	0.0	0.12	0.0539	0.0543	0.0593	0.0541
x_4	0.0	5.00	3.031	3.036	3.027	3.031
x_5	0.0	2.00	2.000	2.000	2.000	2.000
x_6	85.0	93.0	90.1	90.1	90.8	90.1
x_7	90.0	95.0	95.0	95.0	95.0	95.0
x_8	3.0	12.0	10.482	10.476	9.842	10.493
x_9	1.2	4.0	1.561	1.562	1.562	1.562
x_{10}	145.0	162.0	153.5	153.5	153.5	153.5
$f(x^*)$	-	-	-1.771	-1.765	-1.758	-1.769
$\|g^+(x^*)\|_\infty$	-	-	-	-	0.0006054	0
$\|h(x^*)\|_\infty$	-	-	$1.5e-12$	-	0.0029871	$8.9017e-11$
# Hessian ev.	-	-	40	-	18	31

Table D.1: Comparison of the SQP algorithm with two other NLP solvers given in [15]. The optimal values for the latter two NLP solvers are given under "Reported values I" and "Reported values II". Those obtained with the SQP algorithm are given, for two different values of the termination criterion in order to show the influence of the choice for this value, under "SQP I" and "SQP II": see the text for more details. Here, x^* refers to the solution obtained with the NLP solvers.

Appendix E

Moving Horizon Estimation

E.1 Introduction

In this appendix, first the main idea behind moving-horizon estimation (MHE) is presented after which a simulation-based MSWC application of this state estimation strategy is discussed. This application involves the comparison of MHE with an extended Kalman filter (EKF: see section 6.3.4) on a MSWC plant model and has also been discussed in [57] to which one is referred to for more details.

E.2 The main idea

The main idea behind MHE can be stated as to obtain the optimal state estimate(s) via a fit, using some numerical optimization routine, on observed (input and) output measurements. This leads to an optimal control problem similar to the one solved in NMPC controllers to find the optimal MVs, with the criterion to be minimized typically being a least squares (LS) one reflecting the deviation between observed and predicted outputs and with the ability to include constraints.

In order to illuminate the main idea behind MHE, it is assumed that the plant to be monitored fullfils the following equations:

$$\begin{aligned} x_{k+1} &= f(x_k) + v_k \\ y_k &= h(x_k) + w_k \end{aligned}$$

with the aim set on finding an optimal estimate for the state x_{k+1} using measurements given up till (current) time k. Then, note that this state can be computed from integrating the system equations $f(.)$ over a certain time period $[k - N_{MHE}, ..., k]$ if the past state $x_{k-N_{MHE}}$ and system noise trajectory $\{v_{k-N_{MHE}}, ..., v_k\}$ are perfectly known:

$$\begin{aligned} x_{k-N_{MHE}+1} &= f(x_{k-N_{MHE}}) + v_{k-N_{MHE}} \\ \vdots \quad &\vdots \quad \vdots \end{aligned}$$

$$
\begin{aligned}
x_{k-1} &= f(x_{k-2}) + v_{k-2} \\
x_k &= f(x_{k-1}) + v_{k-1} \\
x_{k+1} &= f(x_k) + v_k
\end{aligned}
\tag{E.1}
$$

MHE exploits this fact by estimating the values for $x_{k-N_{MHE}}$ and $\{v_{k-N_{MHE}}, ..., v_k\}$ and, subsequently, use these estimates to find the desired estimate for x_{k+1} via integration of the system equations in the way described above. The estimation is performed via a LS fit on the output data observed over the same period $[k - N_{MHE}, ..., k]$. Denoting the desired estimates as $\hat{x}_{k-N_{MHE}}$ and $\{\hat{v}_{k-N_{MHE}}, ..., \hat{v}_k\}$ and additionally denoting the states resulting from the integration as $\{\hat{x}_{k-N_{MHE}+1}, ..., \hat{x}_{k+1}\}$, this fit corresponds to (simultaneously) minimizing, in LS sense, the following quantities:

$$
\begin{aligned}
\hat{w}_{k-N_{MHE}} &= y_{k-N_{MHE}} - h(\hat{x}_{k-N_{MHE}}) \\
\vdots \quad &\vdots \quad \vdots \\
\hat{w}_{k-1} &= y_{k-1} - h(\hat{x}_{k-1}) \\
\hat{w}_k &= y_k - h(\hat{x}_k)
\end{aligned}
\tag{E.2}
$$

This estimation problem is an optimal control problem similar to that employed in NMPC controllers for finding the optimal MVs and similar numerical optimization routines can be used to solve it. An important consequence of this numerical optimization approach to the estimation problem is that constraints can be included in this problem. Such constraints can be constraints on *e.g.* the estimated noise sequences or states:

$$
\begin{aligned}
\hat{v}_{k-i} \in \mathcal{V}; &\qquad i = 0...N_{MHE} \\
\hat{w}_{k-i} \in \mathcal{W}; &\qquad i = 0...N_{MHE} \\
\hat{x}_{k-i+1} \in \mathcal{X}; &\qquad i = 0...(N_{MHE}+1)
\end{aligned}
\tag{E.3}
$$

The resulting optimal control problem then typically takes the following form:

$$
\begin{aligned}
\min_{\hat{x}_{k-N_{MHE}}, \{\hat{v}_{k-N_{MHE}}, ..., \hat{v}_k\}} \quad & \textstyle\sum_{i=0}^{N_{MHE}} \|\hat{w}_{k-i}\|^2_{Q_w} + \sum_{i=0}^{N_{MHE}} \|\hat{v}_{k-i}\|^2_{Q_v} + ... \\
& \|\hat{x}_{k-N_{MHE}} - \overline{x}_{k-N_{MHE}}\|^2_{Q_x} \\
s.t. \quad & \\
& \hat{x}_{k-i+1} = f(\hat{x}_{k-i}) + \hat{v}_{k-i}; \qquad i = 0...N_{MHE} \\
& \hat{w}_{k-i} = y_{k-i} - h(\hat{x}_{k-i}); \qquad i = 0...N_{MHE} \\
& \hat{v}_{k-i} \in \mathcal{V}; \qquad i = 0...N_{MHE} \\
& \hat{w}_{k-i} \in \mathcal{W}; \qquad i = 0...N_{MHE} \\
& \hat{x}_{k-i+1} \in \mathcal{X}; \qquad i = 0...(N_{MHE}+1)
\end{aligned}
\tag{E.4}
$$

where $\overline{x}_{k-N_{MHE}}$ is an estimate for the initial state $x_{k-N_{MHE}}$ available from the previous estimation step. As can be seen from this optimal control problem formulation, typical additional terms in its objective function are one that punishes large values for the estimated noises $\{\hat{v}_{k-N_{MHE}}, ..., \hat{v}_k\}$ and one that punishes large deviations from the prior initial state $\overline{x}_{k-N_{MHE}}$.

It is noted that the horizon N_{MHE} is chosen fixed and, thereby, is a moving horizon similar to the control and prediction horizon in the NMPC optimal control problem for computing the MVs. Similar to the latter horizons, N_{MHE} is determined via tuning. This also applies to the weighting matrices Q_v, Q_w and Q_x. From an accuracy and stability point of view, a particularly important choice is that for Q_x, in particular for short horizons N_{MHE}. This weighting matrix determines to what extent past data, *i.e.* data before $k - N_{MHE}$, is taken into account in the MHE optimal control problem. These past data are reflected in the prior initial state $\overline{x}_{k-N_{MHE}}$. Too much or too little confidence in this state, reflected in the choice for Q_x, may give inaccurate results or even divergence. The effect of this initial state on the estimation result can be diminished by choosing a large value for N_{MHE}. However, this also results in a larger computation time.

Even though MHE is depicted here as a deterministic approach to the state estimation problem, attempts can be found in the literature to provide it with a probabilistic interpretation, in particular by introducing the notion of the *truncated* Gaussian distribution. See *e.g.* [84]. One motivation for these attempts may be the existing LS, hence deterministic, interpretation of the Kalman filter, which is based on probability theory (see *e.g.* [10] for this interpretation).

The main motivation for employing MHE over other state estimation techniques is the fact that this estimation technique can deal with constraints. More specific, it can deal *optimally* with constraints as other estimation techniques such as *e.g.* Kalman filter based techniques can deal with them in a *sub-optimal* way by means of clipping. The option of incorporating constraints may be used to improve the estimation result by *e.g.* incorporating knowledge that certain states with a physical interpretation, such as masses or concentrations, cannot become smaller than zero. Examples of the improvement in estimation performance that can be obtained by including (inequality) constraints in the estimation problem can be found in *e.g.* [84] and [85]. Apart from its optimal handling of constraints, another motivation for using MHE may be its good handling of system nonlinearities. In this respect, it performs better than the EKF, as can be seen from the MSWC plant example discussed in section E.3 (and [57]). However, other state estimators such as the Ensemble and Unscented Kalman filter may also handle system nonlinearities well.

It is outside the scope of this thesis to fully discuss all issues relevant to MHE. For that, one is referred to the literature dealing with this subject, *e.g.* [84, 85, 87, 91].

E.3 MHE versus EKF for an MSWC plant NMPC application

In this section, a simulation based comparison of the estimation accuracy of an MHE with that of an EKF is discussed for an MSWC plant application. Even more so, being a thesis about NMPC of such plants, this comparison is done in a closed-loop setting where the control performance of an NMPC controller employing the considered MHE is compared to that of an NMPC controller employing the considered EKF. The latter comparison is performed to not only show the difference in estimation accuracy of the two considered state estimators but also to show the effect of this accuracy on NMPC control performance.

The comparison is performed with an MSWC plant model of the type considered in chapter 2 being used as the plant to be controlled. The same model is employed by both NMPC controllers, including within the MHE and EKF employed by these controllers. More specific, this MSWC plant model has 4 MVs (waste inlet flow, speed of grate, primary air flow, secondary air flow), 4 states (amongst others, the waste layer mass and temperature) and 2 CVs (steam production and O_2-concentration). Both NMPC controllers are of the type considered in chapter 6 and are compared under the same (simulation based) experimental conditions and for the same typical MSWC plant control problem of suppressing as much as possible the disturbances on steam production and O_2 concentration. Further specifics of the simulations can be found in [57].

The results of the comparison are depicted in figures E.1 and E.2 and in table E.1 (which all are taken from [57]).

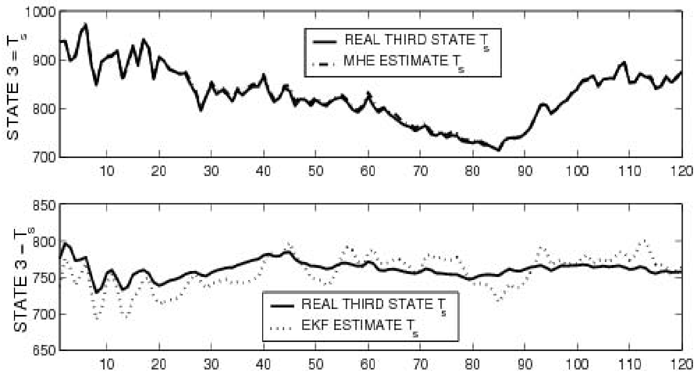

Figure E.1: EKF versus MHE when used in MSWC plant NMPC control- comparison of estimator performance: estimated versus real (third) state for an EKF (below) and MHE (above).

From these results it can be seen that, for the considered MSWC plant application, the MHE based NMPC controller performs considerably better than its EKF counterpart, both from an estimation and a control performance point of view, with the latter obviously directly proportional to the first. As constraints did not come into play for the estimators, this observed better performance for the MHE is subscribed to the better handling of model nonlinearities by this state estimator.

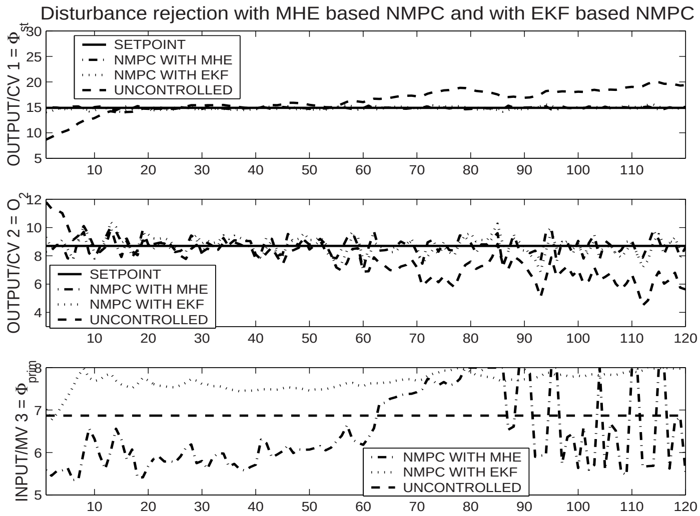

Figure E.2: EKF versus MHE when used in MSWC plant NMPC control - comparison of control c.q. disturbance rejection performance: steam (CV; upper part), O_2 (CV; middle part) and primary air flow (MV; lower part).

	PERF. MEASURE (in standard deviation) (sp = setpoint)	*MHE*	*EKF*
CONTROL PERFORMANCE	std(Φ_{st}^{sp} - $\Phi_{st,k}$)	0.19	0.31
	std(O_2^{sp} - $O_{2,k}$)	0.58	0.63
ESTIMATION PERFORMANCE	std($\hat{x}_{k+1}^1$-x_{k+1}^1)	0	$1.2 * 10^{-19}$
	std($\hat{x}_{k+1}^2$-x_{k+1}^2)	0.12	0.24
	std($\hat{x}_{k+1}^3$-x_{k+1}^3)	3.3	19.7
	std($\hat{x}_{k+1}^4$-x_{k+1}^4)	0.05	0.18

Table E.1: EKF versus MHE based MSWC plant NMPC control - estimator and control performances.

Bibliography

[1] *Physical properties of natural gases*. N.V. Nederlandse Gasunie, 1980.

[2] B.D.O. Anderson and J.B. Moore. *Optimal Filtering*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1979.

[3] E.M.B. Aske, S. Strand, and S. Skogestad. Coordinator mpc for maximizing plant throughput. *Computers and Chemical Engineering*, 32:195–204, 2008.

[4] K.J. Aström and R.D. Bell. Drum-boiler dynamics. *Automatica*, 36:363–378, 2000.

[5] B. Kouvaritakis, J.A. Rossiter, and M. Cannon. Closed-loop predictions in model based predictive control of linear and nonlinear systems (chapter 7). In *Nonlinear predictive control - theory and practice* (B. Kouvaritakis and M. Cannon (Eds.)), IEE Control Engineering Series, Volume 61, London, United Kingdom, pages 153–172, 2001.

[6] L.T. Biegler. An overview of simultaneous strategies for dynamic optimization. *Chemical Engineering and Processing*, 46:1043–1053, 2007.

[7] T. Binder, L. Blank, H.G. Bock, R. Bulirsch, W. Dahmen, M. Diehl, T. Kronseder, W. Marquardt, J.P. Schlöder, and O. von Stryk. Introduction to model based optimization of chemical processes on moving horizons. In *Online Optimization of Large Scale Systems. Eds.: M. Grötschel, S.O. Krumke and J. Rambau. Springer-Verlag*, pages 295–339, 2001.

[8] P.T. Boggs and J.W. Tolle. Sequential quadratic programming. *Acta Numerica*, 4:1–51, 1995.

[9] X. Bombois, G. Scorletti, M. Gevers, P. Van den Hof, and R. Hildebrand. Least costly identification experiment for control. *Automatica*, 42:1651–1662, 2006.

[10] R. Bos. *Monitoring of industrial processes using large-scale first principles models*. PhD thesis, Delft University of Technology, Delft, The Netherlands, 2006.

[11] R. Cagienard, P. Grieder, E.C. Kerrigan, and M. Morari. Move blocking strategies in receding horizon control. *Journal of Process Control*, 17:563570, 2007.

[12] A. Cervantes and L.T. Biegler. Optimization strategies for dynamic systems. In *Encyclopedia of Optimization. Eds.: C. Floudas and P. Pardalos. Kluwer*, 2000.

[13] N.-B. Chang and W.C. Chen. Fuzzy controller design for municipal incinerators with the aid of genetic algorithms and genetic programming techniques. *Waste Management and Research*, 18:429–443, 2000.

[14] W.C. Chen, N.-B. Chang, and J.-C. Chen. Ga-based fuzzy neural controller design for municipal incinerators. *Fuzzy Sets and Systems*, 129:343–369, 2002.

[15] W.C. Chen and V.S. Vassiliadis. Solution of general nonlinear optimization problems using the penalty/modified barrier method with the use of exact hessians. *Computers and Chemical Engineering*, 27:501–525, 2003.

[16] D. Manca and M. Rovaglio. Numerical modeling of a discontinuous incineration process with on-line validation. *Ind. Eng. Chem. Res.*, 44:3159–3177, 2005.

[17] D. Van Hessem and O. Bosgra. Stochastic closed-loop model predictive control of continuous nonlinear chemical processes. *Journal of Process Control*, 16:225–241, 2006.

[18] T. de Hoog. *Rational orthonormal bases and related transforms in linear systems modeling*. PhD thesis, Delft University of Technology, 2001.

[19] S. De Wolf and P.M.J. Van den Hof. Identification of a pilot plant crystallization process with output error methods. In *Proc.* 9^{th} *IFAC/IFORS Symp. Identification and System Parameter Estimation*, pages 1208–1213, Budapest, Hungary, 1991.

[20] T.F. Edgar, D.M. Himmelblau, and L.S. Lasdon. *Optimization of chemical processes. Second Edition*. McGraw-Hill, 2001.

[21] R.A. Eek, J.A. Both, and P.M.J. Van den Hof. Closed-loop identification of a continuous crystallization process. *AIChE Journal*, 42:767–776, 1996.

[22] R. El Asri and D. Baxter. Process control in municipal solid waste incinerators: survey and assessment. *Waste Management and Research*, 22:177–185, 2004.

[23] E.N. Pistikopoulos, M.C. Georgiadis, and V. Dua (Eds.). *Multi-Parametric Model-Based Control - Theory and Applications*. Wiley-VCH, Weinheim, Germany, 2007.

[24] G. Evensen. The ensemble kalman filter for combined state and paramater estimation - monte carlo techniques for data assimilation in large systems. *IEEE Control Systems Magazine*, 29:83–104, 2009.

[25] F. Allgöwer, R. Findeisen, and Z.K. Nagy. Nonlinear model predictive control: From theory to application. *J. Chin. Inst. Chem. Engrs.*, 35:299–315, 2004.

[26] R. Fletcher. *Practical Methods of Optimization, Second Edition*. John Wiley & Sons, 1981.

[27] R. Fletcher. An l_1 penalty method for nonlinear constraints. In *Numerical Optimization 1984* (P.T. Boggs, R.H. Byrd and R.B. Schnabel (Eds.)), SIAM Publications, Philadelphia, pages 26–40, 1984.

[28] U. Forssell and L. Ljung. Closed-loop identification revisited. *Automatica*, 35:1215–1241, 1999.

[29] S. Fujii, S. Tomiyama, Y. Nogami, M. Shirai, H. Ase, and T. Yokoyama. Fuzzy combustion control for reducing both co and no_x from flue gas of refuse incineration furnace. *JSME Int. Journal*, 40, No. 2:279–284, 1997.

[30] Y. R. Goh, R. G. Siddall, V. Nasserzadeh, R. Zakaria, J. Swithenbank, D. Lawrence, N. Garrod, and B. Jones. Mathematical modelling of the burning bed of a waste incinerator. *Journal of the Institute of Energy*, 71:110–118, 1998.

[31] Y.R. Goh, C.N. Lim, R. Zakaria, K.H. Chan, G. Reynolds, Y.B. Yang, R.G. Siddall, V. Nasserzadeh, and J. Swithenbank. Mixing, modelling and measurements of incinerator bed combustion. *Trans. IChemE, part B*, 78:21–32, 2000.

[32] R. Gort. *On the Propagation of a Reaction Front in a Packed Bed*. PhD thesis, Twente University, 1995.

[33] S.P. Han. Superlinearly convergent variable metric algorithms for general nonlinear programming problems. *Mathematical Programming*, 11:263–282, 1976.

[34] S.P. Han. A globally convergent method for nonlinear programming. *Journal of Optimization Theory and Applications*, 22:297–309, 1977.

[35] W.F.M. Hesseling and P.L.F. Rademakers. Rendementsverhoging bij AVI's: Evaluatie van ervaringen met ketelcorrosie en corrosiebestrijding (in Dutch). Technical Report R2003/062, TNO-MEP, 2003.

[36] B.L. Ho and R.E. Kalman. Effective construction of linear, state-variable models from input/output functions. *Regelungstechnik*, 14:545–548, 1966.

[37] A.H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, 1970.

[38] S. Julier, J. Uhlmann, and H.F. Durrant-Whyte. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Trans. on Automatic Control*, 45:477–482, 2000.

[39] R.E. Kalman. A new approach to linear filtering and prediction problems. *Trans. ASME, Journal of Basic Engineering*, 82 (Series D):35–45, 1960.

[40] S. Kameswaran and L.T. Biegler. Simultaneous dynamic optimization strategies: Recent advances and challenges. *Computers and Chemical Engineering*, 30:1560–1575, 2006.

[41] L.B.M. Van Kessel. *Stochastic Disturbances and Dynamics of Thermal Processes with application to municipal solid waste combustion*. PhD thesis, Eindhoven University of Technology, 2003.

[42] L.B.M. Van Kessel and G. Brem. A dynamic model of a municipal solid waste incinerator as a tool for process optimization and operator training. In *VDI Berichte 1192: Verbennung und Feuerungen, 16. Deutscher Flammentag*, Düsseldorf, 1995. VDI-Verlag.

[43] L.B.M. Van Kessel and M. Leskens. System for continuous thermal combustion of matter, such as waste, 2002. European Patent Application EP1218669.

[44] J.D. Kilgroe. Control of dioxin, furan, and mercury emissions from municipal waste combustors. *Journal of Hazardous Materials*, 47:163–194, 1996.

[45] T. Klasen, K. Görner, W. Auel, and B. Sudau. Optimierung des Müllheizkraftwerkes (MHW) Bremen auf Basis von CFD-Simulationen und Betriebserfahrungen nach dem Umbau. *VGB PowerTech - International Journal for Electricity and Heat Generation*, 86:45–48, 2006.

[46] B. Krause, C. von Altrock, K. Limper, and W. Schäfers. A neuro-fuzzy adaptive control strategy for refuse incineration plants. *Fuzzy Sets and Systems*, 63:329–338, 1994.

[47] E. Kreyszig. *Introductory functional analysis with applications*. John Wiley & Sons, 1978.

[48] C.S. Krivanek III. Mercury control technologies for mwc's: The unanswered questions. *Journal of Hazardous Materials*, 47:119–136, 1996.

[49] S. Kung. A new identification and model reduction algorithm via singular value decomposition. In *Proc.* 12^{th} *Ann. Asilomar Conf. Circuits Syst. Comp.*, pages 705–714, Pacific Grove, CA, USA, 1979.

[50] J. Landlust. Dynamic optimization of process systems - a literature survey, 2004. Systems and Control Group, Faculty of Mechanical Engineering, Delft University of Technology, Delft, the Netherlands.

[51] L.B.M. van Kessel, M. Leskens, and G. Brem. On-line calorific value sensor and validation of dynamic models applied to municipal solid waste combustion. *Process Safety and Environmental Protection (Trans. IChemE, Part B)*, 80:245–255, 2002.

[52] J.R. Leis and M.A. Kramer. Sensitivity analysis of systems of differential and algebraic equations. *Computers and Chemical Engineering*, 9:93–96, 1985.

[53] M. Leskens. Multivariabele systeemidentificatie van industriële processen (in Dutch), 1997. Literature survey, report S-787, Faculty of Mechanical Engineering, Delft University of Technology.

[54] M. Leskens. Nonlinear model predictive control of MSW combustion plant AGEA - Results ECOTHERM project for deliverable 7a. Technical Report I&T-A R 2005/074, TNO Science and Industry, 2005.

[55] M. Leskens, L.B.M. van Kessel, and O.H. Bosgra. Model predictive control as a tool for improving the process operation of MSW combustion plants. *Waste Management*, 25:788–798, 2005.

[56] M. Leskens, L.B.M. Van Kessel, and P.M.J. Van den Hof. MIMO closed-loop identification of an MSW incinerator. *Control Engineering Practice*, 10:315–326, 2002.

[57] M. Leskens, L.B.M. van Kessel, P.M.J. Van den Hof, and O.H. Bosgra. Nonlinear model predictive control with moving horizon state and disturbance estimation - with application to MSW combustion. In *Proceedings of the 16th IFAC World Congress*, Prague, Czech Republic, 2005.

[58] M. Leskens and P.M.J. Van den Hof. Closed-loop identification of multivariable processes with part of the inputs controlled. *International Journal of Control*, 80:1552–1561, 2007.

[59] M. Leskens, P.P. van 't Veen, R. van Kessel, O.H. Bosgra, and P.M.J. Van den Hof. Improved economic operation of mswc plants with a new model based pid control strategy. In *Proc. IFAC 9th Intern. Symp. on Dynamics and Control of Process Systems*, pages 641–646, Leuven, Belgium, 2010.

[60] M. Leskens and R. van der Linden. Evaluation of MPC technology at AVR - Results NextGenBioWaste project for deliverable D 2.1.5. Technical report, Netherlands Organisation for Applied Scientific Research / TNO, 2008.

[61] M. Leskens, R.J.P. van der Linden, L.B.M. van Kessel, O.H. Bosgra, and P.M.J. Van den Hof. Nonlinear model predictive control of municipal solid waste combustion plants. In *Proc. Intern. Workshop on Assessment and Future Directions of Nonlinear Model Predictive Control*, Pavia, Italy, 2008.

[62] K. Liu. Identification of multi-input and multi-output systems by observability range space extraction. In *Proc.* 31^{st} *IEEE Conf. Decision and Control*, pages 915–920, Tucson, Arizona, USA, 1992.

[63] L. Ljung. *System identification: Theory for the user*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[64] L. Ljung. *System identification: Theory for the user (Second Edition)*. Upper Saddle River, NJ: Prentice-Hall, 1999.

[65] J. Ludlage, R. Oudbier, and T. Backx. Use of a priori knowledge of static gains in MIMO process identification. In *Proc.* 9^{th} *IFAC/IFORS Symp. Identification and System Parameter Estimation*, pages 1196–1201, Budapest, Hungary, 1991.

[66] M. Gevers, L. Ljung, and P.M.J. Van den Hof. Asymptotic variance expressions for closed-loop identification. *Automatica*, 37, no.5:781–786, 2001.

[67] D. Manca, M. Rovaglio, G. Pazzaglia, and G. Serafini. Inverse response compensation and control optimization of incineration plants with energy production. *Computers and Chemical Engineering*, 22:1879–1896, 1998.

[68] D.Q. Mayne, J.B. Rawlings, C.V. Rao, and P.O.M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36:789–814, 2000.

[69] K. Millrath and N. J. Themelis. Waste as a renewable source of energy: Current and future practices. In *Proceedings of the 2003 ASME International Mechanical Engineering Congress & Exposition*, Washington D.C., USA, November 16-21, 2003. IMECE 2003-55258.

[70] K.R. Muske and T.A. Badgwell. Disturbance modeling for offset-free linear model predictive control. *Journal of Process Control*, 12:617–632, 2002.

[71] S.G. Nash and A. Sofer. *Linear and Nonlinear Programming*. McGraw-Hill, 1996.

[72] G. De Nicolao, L. Magni, and R. Scattolini. Stability and robustness of nonlinear receding horizon control. In *Progress in Systems and Control Theory, Vol. 26, Birkhäuser Verlag (Eds.: F. Allgöwer and A. Zheng)*, pages 3–22, Basel, Switzerland, 2000.

[73] H.P. Nielsen, F.J. Frandsen, K. Dam-Johansen, and L.L. Baxter. The implications of chlorine-associated corrosion on the operation of biomass-fired boilers. *Prog. Energy Combust. Sci.*, 26:283–298, 2000.

[74] H. Ono, T. Ohnishi, and Y. Terada. Combustion control of refuse incineration plant by fuzzy logic. *Fuzzy Sets and Systems*, 32:193–206, 1989.

[75] Stichting Bureau Afval Overleg Orgaan. De afvalmarkt: structuur en ontwikkelingen (in dutch. transl.: The waste market: structure and developments, 2000. AOO report nr. 2000-02.

[76] G. Pannocchia. Robust disturbance modeling for model predictive control with application to multivariable ill-conditioned processes. *Journal of Process Control*, 13:693–701, 2003.

[77] G. Pannocchia and J.B. Rawlings. Disturbance models for offset-free model-predictive control. *AIChE Journal*, 49:426–437, 2003.

[78] K.M. Passino and S. Yurkovich. *Fuzzy Control*. Addison-Wesley, 1998.

[79] R.D. Pickens. Add-on control techniques for nitrogen oxide emissions during municipal solid waste combustion. *Journal of Hazardous Materials*, 47:195–204, 1996.

[80] M.J.D. Powell. A fast algorithm for nonlinearly constrained optimization calculations. In *Numerical Analysis Dundee, 1977* (G.A. Watson, ed.), Springer-Verlag, Berlin, pages 144–157, 1977.

[81] M.J.D.. Powell. Algorithms for nonlinear constraints that use Lagrangian functions. *Mathematical Programming*, 14:224–248, 1978.

[82] M.J.D. Powell. The convergence of variable metric methods for nonlinearly constrained optimization calculations. In *Nonlinear Programming 3* (O. Mangasarian, R. Meyer and S. Robinson, eds.), Academic Press, New York, pages 27–64, 1978.

[83] S.J. Qin and T.A. Badgwell. A survey of industrial model predictive control technology. *Control Engineering Practice*, 11:733–764, 2003.

[84] C.V. Rao and J.B. Rawlings. Constrained process monitoring: Moving-horizon approach. *AIChE Journal*, 48:97–108, 2002.

[85] C.V. Rao, J.B. Rawlings, and D.Q. Mayne. Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *IEEE Transactions on Automatic Control*, 48:246–258, 2003.

[86] L.P.M. Rijpkema. Rookgasreiniging gevudo (in dutch). Technical Report 94-284, TNO Environment, Energy and Process Innovation, 1994.

[87] D.G. Robertson and J.H. Lee. A least squares formulation of state estimation. *Journal of Process Control*, 5:291–299, 1995.

[88] O. Rosen and R. Luus. Evaluation of gradients for piecewise constant optimal control. *Computers and Chemical Engineering*, 15:273–281, 1991.

[89] M. Rovaglio, D. Manca, and F. Rusconi. Supervisory control of a selective catalytic reactor for NOx removal in incineration plants. *Waste Management*, 18:525–538, 1998.

[90] M. Rovaglio, D. Manca, and G. Biardi. Dynamic modeling of waste incineration plants with rotary kilns: Comparisons between experimental and simulation data. *Chemical Engineering Science*, 53:2727–2742, 1998.

[91] L.P. Russo and R.E. Young. Moving-horizon state estimation applied to an industrial polymerization process. In *Proc. 1999 American Control Conference*, San Diego, CA, USA, 1999.

[92] L.A. Ruth. Energy from municipal solid waste: a comparison with coal combustion technology. *Prog. Energy Combust. Sci.*, 24:545–564, 1998.

[93] S. Skogestad and I. Postlethwaite. *Multivariable Feedback Control - Analysis and Design*. John Wiley & Sons Ltd., Chichester, England, 1996.

[94] R.W.H. Sargent. Optimal control. *Journal of Computational and Applied Mathematics*, 124:361–371, 2000.

[95] K. Shen, J. Lu, Z. Li, and G. Liu. An adaptive fuzzy approach for the incineration temperature control process. *Fuel*, 84:1144–1150, 2005.

[96] D. Shin, J.H. Choi, V. Nasserzadeh, S. Choi, and j. Swithenbank. Experimental measurements and computational modelling of a pilot scale incinerator. *Journal of the Institute of Energy*, 72:56–63, 1999.

[97] T. Söderström and P. Stoica. *System Identification*. Englewood Cliffs, NJ: Prentice Hall, 1989.

[98] J. Swithenbank, V. Nasserzadeh, Y.R. Goh, and Y.B. Yang Et.al. Developments in incinerator bed modelling. In *Proceedings of the 3rd International Symposium on Incineration and Flue Gas Treatment Technologies*. IChemE, 2001.

[99] M.J. Tenny. *Computational Strategies for Nonlinear Model Predictive Control*. PhD thesis, University of Wisconsin-Madison, Madison, WI, USA, 2002.

[100] Inc. The MATHWORKS.

[101] P. Van den Hof. Closed-loop issues in system identification. *Annual Reviews in Control*, 22:173–186, 1998.

[102] P.M.J. Van den Hof and R.J.P. Schrama. An indirect method for transfer function estimation from closed loop data. *Automatica*, 29, no.6:1523–1527, 1993.

[103] E.T. van Donkelaar. *Improvement of efficiency in identification and model predictive control of industrial processes - A flexible linear parametrization approach*. PhD thesis, Delft University of Technology, Delft, the Netherlands, 2000.

[104] J.B. Van Helmont, A.J.J. Van der Weiden, and H. Anneveld. Design of an optimal controller for a coal fired benson boiler based on a modified approximate realisation algorithm. In *R. Whalley (Ed.), Application of multivariable system techniques (AMST90)*, pages 313–320. London: Elsevier Publishing Company, 1990.

[105] D. Van Hessem. *Stochastic inequality constrained closed-loop model predictive control*. PhD thesis, Delft University of Technology, 2004.

[106] L.B.M. Van Kessel, A.R.J. Arendsen, and G. Brem. On-line determination of the calorific value of solid fuels. *Fuel*, 83(1):59–71, 2004.

[107] S. van Loo and J. Koppejan (eds.). *Handbook of Biomass Combustion and Co-Firing*. Twente University Press, Enschede, 2002.

[108] P. Van Overschee and B. de Moor. N4sid: Subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica*, 30:75–93, 1994.

[109] P. Van Overschee and B. de Moor. *Subspace Identification for Linear Systems*. Kluwer Academic Publ., Dordrecht, The Netherlands, 1996.

[110] Various authors. *Handboek Kwaliteitsbeheersing Afvalverbranding (in Dutch. English: Handbook for quality control of MSW combustion)*. Veabrin, Rotterdam, 1990.

[111] J. Vehlow, B. Bergfeldt, R. Visser, and C. Wilen. European union waste management strategy and the importance of biogenic waste. *J. Mater. Cycles Waste Manag.*, 9:130–139, 2007.

[112] M. Verhaegen. A novel non-iterative mimo state-space model identification technique. In *Proc.* 9^{th} *IFAC/IFORS Symp. Identification and System parameter estimation*, pages 1453–1458, Budapest, Hungary, 1991.

[113] E.A. Wan and R. van der Merwe. The unscented kalman filter for nonlinear estimation. In *Proc. IEEE Symposium on Adaptive Systems for Signal Processing, Communications and Control*, pages 153–158, Lake Louise, Alberta, Canada, 2000.

[114] K.R. Westerterp, W.P.M. van Swaaij, and A.A.C.M. Beenackers. *Chemical Reactor Design and Operation*. John Wiley & Sons, 1984.

[115] R.B. Wilson. *A Simplical Algorithm for Concave Programming*. PhD thesis, Harvard University, 1963.

[116] Y.B. Yang, C.N. Lim, J. Goodfellow, V.N. Sharifi, and J. Swithenbank. A diffusion model for particle mixing in a packed bed of burning solids. *Fuel*, 84:213–225, 2005.

[117] Y.B. Yang, H. Yamauchi, V. Nasserzadeh, and J. Swithenbank. Effects of fuel devolatilisation on the combustion of wood chips and incineration of simulated municipal solid wastes in a packed bed. *Fuel*, 82:2205–2221, 2005.

[118] Y.B. Yang, J.B. Goodfellow, Y.R. Goh, V. Nasserzadeh, and J. Swithenbank. Investigation of channel formation due to random packing in a burning waste bed. *Trans. IChemE, part B*, 79:267–277, 2001.

[119] Y.B. Yang, Y.R. Goh, R. Zakaria, V. Nasserzadeh, and J. Swithenbank. Mathematical modelling of msw incineration on a travelling bed. *Waste Management*, 22:369–380, 2002.

[120] Y.C. Zhu, A.C.P.M. Backx, and P. Eykhoff. Multivariable process identification for robust control. Technical report, EUT Report 91-E-249. Faculty of Electrical Engineering, Eindhoven University of Technology, 1991.

[121] H.P. Zeiger and A.J. McEwen. Approximate linear realizations of given dimension via ho's algorithm. *IEEE Trans. Automat. Contr.*, 19:153, 1974.

[122] Y. Zhu. *Multivariable System Identification for Process Control*. Pergamon, 2001.

[123] Y. Zhu and T. Backx. *Identification of Multivariable Industrial Process for Simulation, Diagnosis and Control*. Springer-Verlag, 1993.

[124] Y. Zhu and X. Xu. A method of lpv model identification for control. In *Proceedings of the 17th IFAC World Congress*, pages 5018–5023, Seoul, Korea, 2008.

Summary

The combustion of municipal solid waste (MSW) is used for its inertisation, reduction of its volume and the conversion of its energy content into heat and/or electricity. Operation and control of modern large scale MSW combustion (MSWC) plants is determined by economic and environmental objectives and constraints that have to be fulfilled in the presence of large unmeasured disturbances resulting from a large variation in waste composition, which make the fulfillment of these objectives and constraints very difficult, and under varying market conditions and, thereby, economic priorities. Specifically, the operation and control objectives for MSWC plants are (i) revenue maximization, by maximization of the waste throughput and steam production, and (ii) minimization of process variations and fulfillment of constraints to minimize operational and maintenance costs and to fulfill environmentally motivated limits.

MSWC plant operators and managers are under an increasing pressure to operate economically more optimally, within the operating envelope determined by the environmental limits, due to the increasing business character of the environment they have to operate in, with market forces and competition increasingly dictating the plant operation. A direction with high potential for obtaining a more optimal economic MSWC plant performance is the application of model based combustion control, which has not been done so far at MSWC plants where PID type of combustion control is the standard. This potential is due to the fact that the combustion control system of an MSWC plant has a major influence on its overall economic performance and that model based control strategies allow for a systematic handling of the main characteristics of the MSWC plant combustion control problem, more specific the presence of constraints and the multivariable, interacting nature of the MSWC plant combustion dynamics. In particular the model based control strategy called Model Predictive Control (MPC) allows for this systematic handling and has potential for improvement.

Motivated by the observed need for an improved economic MSWC plant operation and the potential of model based combustion control for achieving this goal, the main research objective addressed in this thesis is to explore the opportunities of this type of control for improving the economic performance of these plants.

For the purpose of tackling this objective, three types of model based MSWC plant combustion control strategies are evaluated and compared: linear MPC, nonlinear MPC and a new PID type of combustion control strategy developed in this thesis. Additionally, modeling issues are addressed by exploring the opportunities of both first-principles modeling and linear empirical modeling *c.q.* system identification for obtaining a model suitable for model based MSWC plant combustion control, where for

the system identification part data are used that have been obtained from dedicated experiments performed at large scale MSWC plants.

The main conclusion from this thesis is that the economic performance of an MSWC plant can significantly be improved by means of model based combustion control, in particular by means of MPC based combustion control. More specific, usage of the latter type of combustion control particularly allows for an improved constraint handling of controlled variables not subject to setpoint tracking. This ability can be used to significantly reduce MSWC plant downtime and maintenance costs, *e.g.* by maintaining furnace temperatures below a certain maximum level to increase the lifetime of furnace components. MPC also allows for a significant reduction in process variability in MSWC plant combustion control applications even when constraints do not come into play, which can be used to decrease operational and maintenance costs and to operate the MSWC plant closer to the economically optimal operating point. A well performing state and disturbance estimator is crucial for obtaining this variability reduction. Additionally, the ability to include constraints in MPC together with the flexibility in formulating its optimal control problem allows for application of constraint pushing type of MSWC plant combustion control problems, which also allows for operating the MSWC plant closer to the economically optimal operating point.

In addition, it has been found that by means of the new model based PID-type of combustion control strategy developed in this thesis the setpoint tracking properties of commonly used PID combustion control strategies can be improved substantially, and thereby the overall economic performance of MSWC plants. However, to improve upon the process variability minimization properties of the currently used PID combustion control strategies, and thereby making a further improvement in economic performance, non-PID type of combustion control strategies are required.

Also, in order to obtain models suitable for model based MSWC plant combustion control, a new first-principles model suitable for that purpose has been developed that is an extended version of an available literature model. The extension involves the incorporation of the equations underlying the so called calorific value sensor (CVS), which is an on-line estimator of the composition and calorific value of the burning part of the waste. This incorporation leads to a more detailed description of the waste composition in the model which, combined with the ability to estimate the main parameter of this description from large scale MSWC plant plant data through the CVS, allows for an improved simulation, validation and model based combustion control performance compared to existing models.

Additionally, a specific system identification methodology has been developed to derive models suitable for model based MSWC plant combustion control. This methodology has been derived via identification of opportunities in the literature for overcoming all potential obstacles for arriving at such models through system identification.

Through application of first-principles modeling and system identification on data experimentally obtained from a large scale Dutch MSWC plant it was found that with both these modeling approaches it is possible to derive a model suitable for model based MSWC plant combustion control.

Another specific contribution of this thesis is an analysis of and solution approach to the occurrence of bias in case of estimating a model with the so called direct method on the basis of partial closed-loop identification data, which type of data may be experimentally obtained at MSWC plants. Within the prediction error framework, it has been shown that, whereas with the direct method (particular transfers in) multivariable plant models can be identified consistently without consistent noise modeling if the data are obtained in open-loop, in a general closed-loop situation this method looses this property already if one single control loop is present. The loss of this property, which also occurs at MSWC plants due to the specific experimental conditions at these plants, imposes restrictions on the model structure that may have significant negative consequences for the model variance, computation time and ability to tune the bias. The so called two stage method can be used to overcome these restrictions and consequences.

Also, validation of first-principles MSWC plant models directly on the basis of comparing simulated model outputs with their measured counterparts is generally impossible due to the presence of large nonmeasured disturbances on MSWC plant data. With other validation methods not being directly available, a new first-principles MSWC plant model validation method based on system identification has been developed in this thesis to overcome this validation problem.

Other main contributions of this thesis are both a linear and nonlinear MPC strategy for properly tackling MSWC plant combustion control problems. These strategies follow a standard moving horizon control strategy and employ a specific state and disturbance estimator developed in this thesis.

Samenvatting

De verbranding van huishoudelijk afval wordt gebruikt om het inert te maken, het volume ervan te verkleinen en om de energie-inhoud ervan in warmte en/of electrictieit om te zetten. De bedrijfsvoering en regeling van moderne grootschalige afvalverbrandingsinstallaties (AVIs) wordt bepaald door economische en milieu gerelateerde doelstellingen en begrenzingen die moeten worden vervuld in de aanwezigheid van grote ongemeten vestoringen, welke het gevolg zijn van een grote variatie in afvalsamenstelling en welke de vervulling van deze doelstellingen en begrenzingen erg moeiljk maken, en onder variërende martkcondities en, daardoor, economische prioriteiten. Specifiek, de bedrijfsvoerings- en regeldoelstellingen voor AVIs zijn (i) maximalisatie van de inkomsten, door middel van maximalisatie van de afvaldoorvoer en stoomproductie, en (ii) minimalisatie van procesvariaties en vervulling van begrenzingen om operationele en onderhoudskosten te minimaliseren en om aan de milieu-eisen te voldoen.

Operators en managers van AVIs staan onder een toenemende druk om economisch optimaler te opereren, binnen de door milieu-eisen bepaalde begrenzingen, door het toenemende zakelijke karakter van de omgeving waarin zij werken, waarin marktkrachten en concurrentie in toenemende mate de bedrijfsvoering bepalen. Een richting voor AVIs met grote mogelijkheden voor het verkrijgen van een economisch optimalere bedrijfsvoering is de toepassing van *model gebaseerde regeling van het verbrandingsproces*, waar tot nu toe nog geen gebruik van is gemaakt door AVIs waar PID type verbrandingsregelingen de standaard zijn. Deze mogelijkheden zijn het gevolg van het feit dat de verbrandingsregeling van een AVI een grote invloed heeft op de economische prestatie van de AVI als geheel en dat model gebaseerde regelstrategieën systematisch om kunnen gaan met de hoofdkenmerken van het afvalverbrandingsregelprobleem, meer specifiek de aanwezigheid van begrenzingen en het multi-variabele karakter van de verbrandingsdynamica van een AVI. In het bijzonder is de model gebaseerde regelstrategie genaamd Model Predictive Control (MPC) hiertoe in staat en kan toepassing van deze regelstrategie tot verbetering leiden.

Gemotiveerd door de waargenomen behoefte aan een verbeterde economische bedrijfsvoering van AVIs en de mogelijkheden van model gebaseerde regeling van het verbrandingsproces voor het bereiken van dat doel is de belangrijkste onderzoeksdoelstelling van dit proefschrift het onderzoeken van de mogelijheden van dit type regeling voor het verbeteren van de economische prestaties van AVIs.

Ten behoeve van de aanpak van deze doelstelling zijn drie model gebaseerde verbrandingsregelstrategieën geevalueerd en vergeleken: lineaire MPC, niet-lineaire MPC en een nieuwe PID type verbrandingsregelstrategie die ontwikkeld is in dit proefschrift. Daarnaast zijn er modelleringszaken onderzocht via het bekijken van de mogelijkheden van zowel fysich-chemische modellering als ook lineaire empirische modellering *c.q.* systeemidentificatie voor het verkrijgen van modellen die geschikt zijn voor model gebaseerde regeling van het AVI verbrandingsproces, waarbij voor dit laatste deel data zijn gebruikt die verkregen zijn via gerichte experimenten uitgevoerd op grootschalige AVIs.

De belangrijkste conclusie van dit proefschrift is dat de economische prestatie van een AVI significant kan worden verbeterd door middel van model gebaseerde regeling van het verbrandingsproces, in het bijzonder door middel van MPC gebaseerde regeling van dit proces. Meer specifiek kan het gebruik van dit laatste type verbrandingsregeling met name leiden tot een verbeterde vervulling van limieten op procesvariabelen die geen gewenste/setpoint waarde hoeven te volgen. Dit vermogen kan worden gebruikt om uitvaltijd gerelateerde kosten en onderhoudskosten significant te verlagen, bijv. door oventemperaturen beneden een zekere maximale waarde te houden om daarmee de levensduur van de componenten van deze oven te verhogen. Zelfs als limieten op procesvariabelen geen rol van betekenis spelen kan met MPC een significante reductie in procesvariabiliteit worden verkregen in AVI verbrandingsregelingstoepassingen. Deze mogelijkheid kan worden gebruikt om operationele en onderhoudskosten te verlagen en om de AVI dichter tegen zijn/haar economisch optimale werkpunt aan te bedrijven. Een goed werkende toestands- en verstoringsschatter is cruciaal voor het verkrijgen van deze procesvariabiliteitsreductie. Daarnaast kan het vermogen om procesbegrenzingen in MPC mee te nemen in combinatie met de flexibiliteit die MPC biedt in het formuleren van zijn/haar regelprobleem gebruikt worden voor de toepassing van zogenaamde *constraint pushing* type AVI verbrandingsregelproblemen, wat ook kan worden gebruikt voor het dichter tegen het economische optimale werkpunt aan bedrijven van de AVI.

Daarnaast is gevonden dat door middel van de nieuwe in dit proefschrift ontwikkelde PID type AVI verbrandingsregelstrategie de referentie waarde /setpoint volgende eigenschappen van de nu toegepaste PID type verbrandingsregelstrategieën substantieel kunnen worden verbeterd, en daarmee de economische prestaties van AVIs als geheel. Echter, om de eigenschappen van de huidige PID type verbrandingsregelstrategieën ten aanzien van reductie van de procesvariabiliteit te kunnen verbeteren, en daarmee een verdere verbetering in economische prestatie te bewerkstelligen, zijn regelstrategieën vereist die niet van het PID type zijn.

Verder, met als doel het verkrijgen van modellen die geschikt zijn voor model gebaseerde regeling van het AVI verbrandingsproces, is er een nieuw fysisch-chemisch model ontwikkeld dat geschikt is voor dat doel en welke een uitgebreidere versie is van een model dat al beschikbaar is in the literatuur. De uitbreiding heeft betrekking op het invoegen van de vergelijkingen die ten grondslag liggen aan de zogenaamde stookwaardesensor, wat een on-line schatter is van de samenstelling en stookwaarde van het brandende deel van het afval. Deze invoeging leidt tot een gedetailleerdere beschrijving van de afvalsamenstelling in het model wat, gecombineerd met de mogelijkheid tot het schatten van de belangrijkste parameter van deze beschrijving met de stookwaardesensor op basis van gemeten AVI data, gebruikt kan worden om de simulatie, validatie- en model gebaseerde verbrandingsregelingsprestatie te verbeteren in vergelijking met bestaande modellen.

Daarnaast is er een specifieke systeemidentificatiemethodologie ontwikkeld voor het verkrijgen van modellen die geschikt zijn voor model gebaseerde regeling van het AVI verbrandingsproces. Deze methodologie is afgeleid via identificatie in de literatuur van mogelijkheden voor het overwinnen van potentiële obstakels voor het verkrijgen van zulke modellen met systeemidentificatie.

Via toepassing van zowel fysisch-chemische modellering als systeemidentificatie op data die experimenteel verkregen zijn bij een grootschalige AVI, is gevonden dat het met beide modelleringsmethoden mogelijk is om een model te verkrijgen dat geschikt is voor model gebaseerde regeling van het AVI afvalverbrandingsproces.

Een andere bijdrage van dit proefschrift is een analyse van, en oplossingsbenadering voor, de aanwezigheid van bias (systematische modelafwijkingen) in het geval van het schatten van een model met de zogenaamde direct methode op basis van partieel gesloten-lus data, welke type data experimenteel verkregen kan worden bij AVIs. Binnen het prediction error raamwerk is aangetoond dat, terwijl met de directe methode (specifieke overdrachten in) multi-variabele procesmodellen consistent kunnen worden geidentificeerd zonder consistente ruismodellering als de data verkregen zijn in open-lus, deze methode deze eigenschap in een algemene gesloten-lus situatie al verliest als een enkelvoudige regelkring gesloten is. Het verlies van deze eigenschap, wat ook het geval is bij AVIs ten gevolge van de specifieke experimentele condities bij deze installaties, leidt tot restricties op de modelstructuur welke significant negatieve gevolgen kunnen hebben voor de modelvariantie (toevallige modelafwijkingen), rekentijd en de mogelijkheden tot het afstemmen van de bias. De zogenaamde tweestapsmethode kan worden gebruikt om deze restricties en negatieve gevolgen te ontwijken.

Tevens is gevonden dat validatie van fysisch-chemische modellen van het AVI verbrandingsprocess direct op basis van het vergelijken van gesimuleerde modeluitgangen met hun gemeten versies over het algemeen onmogelijk is dankzij de aanwezigheid van grote ongemeten verstoringen op de AVI data. Omdat andere validatiemethoden niet direct voorhanden zijn is er een nieuwe validatiemethode ontwikkeld in dit proefschrift om dit probleem op te lossen welke gebaseerd is op systeemidentificatie.

Andere substantiële bijdragen van dit proefschrift zijn zowel een lineaire als een niet-lineaire MPC strategie voor de aanpak van AVI verbrandingsregelproblemen. Deze strategieën volgen een standaard *moving horizon* regelstrategie en maken gebruik van een specifieke toestands- en verstoringsschatter die ontwikkeld is in dit proefschrift.

Curriculum Vitae

Martijn Leskens was born on the 8^{th} of December 1971 in The Hague, The Netherlands. In 1990 he finished his pre-university education (VWO) at the Stedelijk Gymnasium in Breda, The Netherlands. He subsequently studied for one year Industrial Design at the Delft University of Technology where he received the degree of Propedeuse. In 1991 he changed faculty at this university to study Mechanical Engineering, for which he received in 1998 the degree of Master of Science with specialization Measurement and Control. In that same year he started working for the Netherlands Organisation for Applied Scientific Research TNO in Apeldoorn, The Netherlands, at the TNO Environment, Energy and Process Innovation institute. In May 1999 he started to pursue (part-time) a Ph.D. degree in systems and control at the Delft University of Technology on the subject of model based control of municipal solid waste incinerators, of which the results are described in this thesis. This project was initiated and sponsored by TNO Environment, Energy and Process Innovation and was supervised by Prof. Paul Van den Hof, Prof. Okko Bosgra and Dr. Robert van Kessel. Until 2003 it was performed at the Mechanical Engineering and Applied Physics Systems and Control groups of the Delft University of Technology and, after that, at the Delft Center for Systems and Control (DCSC), also at the Delft University of Technology. Having worked for 10 years at TNO on pre-dominantly model based control applications in the fields of waste combustion and upstream oil and gas, Martijn started in 2009 as a post-doctoral researcher at DCSC on the subject of model based control of compact gas/liquid separators. After having finished this project in 2011 he started working for FMC Separation Systems, part of FMC Technologies, in Arnhem, The Netherlands, as an R&D Technologist, where he still works. During his career, Martijn has followed numerous courses on process control and engineering related subjects, both with an academic and industrial focus, including DISC courses and an industrial course on model predictive control.