Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace

authored by **S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi**

and published in *Control Engineering Practice*.

Contact:

Automation and Control Institute (ACIN)
TU Wien
Gusshausstrasse 27-29/E376
1040 Vienna, Austria

Internet: www.acin.tuwien.ac.at
E-mail: office@acin.tuwien.ac.at
Phone: +43 1 58801 37601
Fax: +43 1 58801 37699

ACIN

Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace

S. Strommer[a,*], M. Niederer[a], A. Steinboeck[b], A. Kugi[a,b]

[a]*AIT Austrian Institute of Technology, Center for Vision, Automation and Control, Argentinierstrasse 2/4, 1040 Vienna, Austria*
[b]*Automation and Control Institute, Technische Universität Wien, Gusshausstrasse 27–29, 1040 Vienna, Austria*

Abstract

Continuous strip annealing furnaces are complex multi-input multi-output nonlinear distributed-parameter systems. They are used in industry for heat treatment of steel strips. The product portfolio and different materials to be heat-treated is steadily increasing and the demands on high throughput, minimum energy consumption, and minimum waste have gained importance over the last years. Designing a furnace control concept that ensures accurate temperature tracking under consideration of all input and state constraints in transient operations is a challenging task, in particular in view of the large thermal inertia of the furnace compared to the strip. The control problem at hand becomes even more complicated because the burners in the different heating zones of the considered furnace can be individually switched on and off. In this paper, a real-time capable optimization-based hierarchical control concept is developed, which consists of a static optimization for the selection of an operating point for each strip, a trajectory generator for the strip velocity, a dynamic optimization routine using a long prediction horizon to plan reference trajectories for the strip temperature as well as switching times for heating zones, and a nonlinear model predictive controller with a short prediction horizon for temperature tracking. The mass flows of fuel and the strip velocity are the basic control inputs. The underlying optimization problems are transformed to unconstrained problems and solved by the Gauss-Newton method. The performance of the proposed control concept is demonstrated by an experimentally validated simulation model of a continuous strip annealing furnace at voestalpine Stahl GmbH, Linz, Austria.

Keywords: Nonlinear model predictive control, direct- and indirect-fired strip annealing furnace, reheating and heat treatment of metal strips, nonlinear MIMO system, unconstrained optimization, Gauss-Newton method

1. Introduction

Continuous annealing processes are used for heat treatment of steel strips. Controllers should ensure that the strip temperature follows a set-point temperature trajectory as closely as possible. The set-point trajectories depend on metallurgical requirements and may vary from strip to strip. Typically, a diverse portfolio of products is heat-treated in continuous strip annealing furnaces (CSAF). Therefore, a variety of different CSAF can be found in industry (Mullinger and Jenkins, 2014; Imose, 1985). The CSAF considered in this paper is part of a hot-dip galvanizing line of voestalpine Stahl GmbH, Linz, Austria.

The accurate temperature control of a CSAF is essential to ensure a high product quality. This is in particular challenging in transient operational situations when a welded joint moves through the furnace. In this case, the strip dimensions (thickness, width), the steel grade, the set-point strip temperature, and the strip velocity may change. Since the strip temperature is a distributed process variable, which can only be measured at a very few discrete points, the control task is further complicated. Moreover, the thermal inertia of the furnace is rather high compared to that of the strip. Thus, the response time of the furnace is also high compared to the processing time of a strip.

The CSAF constitutes a multiple-input multiple-output nonlinear distributed-parameter system. The main control inputs are the fuel supplies of the heating zones. They can be individually switched on/off depending on the required heat input, which makes the task of finding optimal control inputs a mixed-integer programming problem (Grossmann and Kravanja, 1997). The strip velocity serves as an additional control input. It is subject to several restrictions which are mainly defined by downstream process steps. All control inputs are bounded from above and below. In this work, a nonlinear optimization-based hierarchical control strategy for the considered CSAF of voestalpine Stahl GmbH is presented.

1.1. Continuous strip annealing furnace

Figure 1 shows a schematic of the considered CSAF, which consists of a direct- and an indirect-fired furnace

*Corresponding author. *Tel.:* +43 50550 6813, *fax:* +43 50550 2813.
Email address: stephan.strommer@ait.ac.at (S. Strommer)

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005

Nomenclature

Latin symbols

b	width
c	specific heat capacity
$\mathcal{D} = \{hza, hzb, hzc, hzd\}$	set of abbreviations for the heating zones a–d
d	thickness
$\mathbf{d}$	search direction
$\mathcal{F} = \{dff, rth, rts\}$	set of abbreviations for furnace sections
$\mathbf{g}$	gradient
$\dot{H}$	enthalpy flow
$\mathbf{H}$	Hessian
h	specific enthalpy
i	index
J	objective function
k	index with respect to time
$\mathbf{L}$	vector-valued Lagrange function
l	index
$\dot{M}$	mass flow
$\bar{M}$	molar mass
$\mathbf{m}$	vector of the mass flows of fuel
$\dot{\mathbf{m}}$	slope of the mass flows of fuel
N	number of discretized elements
$\dot{Q}$	heat flow
$\dot{q}$	heat flux
$\mathbf{R}$	objective function
S	surface
s	switching state
T	temperature
t	time
$\mathbf{U}$	system input
$\mathbf{u}$	optimization variables
v_s	strip velocity
$\dot{v}_s$	slope of the strip velocity
$\mathbf{W}$	weighting matrix
$\mathbf{x}$	algebraic and state variables
$\mathbf{y}$	system outputs
z	spatial coordinate

Greek symbols

ε	emissivity
$\mathbf{\Gamma}$	system dynamics
$\mathbf{\Lambda}$	Lagrange multiplier
λ	air-fuel equivalence ratio
φ	nonlinear transformation
ρ	mass density
τ	switching time
υ	unconstrained optimization variable
χ	stoichiometric coefficient

Subscripts

g	flue gas
h	roll
r	radiant tube
s	strip
w	wall

Superscripts

$+$	upper bound
$-$	lower bound
ad	adiabatic flame
a	air
cb	combustion
d	set point
f	fuel
in	incoming
out	outgoing
r	reference
sp	nitrogen flushing
$\hat{}$	observer

separated by an air lock. In the considered CSAF, standard steel for the automotive area (bodywork) is produced. The strip width varies from 0.8 m up to 1.8 m, whereas the strip thickness typically varies from 0.4 mm up to 1.2 mm due to the production needs but can accept materials having lower thickness. The steel strip, which is conveyed through the furnace by rolls, couples both parts. This furnace type is designed for heat treatment of steel strips in terms of throughput, energy consumption, and product quality (Imose, 1985).

In the direct-fired furnace (DFF), the strip is heated by means of hot flue gas, which comes from the combustion of fuel. The fuel is burnt fuel rich in the four heating zones (hz a–d) to avoid scale formation of the strip. Thus, the flue gas contains unburnt products, which are burnt in a post combustion chamber (PCC) by adding fresh air via an air intake. The flue gas leaving the PCC contains excess oxygen and streams into the preheater, where it is used to preheat the incoming strip. The DFF is a counterflow heat exchanger because the flue gas streams in the opposite direction of the strip motion.

In the heating zone a and b, an array of burners is used, where a defined number of burners can be deactivated depending on the width of the strip (narrow, middle, wide). These heating zones are responsible for the base load. Using shut-off valves, the fuel supply of the heating zones a–d can be individually switched on/off. Deactivated burners are flushed with cold nitrogen to protect the burner nozzles from thermal damage (Strommer et al., 2014b). The media supplies of air and fuel are coupled by the air-fuel

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005

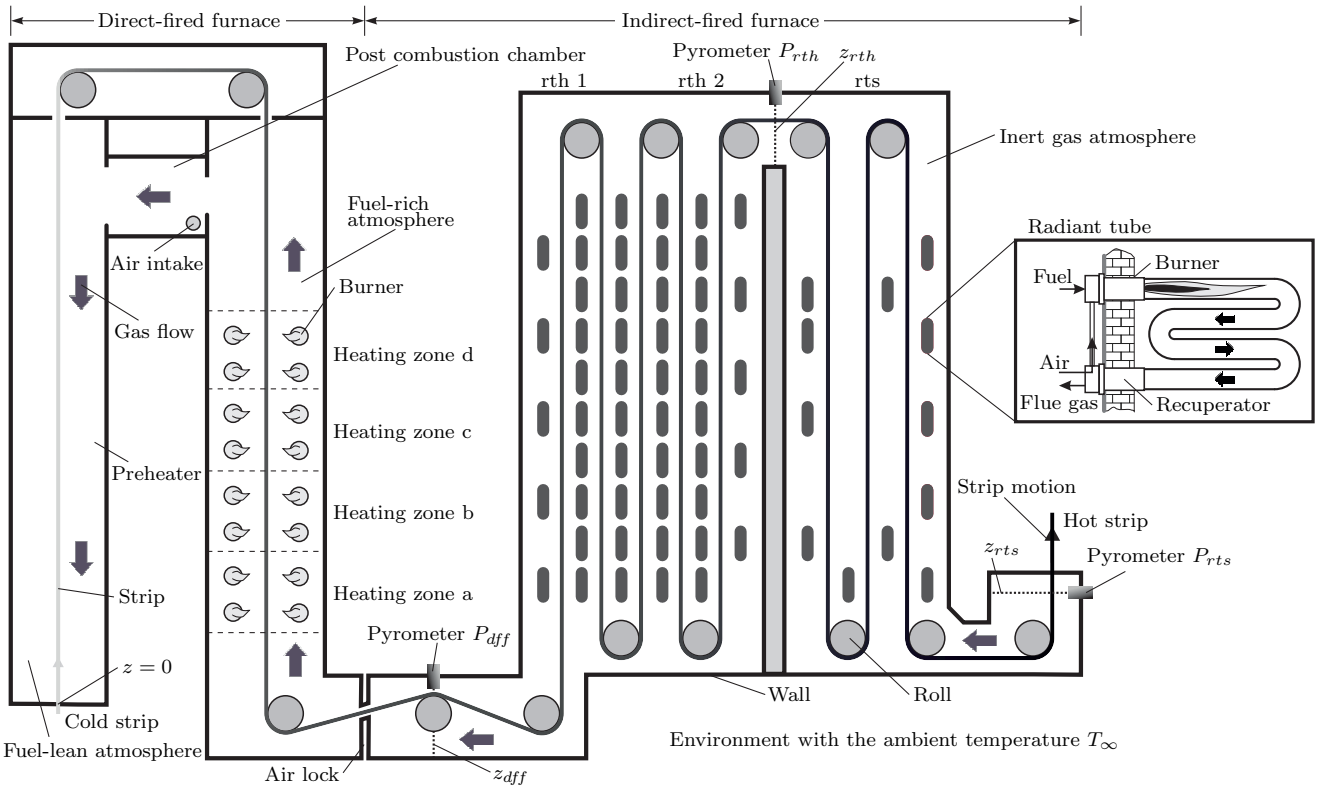

Figure 1: Combined direct- and indirect-fired strip annealing furnace.

equivalence ratio, which is controlled by a cross-limiting controller.

The indirect-fired furnace (IFF) is separated into three sections, the radiant tube heating sections 1 and 2 (rth 1 and 2) and the radiant tube soaking section (rts). Each of these sections is equipped with W-shaped radiant tubes and can be separately controlled. The tubes are permanently supplied with fuel and air to avoid flame extinction. Inside the IFF, an inert gas atmosphere is established to prevent scale formation of the strip. Due to a controlled pressure gradient, a gas flow in the direction of the DFF is always assured.

The strip temperature is measured by three pyrometers (P_{dff}, P_{rth}, and P_{rts}), see Fig. 1. Additionally, thermocouples, which measure several local flue gas temperatures, wall temperatures, and radiant tube temperatures, are installed for safety reasons.

1.2. Existing solutions

Control concepts that can be found in the literature significantly differ in their complexity and application. In CSAFs, simple PID controllers are widespread to control the heating of the strip. In (Dunoyer et al., 1998; Martineau et al., 2004), PID controllers based on mathematical models of the furnace are applied. The PID controller proposed by Li et al. (2004) is used to control the temperature of a workpiece inside a furnace. The parameters of this controller were determined by an optimization problem. In (Kelly et al., 1988), a nonlinear model of an IFF was introduced. Moreover, a Kalman filter and a linear quadratic controller were designed based on a linear system model.

To support the operators, mathematical models are employed to calculate optimal process parameters, e.g., the strip velocity and the target strip temperature, see (de Pisón et al., 2011; Yahiro et al., 1993). The operators use this information to adjust the process, in particular when a welded joint moves through the furnace. In (Bitschnau et al., 2010), a nonlinear model of an IFF was derived with a controller consisting of a feedback and a feedforward part.

In (Norberg, 1997), the challenges of controlling a CSAF are discussed in more detail, in particular the problems and the requirements with respect to the strip velocity, the combustion process, the gas atmosphere, and the scaling process are considered. As a conclusion, model predictive control is highly recommended to tackle all these issues.

To control the strip temperature in an IFF, real-time implementations of linear model predictive controllers (MPC) are given in (Bitschnau and Kozek, 2009; Lewis et al., 1994; Wu et al., 2014). For all these approaches, only the mass flows of fuel serve as control inputs, whereas the strip velocity and other process parameters are defined by the operator.

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005

In (Guo et al., 2009), a particle swarm optimization is proposed to ensure an accurate heating of the strip according to a desired trajectory. Here, a neural network was used for modeling the system.
In (Niederer et al., 2016), a nonlinear model predictive controller for the considered strip annealing furnace was designed, where the strip velocity can be arbitrarily varied within a permissible range. The control inputs are the mass flows of fuel and the strip velocity to realize an optimal tracking control of the strip temperature.
In furnace control, hierarchical control structures are quite common. In (Ming et al., 2008), a control structure with three layers is used. One layer adapts the weakly known parameters of the mathematical model based on measurements. The model is used in a second layer to determine a target trajectory for the strip temperature. Finally, in a third layer, a nonlinear MPC uses this target trajectory to calculate the optimal fuel supply. Based on a nonlinear model of an IFF, Ueda et al. (1991) solved an optimal control problem. The control inputs are the adjustment times of the strip velocity and the furnace temperature. A hierarchical structure is also used by Yoshitani and Hasegawa (1998). First, a reference speed is determined based on a static model. Then, the optimal time of the speed change is calculated. Furthermore, a target trajectory of the strip temperature is specified. Finally, a tracking controller (MPC) is used to determine the mass flows of fuel.

1.3. Motivation and objectives

Usually, CSAFs are not equipped with a DFF. However, the key advantage of a DFF compared to an IFF is its fast response characteristic. With a DFF, a nearly instantaneous heating of the strip can be achieved (Delaunay, 2007; Imose, 1985; Mould, 1982). Most of the existing control strategies only concern IFFs and are based on simple furnace models that capture the heaters and the strip. Although such models may simplify the overall control design, this approach often restricts the capabilities and accuracy of the controlled furnace.
A CSAF is a rather complex dynamical system, where typically a diverse portfolio of products having a notable variety of steel grades is processed. Moreover, the requirements in terms of product quality may vary from application to application. This is why a tailored control concept has to be developed to fully utilize the potential of the CSAF under consideration.
In this paper, the strip velocity is not allowed to vary in an arbitrary way due to downstream process steps. Moreover, the heating zones can be switched on/off. This is a substantial extension to the control strategy proposed by Niederer et al. (2016).
A control strategy based on predictive control seems promising for the considered task because the strips are known in advance (Norberg, 1997). Thus, the requirements and the parameters of these strips can be incorporated into the design of optimized control inputs at an early stage. Pure feedback control does not appear to be suitable for the considered control task because of rapidly changing operating conditions and the high thermal inertia of the CSAF.
All these aspects motivate the development of a new furnace temperature controller, which utilizes the advantages of MPC. Moreover, a control structure with several modules seems to be useful to tackle the challenges associated with mixed-integer programming, the selection of the strip velocity, the high thermal inertia of the CSAF, the computational effort, and the real-time requirements (Ming et al., 2008; Ueda et al., 1991; Yoshitani and Hasegawa, 1998).
The proposed furnace temperature controller consists of a nonlinear temperature regulator (TR), an optimization-based trajectory planner (OTP), a static optimization module, and a trajectory generator for the strip velocity. The static optimization chooses an optimal operating point characterized by the strip velocity and the switching state of the heating zones (on or off) for each strip. The optimal strip velocities are used by the trajectory generator for the strip velocity to design a desired trajectory, which is then utilized by the TR and the OTP. The OTP calculates an optimized reference trajectory of the strip temperature and optimized switching times for turning on/off the heating zones. The OTP uses a long time horizon (15 min) to take into account the high thermal inertia of the CSAF. The OTP is required to plan ahead for long-term transient changes of operating situations, e.g., when a new strip with other properties than the preceding strip enters the furnace. The TR is based on MPC technology, uses the fuel supplies as control inputs, and performs tracking control for the strip temperatures depending on their optimized reference trajectories obtained from the OTP. The TR uses a short time horizon (3 min). A detailed description of the choice of the time horizons is presented in Sec. 4.6.
The main control objectives are:

- Accurate heating of the strip
- Minimum scrap of material
- Maximum throughput of strip
- Minimum energy consumption and CO_2 emissions

Clearly, by minimizing the energy consumption also the CO_2 emissions are reduced. In the course of the controller design, a number of constraints have to be considered in terms of a reliable and safe operation of a hot-dip galvanizing line:

- The temperatures of the gas, the radiant tubes, and the walls inside the furnace are limited (damage and wear prevention).
- Limitations of control inputs have to be respected.
- The strip temperature is bounded.

The strip velocity has a significant influence on the evolution of the strip temperature. A change of the speed causes a much quicker response of the strip temperature

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005

compared to a change of the fuel supply (Yoshitani and Hasegawa, 1998). However, the speed is subject to several restrictions:

- A change of the strip velocity is programmed only in case a new strip enters the furnace. However, certain production exceptions might require sudden variations of the strip velocity, which may be considered as unforseen disturbances to be compensated by the temperature control loop.
- A change of the strip velocity should always be monotonous, i.e., the target acceleration should not change its sign during coil transition.
- The heat-treatment time of a strip section has to be limited due to annealing and material aging, which may reduce the product quality.
- There is a maximum and minimum speed, which are caused by metallurgical requirements and motor sizing.
- Speed limitations demanded by the operator or due to downstream process steps have to be respected.

A big difference between the roll and the strip temperature may cause mechanical tensions in the strip, which result in a distortion of the strip. Thus, the product quality suffers and in the worst case, scrap is produced. This phenomenon is called heat buckling (Paulus and Laval, 1985; Sasaki et al., 1984). Such a temperature difference follows from a rapid change of the strip velocity as well as a change of the strip width. A heating zone may be switched on/off depending on the required heat input. Since the switching of the burners may cause several disadvantages like reduced energy efficiency and a high temperature gradient inside the furnace, the switching should be restricted:

- Ensure a minimum time between two switching cycles.
- Switching of burners is only allowed in the vicinity of strip transitions.

There exist further control objectives and requirements which should be incorporated into an advanced model-based furnace control strategy:

- Consideration of recuperators, heat recirculation (Katsuki and Hasegawa, 1998).
- The mathematical model aims at predicting the process behavior in all operating conditions and for all product types.
- The control concept has to realize a fuel-rich gas atmosphere to ensure that the flue gas does not contain oxygen, which may cause undesirable scale formation (product quality).

1.4. Contents

This work has the following structure: In Section 2, the mathematical model of the CSAF is briefly summarized. The controller is based on a hierarchical structure, which is outlined in Section 3 together with the control tasks and objectives. Section 4 is devoted to the formulation and the numerical solution of the associated unconstrained nonlinear optimization problem. Finally in Section 5, the performance of the proposed control concept is demonstrated by simulation on an industrially validated model.

2. Mathematical model

The mathematical modeling of the overall process is a complex task due to the underlying nonlinear physical effects and the variety of products and materials yielding to a large range of operating conditions. Nonlinear models of the DFF and IFF were developed and validated by measurements in Strommer et al. (2014a) and Niederer et al. (2014), respectively. Moreover, a combined model of the DFF and IFF is presented in (Niederer et al., 2015). It serves as a simulation model, however, it is not suitable for control design due to its high system dimension and complexity. Therefore, a reduced model in terms of complexity, dimension, and computational effort will be used for control purposes. This model consists of the subsystems flue gas, radiant tube, roll, strip, and wall. The following reduction steps are performed:

- Coarser spatial discretization of the furnace
- Simplified calculation of the heat transfer coefficient
- Simplified combustion in the DFF
- Simplified model of the radiant tubes

The coarser spatial discretization significantly reduces the system dimension from 850 for the full simulation model to 150 for the reduced controller design model. The calculation of the heat transfer coefficient is based on empirical relations, which depend on the properties of the flue gas (Baehr and Stephan, 2006; Incropera et al., 2007). To determine these properties for a gaseous mixture, the formula suggested by Buddenberg and Wilke (1949) is typically applied. This formula is complex and causes a significant computational effort. Therefore, the gaseous mixture is only represented by nitrogen for calculating the heat transfer coefficient and hence, the complexity can be reduced substantially.

2.1. Fuel supplies

Generally, the fuel supply of the four heating zones a–d of the DFF can be switched off using shut-off valves and thus, $2^4 = 16$ different possibilities have to be considered. However, switching off the heating zones a and b is not useful because they cover the base load. Therefore, only the heating zones c and d are switched depending on the required heat input. Based on a thorough energy analysis (Strommer et al., 2013) and the fact that the furnace operates as a counterflow heat exchanger (Incropera et al., 2007), it can be shown that it is not useful to fire heating zone d if heating zone c is off. Thus, only three different cases can occur, which are indicated by the switching state $s \in \{1, 2, 3\}$, see Tab. 1. In case 1 ($s = 1$), all heating

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005

s	$\dot{M}^f_{hzc}$	$\dot{M}^f_{hzd}$
1	$\left[\dot{M}^{f,min}_{hzc}, \dot{M}^{f,max}_{hzc}\right]$	$\left[\dot{M}^{f,min}_{hzd}, \dot{M}^{f,max}_{hzd}\right]$
2	$\left[\dot{M}^{f,min}_{hzc}, \dot{M}^{f,max}_{hzc}\right]$	$\{0\}$
3	$\{0\}$	$\{0\}$

Table 1: Different cases and allowed ranges of the mass flows of fuel to the heating zones c and d.

s	$\dot{M}^{sp}_{hzc}$	$\dot{M}^{sp}_{hzd}$
1	0	0
2	0	$\dot{M}^{sp}_{hzd}$
3	$\dot{M}^{sp}_{hzc}$	$\dot{M}^{sp}_{hzd}$

Table 2: Different cases and the corresponding mass flows due to nitrogen flushing to the heating zones c and d.

zones of the DFF are active, in case 2 ($s = 2$), heating zone d is switched off, and in case 3 ($s = 3$), the heating zones c and d are switched off. The switching state s also defines the mass flows $\dot{M}^f_\alpha$ of fuel to the heating zones c and d, $\alpha \in \{hzc, hzd\}$, respectively. If a heating zone is switched off, it is flushed with a constant amount $\dot{M}^{sp}_\alpha$ of nitrogen, see Tab. 2. In the IFF, the radiant tubes are permanently supplied by fuel and air. Thus, the mass flow $\dot{M}^f_\alpha$ of fuel to an active heating zone of the DFF and to the IFF can vary in the range $\dot{M}^f_\alpha \in \left[\dot{M}^{f,min}_\alpha, \dot{M}^{f,max}_\alpha\right]$, $\alpha \in \mathcal{D} \cup \{rth1, rth2, rts\}$, with $\mathcal{D} = \{hza, hzb, hzc, hzd\}$ and the minimum and maximum mass flow $\dot{M}^{f,min}_\alpha$ and $\dot{M}^{f,max}_\alpha$ of fuel, respectively. The bounds $\dot{M}^{f,min}_\alpha$ and $\dot{M}^{f,max}_\alpha$, $\alpha \in \{hza, hzb\}$, depend on the width of the strip.

2.2. Flue gas in the DFF

The DFF is discretized into $N_g = 13$ volume zones, where each zone is considered as a well-stirred reactor with a uniform flue gas temperature $T_{g,i}$, $i = 1, \ldots, N_g$. The volume zones 2–5 correspond to the heating zones a–d and zone 12 represents the PCC. The flue gas is considered quasi-stationary due to its fast response characteristic compared to the remaining subsystems (Strommer et al., 2014a). It is assumed that the combustion occurs right at the burner nozzle and no further chemical reactions are considered within a volume zone. Furthermore, it is assumed that natural gas consists only of methane CH_4. The stationary combustion reaction in each heating zone reads as (Turns, 2006)

$$\begin{aligned} &CH_4 + 2\lambda_i\,(O_2 + 3.76N_2) \longrightarrow \chi_i^{CO_2}CO_2 + \chi_i^{CO}CO \\ &+ \chi_i^{H_2O}H_2O + \chi_i^{H_2}H_2 + \chi_i^{O_2}O_2 + \chi_i^{N_2}N_2 \end{aligned} \tag{1}$$

with $i \in \{2, 3, 4, 5\}$. Methane is oxidized into carbon dioxide CO_2, carbon monoxide CO, water H_2O, hydrogen H_2, oxygen O_2, and nitrogen N_2. These reaction products are summarized in the set $\mathcal{G} = \{CO_2, CO, H_2O, H_2, O_2, N_2\}$. The parameters χ_i^ν denote the stoichiometric coefficients of the component $\nu \in \mathcal{G}$ and depend on the air-fuel equivalence ratio λ_i. In the heating zones of the DFF, a fuel-rich combustion is realized, which implies $\lambda_i < 1$ and $\chi_i^{O_2} = 0$. The remaining parameters χ_i^ν, $\nu \in \mathcal{G} \setminus \{O_2\}$, are determined by mole balances and the equilibrium equation

$$\frac{\chi_i^{CO_2}\chi_i^{H_2}}{\chi_i^{CO}\chi_i^{H_2O}} = K_c(T_g^{ad}) \tag{2}$$

of the water-gas-shift reaction (Moe, 1962) with the equilibrium constant $K_c(T_g^{ad})$. In contrast to Strommer et al. (2014a), the adiabatic flame temperature $T_g^{ad} = \text{const.}$ is used in (2). This simplification is justified because $K_c(T_g^{ad}) \approx K_c(T_{g,i})$. Thus, the stoichiometric coefficients χ_i^ν, $\nu \in \mathcal{G}$, can be determined from mole balances, (1), and (2) independently of the flue gas temperature.
The mass flow $\dot{M}_i^{cb,\nu}$ of a combustion product $\nu \in \mathcal{G}$ which enters the furnace zone i can be calculated by

$$\dot{M}_i^{cb,\nu} = \frac{\bar{M}^\nu}{\bar{M}^{CH_4}}\chi_i^\nu \dot{M}_i^f,$$

with the mass flow $\dot{M}_i^f$ of fuel supplied to this zone and the molar mass $\bar{M}^\nu$ of the component $\nu \in \mathcal{B} = \mathcal{G} \cup \{CH_4\}$. The mass flow of combustion air to the zone i is given by $\dot{M}_i^a = \dot{M}_i^{a,O_2} + \dot{M}_i^{a,N_2}$ with

$$\dot{M}_i^{a,\kappa} = \frac{\bar{M}^\kappa}{\bar{M}^{CH_4}}\chi_i^{a,\kappa}\dot{M}_i^f,$$

$\kappa \in \mathcal{A} = \{O_2, N_2\}$, $\chi_i^{a,O_2} = 2\lambda_i$, and $\chi_i^{a,N_2} = 7.52\lambda_i$. The mass flow $\dot{M}_{12}^a = \dot{M}_{12}^{a,O_2} + \dot{M}_{12}^{a,N_2}$ of air supplied to the PCC reads as

$$\dot{M}_{12}^{a,\kappa} = \frac{\bar{M}^\kappa}{\bar{M}^{CH_4}}\sum_{i=2}^{5}\left(\chi_{12}^{a,\kappa} - \chi_i^{a,\kappa}\right)\dot{M}_i^f,$$

with $\kappa \in \mathcal{A}$, $\chi_{12}^{a,O_2} = 2\lambda_{12}$, $\chi_{12}^{a,N_2} = 7.52\lambda_{12}$, and the air-fuel equivalence ratio λ_{12} in the PCC.

Remark 1. To ensure that the flue gas which leaves the furnace contains a desired amount of excess oxygen, λ_{12} is feedback-controlled. This ensures a fuel-lean gas mixture, i.e., $\lambda_{12} > 1$ (Strommer et al., 2017).

The stationary mass balance of each component $\nu \in \mathcal{G}$ can be utilized to determine the outgoing mass flow $\dot{M}_i^{out,\nu}$ from an individual volume zone i (Baehr and Stephan, 2006; Incropera et al., 2007)

$$\dot{M}_i^{out,\nu} = \dot{M}_i^{in,\nu} + \dot{M}_i^{cb,\nu} + \dot{M}_i^{sp,\nu}, \tag{3}$$

where $\dot{M}_i^{in,\nu}$ is the incoming mass flow from the upstream zone, i.e., $\dot{M}_i^{in,\nu} = \dot{M}_{i-1}^{out,\nu}$. $\dot{M}_i^{sp,\nu}$ is the mass flow due to nitrogen flushing, i.e., $\dot{M}_i^{sp,\nu} = 0$ for $\nu \neq N_2$ and $\dot{M}_i^{sp,N_2} = \dot{M}_i^{sp}$, see Tab. 2.

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

The flue gas temperature $T_{g,i}$ of volume zone i is calculated based on the stationary enthalpy balance (Baehr and Stephan, 2006; Incropera et al., 2007)

$$0 = \dot{H}_i^{in} + \dot{H}_i^{a} + \dot{H}_i^{f} + \dot{H}_i^{sp} - \dot{H}_i^{out} + \dot{Q}_{g,i}.$$

Here, $\dot{H}_i^{in}$, $\dot{H}_i^{a}$, $\dot{H}_i^{f}$, $\dot{H}_i^{sp}$, and $\dot{H}_i^{out}$ correspond to the enthalpy flows of the incoming bulk flow, combustion air, fuel, nitrogen flushing, and the outgoing flue gas stream, respectively. $\dot{Q}_{g,i}$ is the net heat flow into the flue gas and includes convection and thermal radiation. An enthalpy flow can be determined by $\dot{H} = \sum_{\nu\in\mathcal{B}} \dot{M}^{\nu} h^{\nu}(T)$ with the specific enthalpy $h^{\nu}(T)$ and the mass flow $\dot{M}^{\nu}$ of component $\nu \in \mathcal{B}$ (Turns, 2006). This yields a nonlinear equation

$$\begin{aligned} 0 = &\sum_{\nu\in\mathcal{G}} \dot{M}_{i-1}^{out,\nu} h^{\nu}(T_{g,i-1}) + \sum_{\kappa\in\mathcal{A}} \dot{M}_i^{a,\kappa} h^{\kappa}(T_i^{a}) \\ &+ \dot{M}_i^{f} h^{\mathrm{CH_4}}(T_i^{f}) + \dot{M}_i^{sp} h^{\mathrm{N_2}}(T_i^{sp}) \\ &- \sum_{\nu\in\mathcal{G}} \dot{M}_i^{out,\nu} h^{\nu}(T_{g,i}) + \dot{Q}_{g,i}, \end{aligned} \tag{4}$$

where T_i^{a}, T_i^{f}, and T_i^{sp} are the temperatures of combustion air, fuel, and nitrogen, respectively. For determining the flue gas temperatures $T_{g,i}$ by means of (4), the outgoing mass flows $\dot{M}_i^{out,\nu}$ are required. Since these mass flows only depend on the mass flows of fuel supplied to the four heating zones $\mathbf{m}_D = \left[\dot{M}_{\alpha}^{f}\right]_{\alpha\in\mathcal{D}} \in \mathbb{R}^4$, the air-fuel equivalence ratios $\boldsymbol{\lambda} = [\lambda_{\alpha}]_{\alpha\in\mathcal{D}\cup\{pcc\}} \in \mathbb{R}^5$, and the mass flows of nitrogen flushing $\mathbf{m}_N = \left[\dot{M}_{\alpha}^{sp}\right]_{\alpha\in\mathcal{D}} \in \mathbb{R}^4$, they directly follow from (3).

2.3. W-shaped radiant tube

The N_r W-shaped radiant tubes are equipped with gas-fired burners and local recuperators, cf. Fig 1. In the radiant tubes, the combustion is fuel lean, i.e., $\lambda \geq 1$. In (Niederer et al., 2014), a semi-empirical nonlinear mapping Ψ_i depending on the mass flows $\mathbf{m}_I = \left[\dot{M}_{rth1}^{f}, \dot{M}_{rth2}^{f}, \dot{M}_{rts}^{f}\right]^{\mathrm{T}} \in \mathbb{R}^3$ of fuel is suggested to calculate the heat input due to the combustion of fuel in the form $\dot{Q}_{c,i} = \Psi_i(\mathbf{m}_I)$, $i = 1, \ldots, N_r$. Note, the effect of the recuperator is incorporated into the mapping Ψ_i.
In (Niederer et al., 2014), a radiant tube is approximated by four straight pipes each with a thickness d_r, a mass density ρ_r, and a temperature-dependent specific heat capacity c_r. In the current work, only the temperature $T_{r,i,1}$ of the first pipe of the radiant tube i is calculated and the remaining temperatures are chosen proportional to $T_{r,i,1}$. Starting from the first pipe, a nearly linear temperature drop over the pipes can be observed (Imose, 1985).
The temperature of the first pipe of the radiant tube i is based on the heat balance, resulting in

$$\frac{\mathrm{d}}{\mathrm{d}t} T_{r,i,1} = \frac{\dot{q}_{c,i,1} + \dot{q}_{r,i,1}}{\rho_r c_r(T_{r,i,1}) d_r}, \tag{5}$$

with $i = 1, \ldots, N_r$ and the radiative heat flux $\dot{q}_{r,i,1}$ from the furnace chamber to the first pipe. The heat flux $\dot{q}_{c,i,1} = w\dot{Q}_{c,i}$ into the inner surface of the tube is assumed to be uniform, with a weighting factor $w \in (0, 1)$, see (Niederer et al., 2014) for more details.

2.4. Furnace wall

The layered furnace wall is discretized into N_w wall segments. The temperature of the outer wall surface is assumed to be equal to the ambient temperature T_{∞}. On the inner surface, the boundary condition is defined by the heat flux $\dot{q}_w$ due to convection and radiation. The heat conduction through the wall can be characterized by the one-dimensional heat conduction equation (Baehr and Stephan, 2006; Incropera et al., 2007). To obtain a low-dimensional and computationally undemanding model, the Galerkin weighted residual method is applied (Fletcher, 1984). Here, the stationary solution of the heat conduction equation serves as a trial function. Finally, the dynamic behavior of the wall temperature can be defined by a lumped-parameter model

$$\frac{\mathrm{d}}{\mathrm{d}t} T_{w,i} = \frac{\dot{q}_{w,i}}{K_{1,i}} + \frac{K_{2,i}}{K_{1,i}} (T_{\infty} - T_{w,i}) \tag{6}$$

representing the inner surface temperature $T_{w,i}$ of the wall segment i with $i = 1, \ldots, N_w$. The parameters $K_{1,i}$ and $K_{2,i}$ are defined in (Niederer et al., 2015).

2.5. Roll and strip

The strip enters the furnace at $z = 0$ (cf. Fig. 1) with ambient temperature T_{∞} and moves through the furnace with the velocity v_s. The strip is characterized by the thickness d_s, the width b_s, the mass density ρ_s, and the temperature-dependent specific heat capacity c_s. The dynamic behavior of the strip temperature $T_{s,i}$ of a discretized section i is given in the form (Strommer et al., 2014a)

$$\frac{\mathrm{d}}{\mathrm{d}t} T_{s,i} = \frac{2\dot{q}_{s,i}}{\rho_s c_s(T_{s,i}) d_{s,i}} - v_s \frac{T_{s,i} - T_{s,i-1}}{\Delta z}, \tag{7}$$

with the boundary condition $T_{s,0}(t) = T_{\infty}$ at $z = 0$ and the local heat flux to the strip $\dot{q}_{s,i}(t)$, which comprises the convective, conductive, and radiative heat transfer. In (7), the backward finite difference formula is used to approximate the transport term. The strip is discretized into N_s sections with equidistant length Δz and locally uniform temperature $T_{s,i}$, $i = 1, \ldots, N_s$.
Thermal conduction occurs if the strip touches one of the N_h rolls. The temperature $T_{h,i}$ of a single roll i can be determined based on the heat balance, which gives

$$\frac{\mathrm{d}}{\mathrm{d}t} T_{h,i} = \frac{1}{\rho_h c_h(T_{h,i}) d_h} \left(\left(1 - \frac{S_h^c}{S_h}\right) \dot{q}_{h,i} + \frac{S_h^c}{S_h} \dot{q}_{h,i}^c \right). \tag{8}$$

Here, d_h is the wall thickness of the roll, ρ_h is its mass density, c_h is its temperature-dependent specific heat capacity, S_h is the total surface of the roll, and S_h^c is the

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

contact area between the strip and the roll. Moreover, $\dot{q}_{h,i}$ captures the heat flux to the roll by convection and radiation and $\dot{q}^{c}_{h,i}$ denotes the conductive heat flux.

2.6. Measurement of the strip surface temperature

It is well known that non-contact temperature measurement is a sophisticated task, in particular when the measured object moves. Typically, a pyrometer is used in such cases to measure the intensity. In the considered furnace, three pyrometers are available to measure the intensities I_α with $\alpha \in \mathcal{F} = \{dff, rth, rts\}$ (cf. Fig. 1).

Remark 2. Because the material quality depends on the temperature but not on the intensity and because it is difficult for most furnace operators to interpret intensity values, furnace control systems commonly use the temperature rather than the intensity.

Based on $I_\alpha = \sigma \varepsilon_{s,\alpha} T^4_{s,\alpha}$, the correct strip temperature is given in the form (Iuchi et al., 2010; Michalski et al., 2001)

$$T_{s,\alpha} = \left(\frac{I_\alpha}{\sigma \varepsilon_{s,\alpha}} \right)^{1/4}, \tag{9}$$

with the Stefan-Boltzmann constant σ and the (unknown) real strip emissivity $\varepsilon_{s,\alpha}$ at the corresponding pyrometer position z_α. The strip emissivity $\varepsilon_{s,\alpha}$ can be estimated by means of a state estimator (Strommer et al., 2016).

2.7. State-space model and discrete-time representation

In the previous sections, the individual subsystems were presented. Now, these subsystems are assembled in a reduced mathematical model of the considered CSAF. The subsystems (4)–(8) are interconnected via the heat transfer mechanisms (Niederer et al., 2014, 2015; Strommer et al., 2014a). The reduced model can be written in state-space form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}_1 = \mathbf{f}\left(t, \mathbf{x}_1, \mathbf{x}_2, \mathbf{m}, v_s\right) \tag{10a}$$

$$\mathbf{0} = \mathbf{a}\left(t, \mathbf{x}_1, \mathbf{x}_2, \mathbf{m}, \mathbf{m}_N\right) \tag{10b}$$

$$\mathbf{y} = \mathbf{b}^{\mathrm{T}} \mathbf{x}_1, \tag{10c}$$

with the state vector $\mathbf{x}_1 \in \mathbb{R}^{137}$, the algebraic variables $\mathbf{x}_2 \in \mathbb{R}^{13}$, the vector of mass flows $\mathbf{m}^{\mathrm{T}} = \left[\mathbf{m}_D^{\mathrm{T}}, \mathbf{m}_I^{\mathrm{T}}\right] \in \mathbb{R}^7$ of fuel, the mass flows $\mathbf{m}_N$ of nitrogen, and the initial condition $\mathbf{x}_1(t = t_0) = \mathbf{x}_{1,0}$. The mass flows of fuel, the strip velocity, and the mass flows of nitrogen are the system inputs, where the mass flows of fuel and nitrogen depend on the switching state s, see Tabs. 1 and 2. Equation (10c) defines the system output $\mathbf{y}$ which corresponds to the strip temperatures at the pyrometer positions z_α, $\alpha \in \mathcal{F}$, i.e., $\mathbf{y} = [T_{s,\alpha}]_{\alpha \in \mathcal{F}}$. The linear mapping in (10c) is defined by the vector $\mathbf{b}$. The state vector $\mathbf{x}_1$ summarizes the temperatures of radiant tubes, rolls, strip, and walls. The flue gas temperatures are assembled in the vector $\mathbf{x}_2$. Equation (10) is nonlinear and time variant. The steady state of (10) can be computed from

$$\mathbf{0} = \mathbf{\Pi}(t, \mathbf{x}, \mathbf{U}) = \begin{bmatrix} \mathbf{f}\left(t, \mathbf{x}_1, \mathbf{x}_2, \mathbf{m}, v_s\right) \\ \mathbf{a}\left(t, \mathbf{x}_1, \mathbf{x}_2, \mathbf{m}, \mathbf{m}_N\right) \end{bmatrix}, \tag{11}$$

with the augmented state vector $\mathbf{x}^{\mathrm{T}} = \left[\mathbf{x}_1^{\mathrm{T}}, \mathbf{x}_2^{\mathrm{T}}\right] \in \mathbb{R}^{150}$ and the system input $\mathbf{U}^{\mathrm{T}} = \left[\mathbf{m}_D^{\mathrm{T}}, \mathbf{m}_I^{\mathrm{T}}, \mathbf{m}_N^{\mathrm{T}}, v_s\right] \in \mathbb{R}^{12}$. For computer implementation, (10) has to be integrated in time. In (Niederer et al., 2014, 2015; Strommer et al., 2014a), Euler's explicit method is used with a small sampling time Δt_k. To decrease computational costs, a larger sampling time is preferable, which can, however, jeopardize the accuracy or even cause numerical instability. Numerical methods of higher order may help to keep the accuracy high even for larger sampling times. In (7), the Courant-Friedrichs-Lewy (CFL) condition has to be met for numerical stability (Strikwerda, 2004), i.e., $\Delta t_k v_s \leq \Delta z$ with the sampling time Δt_k, the spatial discretization Δz of the strip, and the strip velocity v_s. Hence, a larger sampling time Δt_k requires an appropriate discretization Δz. Applying the second-order half-explicit Runge-Kutta method (Ascher and Petzold, 1998), the discrete-time system can be expressed by a predictor step

$$\mathbf{X}_{1,k+1} = \mathbf{x}_{1,k} + \Delta t_k \mathbf{f}_k\left(\mathbf{x}_{1,k}, \mathbf{x}_{2,k}, \mathbf{m}_k, v_{s,k}\right) \tag{12a}$$

$$\mathbf{0} = \mathbf{a}_k\left(\mathbf{X}_{1,k+1}, \mathbf{X}_{2,k+1}, \mathbf{m}_k, \mathbf{m}_{N,k}\right) \tag{12b}$$

and a corrector step

$$\begin{aligned} \mathbf{x}_{1,k+1} = \mathbf{x}_{1,k} + \frac{\Delta t_k}{2}\Big(&\mathbf{f}_k\left(\mathbf{x}_{1,k}, \mathbf{x}_{2,k}, \mathbf{m}_k, v_{s,k}\right) \\ &+ \mathbf{f}_{k+1}\left(\mathbf{X}_{1,k+1}, \mathbf{X}_{2,k+1}, \mathbf{m}_{k+1}, v_{s,k+1}\right)\Big) \end{aligned} \tag{13a}$$

$$\mathbf{0} = \mathbf{a}_{k+1}\left(\mathbf{x}_{1,k+1}, \mathbf{x}_{2,k+1}, \mathbf{m}_{k+1}, \mathbf{m}_{N,k+1}\right), \tag{13b}$$

where $\mathbf{X}_{1,k+1}$ and $\mathbf{X}_{2,k+1}$ are intermediate values and $\mathbf{x}_{1,k}$ and $\mathbf{x}_{2,k}$ are the values of $\mathbf{x}_1$ and $\mathbf{x}_2$ at the grid points $t_k = t_0 + \sum_{i=1}^{k} \Delta t_i$. Inserting (12) into (13), the discrete-time system can be written as

$$\mathbf{0} = \mathbf{\Gamma}_k(\mathbf{x}_k, \mathbf{x}_{k+1}, \mathbf{U}_k, \mathbf{U}_{k+1}), \tag{14}$$

with the initial state $\mathbf{x}_0^{\mathrm{T}} = \left[\mathbf{x}_{1,0}^{\mathrm{T}}, \mathbf{x}_{2,0}^{\mathrm{T}}\right]$. Here, the initial state $\mathbf{x}_{2,0}$ follows from solving (10b) with $\mathbf{x}_{1,0}$, $\mathbf{m}_0 = \mathbf{m}(t = t_0)$, and $\mathbf{m}_{N,0} = \mathbf{m}_N(t = t_0)$.

Remark 3. In the dynamic optimization (cf. Sec. 4), the system (14) is evaluated recurrently during a particular time interval $\left[t_b^i, t_e^i\right]$, i.e., $\mathbf{0} = \mathbf{\Gamma}_{k-1}(\mathbf{x}_{k-1}, \mathbf{x}_k, \mathbf{U}_{k-1}, \mathbf{U}_k)$ with the initial condition $\mathbf{x}_{k_0}$, $k \in K$, and the set $K = \{k_0 + 1, \ldots, k_1\}$, see Fig. 2. The parameters $k_0, \ldots, k_1$ are sampling points which correspond to $t_b^i = t_{k_0}, \ldots, t_e^i = t_{k_1}$. At the beginning of the first optimization horizon, $t_b^1 = t_{k_0} = t_0$ and the initial condition $\mathbf{x}_{k_0} = \mathbf{x}_0$.

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005

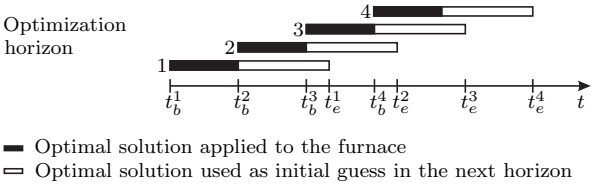

Figure 2: Receding horizon approach.

3. Furnace control system

The mass flow of fuel into each heating zone (hz a, hz b, hz c, hz d, rth 1, rth 2, and rts), the corresponding air-fuel equivalence ratios, the switching state, the strip velocity, and the strip properties (geometry and steel grade) constitute the primary inputs of the CSAF. Next, the control structure is briefly discussed.

The primary control objective is that the strip temperature follows a set-point trajectory at the positions z_{rth} and z_{rts}, where the pyrometers P_{rth} and P_{rts} are installed, see Fig. 1. These strip temperatures are the most important temperatures in terms of material properties and control. Maximization of throughput and minimization of energy consumption are secondary goals. To cater for all these requirements and constraints, a control structure with several hierarchical layers seems appropriate. Additionally, different response times of certain parts of the plant motivate the use of a cascaded structure. For instance, the response time of the control valves of the fuel and air supply is typically much smaller than the response time of the temperatures of the furnace wall and the rolls, see Sec. 1.3.

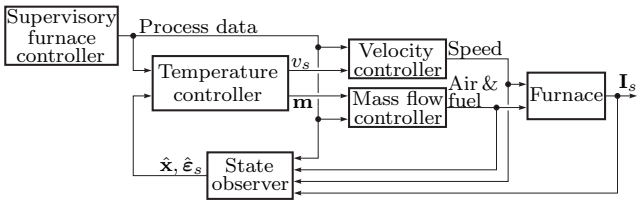

Figure 3: Hierarchical control structure of the furnace control system.

3.1. Hierarchical control layers

Figure 3 shows the hierarchical control system consisting of three layers, i.e., the supervisory furnace controller, the temperature controller, and the subordinate controllers for the mass flow and the strip velocity. The supervisory furnace controller has the highest authority and stipulates all production steps of the hot-dip galvanizing line. It defines the sequence of strips, their set-point temperatures, and the corresponding constraints. This information is referred to as process data. Table 3 shows some process data specified by the supervisory furnace controller for each individual strip l. The set-point values of the air-fuel equivalence ratios $\boldsymbol{\lambda}^l$ are chosen depending on the steel grade and metallurgical requirements. The scalar temperature values $T_{s,\alpha}^{l,d}$ and $T_{s,\alpha}^{l,\pm}$, $\alpha \in \mathcal{F}$, are defined at the positions z_α, i.e., the positions of the pyrometers P_α. Depending on these scalar values and the position trajectory of the strip, the supervisory furnace controller generates piecewise-constant trajectories $T_{s,\alpha}^{d}(t)$ and $T_{s,\alpha}^{\pm}(t)$. In the same way, the piecewise-constant trajectory $\boldsymbol{\lambda}(t)$ of the air-fuel equivalence ratios is designed based on $\boldsymbol{\lambda}^l$.

Variable	Description
$T_{s,\alpha}^{l,d}$	Set-point strip temperature at z_α
$T_{s,\alpha}^{l,-}$	Lower bound on strip temperature at z_α
$T_{s,\alpha}^{l,+}$	Upper bound on strip temperature at z_α
t^l	Time when a strip transition takes place
t_s^l	Beginning of a change of the strip velocity
λ_β^l	Set-point air-fuel equivalence ratio in hz a–d
λ_{pcc}^l	Set-point air-fuel equivalence ratio in PCC
ρ_s^l	Mass density of the strip material
c_s^l	Specific heat capacity of the strip material
d_s^l	Strip thickness
b_s^l	Strip width
L_s^l	Length of strip

Table 3: Process data of the strip l defined by the supervisory furnace controller with $\alpha \in \mathcal{F}$ and $\beta \in \mathcal{D}$.

The second layer (temperature controller) consists of four modules, see Sec. 4 and Fig. 5. In this layer, optimization-based methods are applied to determine the mass flows of fuel and the strip velocity so that a desired heating of the strip is achieved. This layer is the centerpiece of this paper and the corresponding control tasks are specified in Sec. 3.2.

Since the CSAF is only equipped with few measurement devices, a state observer is employed to provide the temperature controller with estimated system states. In fact, the observer estimates the augmented state $\hat{\mathbf{x}}$ and the badly known strip emissivity $\hat{\boldsymbol{\varepsilon}}_s$ (Strommer et al., 2016). The observer is an adaptive estimator, which uses a copy of the system (11) and a heuristic update law $\dot{\hat{\boldsymbol{\varepsilon}}}_s = \mathbf{W}^o(\mathbf{I}_s - \hat{\mathbf{I}}_s)$. $\mathbf{W}^o$ is a positive-definite diagonal matrix, $\mathbf{I}_s$ are the measured intensities of the strip, see Fig. 3, and $\hat{\mathbf{I}}_s$ are the estimated intensities of the strip at the pyrometer positions z_{dff}, z_{rth}, and z_{rts}. The estimated intensities can be computed by $\hat{\mathbf{I}}_s = \left[\sigma \hat{\boldsymbol{\varepsilon}}_{s,\alpha} \hat{T}_{s,\alpha}^4\right]_{\alpha \in \mathcal{F}}$ with the estimated strip emissivity $\hat{\boldsymbol{\varepsilon}}_{s,\alpha}$ and temperature $\hat{T}_{s,\alpha}$, see Sec. 2.6.

The third layer performs subordinate tracking control tasks for the mass flows of fuel and air and the strip velocity. Generally, this layer uses decentralized SISO PI-controllers. Moreover, cross-limiting controllers ensure that the desired air-fuel equivalence ratios are realized (Froehlich et al., 2016; Strommer et al., 2014b, 2017).

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005

For the design of the superordinate temperature controller, it is assumed that the subordinate mass flow and strip velocity controllers are ideal.

3.2. Control task of the temperature controller

The strip temperatures $\mathbf{y}_k = \mathbf{y}(t_k)$ at the positions z_{dff}, z_{rth}, and z_{rts} should follow their set-point trajectories $\mathbf{y}_k^d = \left[T_{s,\alpha}^d(t_k)\right]_{\alpha\in\mathcal{F}}$. If the strip temperatures do not reach the set-point values, they should be at least within an admissible range defined by

$$\mathbf{y}_k^- \leq \mathbf{y}_k \leq \mathbf{y}_k^+, \tag{15}$$

with the vector $\mathbf{y}_k^\pm = \left[T_{s,\alpha}^\pm(t_k)\right]_{\alpha\in\mathcal{F}}$ of lower and upper bounds. The constraint (15) ensures that the final steel product has the desired material properties. In the soaking zone (rts), the strip is held at an elevated temperature (Trinks et al., 2004). This is why the set-point strip temperatures $T_{s,rth}^d$ and $T_{s,rts}^d$ are often chosen identical. Note that for various steel grades this may not be the case. The exact set-point temperatures are the specific know-how of the plant operator. Figure 4 shows an exemplary heating curve of a strip section in the CSAF at the grid point t_k. Moreover, the set-point strip temperatures

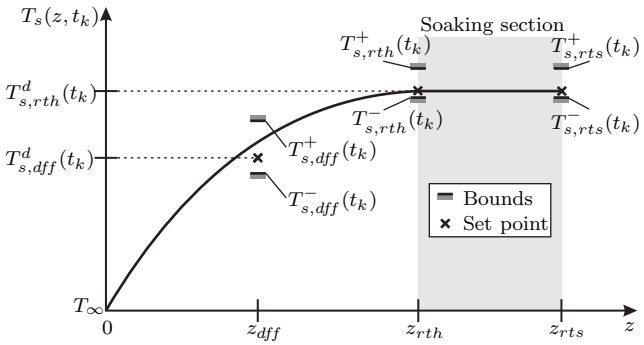

Figure 4: Typical heating curve of a strip section in the furnace.

$\mathbf{y}_k^d$ and the corresponding bounds $\mathbf{y}_k^\pm$ are indicated in this figure.

The mass flows $\mathbf{m}$ of fuel of the active heating zones, cf. Tab 1, and the strip velocity v_s are constrained in terms of their absolute values, which can be written in a time-discrete form as

$$\mathbf{m}_k^- \leq \mathbf{m}_k \leq \mathbf{m}_k^+, \quad k \in K \tag{16a}$$
$$v_{s,k}^- \leq v_{s,k} \leq v_{s,k}^+, \quad k \in K, \tag{16b}$$

with the bounds $\mathbf{m}^-$, $\mathbf{m}^+$, v_s^-, and v_s^+. Moreover, the slopes of the mass flows of fuel and the strip velocity are also constrained, i.e.,

$$\dot{\mathbf{m}}_k^- \leq \frac{\mathbf{m}_k - \mathbf{m}_{k-1}}{\Delta t_{k-1}} \leq \dot{\mathbf{m}}_k^+, \quad k \in K \tag{17a}$$
$$\dot{v}_{s,k}^- \leq \frac{v_{s,k} - v_{s,k-1}}{\Delta t_{k-1}} \leq \dot{v}_{s,k}^+, \quad k \in K, \tag{17b}$$

with the bounds $\dot{\mathbf{m}}^-$, $\dot{\mathbf{m}}^+$, $\dot{v}_s^-$, and $\dot{v}_s^+$. Usually, the bounds $\mathbf{m}_k^\pm$ and $\dot{\mathbf{m}}_k^\pm$ as well as $v_{s,k}^\pm$ and $\dot{v}_{s,k}^\pm$ are independent. The constraints (17) take into account that temperatures inside the CSAF and the electrical drives which convey the strip cannot vary at arbitrary rates due to the limited valve and drive dynamics, the risk of heat buckling, and the possible damage of devices and of the furnace interior (Imose, 1985).

Remark 4. Besides the steel grade and product type, also downstream process steps like the zinc bath may limit the strip velocity. In some cases, the desired velocity $v_{s,k} = v_{s,k}^d$ is prescribed by other process steps, which means that it is no longer a control input.

If there is a change in the trajectory $s_k = s(t_k)$ of the switching state in the heating zones, an optimal switching time has to be determined by the temperature controller. The switching time τ^l, $l \in \mathbb{N}$, is constrained by

$$\tau^{l,-} \leq \tau^l \leq \tau^{l,+} \tag{18}$$

with the bounds $\tau^{l,\pm}$. Equation (18) ensures that the switching occurs in the neighborhood of the corresponding strip transition. Moreover, a certain time $\Delta\tau$ must elapse until a new change of the switching state s is allowed, i.e., $\tau^{l,+} + \Delta\tau < \tau^{l+1,-}$.

The control objectives and constraints stated in this section correspond to the objectives outlined in Sec. 1.3. The primary goal is the accurate control of the strip temperature, which is inherently linked to an optimal product quality. Moreover, the temperature controller should also incorporate the secondary objectives, i.e., maximum throughput and minimum energy consumption. Clearly, the throughput is proportional to the strip velocity and the energy consumption is proportional to the mass flows of fuel.

4. Temperature controller

This section deals with the individual modules of the temperature controller, i.e., the static optimization, the trajectory generator for the strip velocity, the optimization-based trajectory planner (OTP), and the temperature regulator (TR), see Fig. 5. First, the structure of the temperature controller is motivated. Then, the individual modules will be discussed in more detail. Finally, a numerical solution of the underlying optimization problems is presented.

4.1. Structure of the temperature controller

Table 1 indicates that the mass flows of fuel of the heating zones c and d may vary discontinuously. Systematic optimization of the mass flows thus involves a time-consuming mixed-integer programming problem. In this paper, a static optimization problem is formulated instead, which determines the optimum switching state s^l and an optimal strip velocity v_s^l for each strip l. Afterwards, the

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005

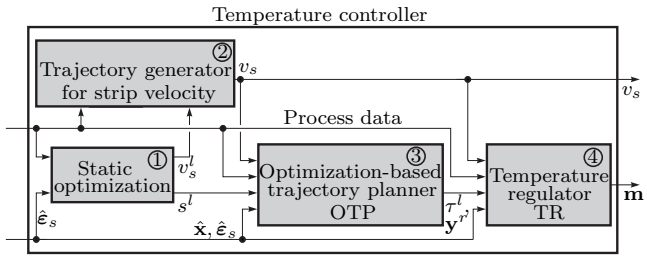

Figure 5: Structure of the temperature controller.

optimal strip velocities v_s^{l-1} and v_s^l of consecutive strips are used to design an appropriate trajectory $v_s(t)$ of the strip velocity in the transition region.
The optimal mass flows $\mathbf{m}$ of fuel are determined by the TR. It is mandatory that the TR is executed in real time. Therefore, the time grid and the time horizon $[t_{k_0}, t_{k_1}]$ must be adequately chosen. If the time horizon is too short, a violation of the bounds $T_{s,\alpha}^{\pm}$ may occur due to the high thermal inertia of the CSAF. In case of a boundary violation, the product quality may suffer and in the worst case, scrap may be produced. Figure 6 gives a graphical representation of this situation. At time t^1, a

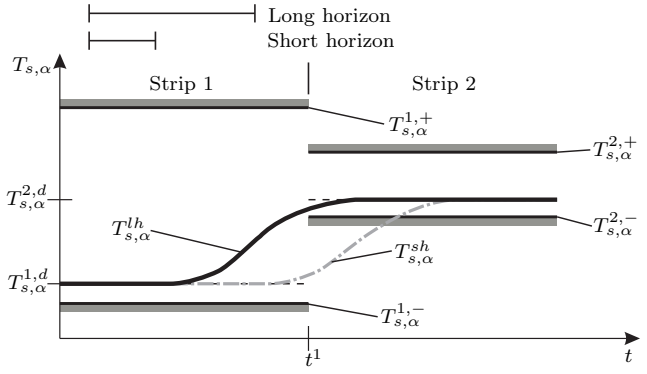

Figure 6: Treatment of a change of the set-point strip temperature in terms of different time horizons, $\alpha \in \mathcal{F}$.

strip transition takes place, i.e., the head of the subsequent strip enters the furnace and thus the set-point strip temperature and the corresponding bounds change. Using a short prediction horizon, the TR is not able to change from $T_{s,\alpha}^{1,d}$ to $T_{s,\alpha}^{2,d}$ without a violation of the lower bound due to the high thermal inertia of the CSAF, see the gray dash-dotted line $T_{s,\alpha}^{sh}$ in Fig. 6. Therefore, the OTP uses a long planning horizon to design reference trajectories $\mathbf{y}^r(t) = \left[T_{s,\alpha}^r(t)\right]_{\alpha\in\mathcal{F}}$ of the strip temperature which do not violate the bounds, see the black solid line $T_{s,\alpha}^{lh}$ in Fig. 6.
The OTP also determines an optimal switching time τ^l if there is a change of the switching state s^l, i.e., if $s^l \neq s^{l+1}$. As indicated in Fig. 7, the discrete-time trajectory s_k depends on the switching time τ^l and the switching states s^l

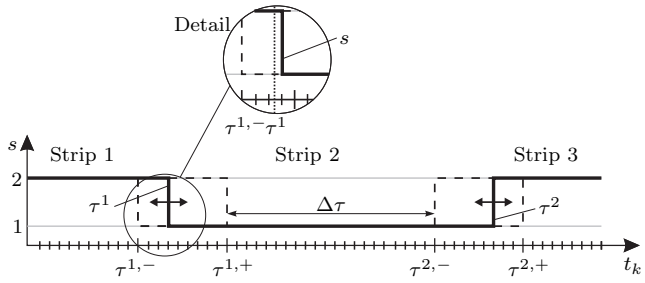

Figure 7: Influence of the switching time on the trajectory of the switching state.

and s^{l+1} of the current and the subsequent strip, respectively. Moreover, the detail in Fig. 7 shows that τ^l is generally round up to the next sampling point t_k (Flasskamp et al., 2012). Clearly, if $s^l = s^{l+1}$, the variable τ^l has no effect and does not need to be defined.
The trajectories of the reference temperatures $\mathbf{y}^r$ and the strip velocity v_s, the optimal switching time τ^l, and the process data are used by the TR to determine the optimal mass flows $\mathbf{m}$ of fuel, see Fig. 3. The TR uses a short prediction horizon to ensure execution in real time.
The tasks and optimization results of the four modules are listed in Tab. 4 and will be discussed in more detail in the next sections.

4.2. Static optimization

The module static optimization calculates an optimal operating point characterized by the strip velocity v_s^l and the switching state s^l for each strip l, see Tab. 4. The calculation of an optimal operating point is based on process data and the estimated strip emissivity $\hat{\boldsymbol{\varepsilon}}_s$ provided by the state observer, see Fig. 3.
As discussed in Sec. 2.1 for switching on/off the heating zones, only three different cases, given in Tab. 1, can occur. From this triple, the optimum switching state s^l has to be selected, which can easily be done by complete enumeration. For each strip, the switching state s^l that is most efficient in terms of energy consumption and throughput is used. In the first iteration ($s^l = 1$), the control inputs $\mathbf{u}_1$ of the static optimization are defined as $\mathbf{u}_1 = \left[\mathbf{m}_D^{\mathrm{T}}, \mathbf{m}_I^{\mathrm{T}}, v_s\right]^{\mathrm{T}}$. In the second iteration ($s^l = 2$), the mass flow $\dot{M}_{hzd}^f$ of fuel to the heating zone d does not constitute a control input because $\dot{M}_{hzd}^f = 0$. In the third iteration ($s^l = 3$), the mass flows $\dot{M}_{hzc}^f$ and $\dot{M}_{hzd}^f$ of fuel to the heating zone c and d are zero.
The furnace process is represented by the steady-state model (11) with the system input $\mathbf{U} = \left[\mathbf{m}_D^T, \mathbf{m}_I^T, \mathbf{m}_N^T, v_s\right]^{\mathrm{T}}$.

Remark 5. The system input $\mathbf{U}$ depends on the control input $\mathbf{u}_1$ and the switching state s^l (see Tab. 1), i.e., $\mathbf{U} = \mathbf{U}(\mathbf{u}_1, s^l)$.

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

Module	Task and optimization results	Description
Static optimization ①	Switching state s^l for each strip l	Section 4.2
	Strip velocity v_s^l for each strip l	
Trajectory generator for strip velocity ②	Design of the strip velocity trajectory v_s	Section 4.3
Optimization-based trajectory planner ③	Target trajectories $\mathbf{y}^r$ of the strip temperatures	Section 4.4
	Switching time τ^l	
Model predictive controller ④	Mass flows $\mathbf{m}$ of fuel	Section 4.5

Table 4: Control tasks and optimization results of the modules of the temperature controller.

The nonlinear optimization problem of the module static optimization can be formulated as

$$\begin{aligned}
\underset{\mathbf{u}_1\in\mathbb{R}^{N_1}}{\text{minimize}} \quad & J_1 = \left\|\mathbf{y}-\mathbf{y}^{l,d}\right\|_{\mathbf{W}^s} + \left\|\mathbf{u}_1-\mathbf{u}_1^d\right\|_{\mathbf{W}^u} && \text{(19a)}\\
\text{subject to} \quad & \mathbf{0} = \mathbf{\Pi}(t,\mathbf{x},\mathbf{U}(\mathbf{u}_1,s^l)) && \text{(19b)}\\
& \mathbf{u}_1^- \le \mathbf{u}_1 \le \mathbf{u}_1^+ && \text{(19c)}\\
& \mathbf{y}^- - \mathbf{y} \le \mathbf{0} && \text{(19d)}\\
& \mathbf{y} - \mathbf{y}^+ \le \mathbf{0} && \text{(19e)}
\end{aligned}$$

with the vector $\mathbf{y}^{l,d} = \left[T_{s,\alpha}^{l,d}\right]_{\alpha\in\mathcal{F}}$ of set-point strip temperatures, the set-point values $\mathbf{u}_1^d = \left[(\mathbf{m}^d)^{\mathrm{T}}, v_s^d\right]^{\mathrm{T}}$ of the control inputs, the bounds $\mathbf{u}_1^{\pm}$ of the control inputs, and the constraints (15) and (16). The second term in (19a) is used to weight the energy consumption and the throughput.

Remark 6. To achieve minimum energy consumption, $\mathbf{m}^d$ is set to zero. To realize maximum throughput, the desired strip velocity v_s^d is set to the upper bound v_s^+ of the strip velocity, i.e., $\mathbf{u}_1^d = \left[\mathbf{0}^{\mathrm{T}}, v_s^+\right]^{\mathrm{T}}$. Note, the secondary goals minimum energy consumption and maximum throughput are antagonistic. By choosing appropriate weighting matrices in (19a), a weighting of these goals can be achieved.

The problem (19) possesses $N_1 = \dim(\mathbf{u}_1) = n_D + n_I + n_{v_s}$ optimization variables with $(s^l, n_D) \in \{(1,4),(2,3),(3,2)\}$, $n_I = \dim(\mathbf{m}_I) = 3$, and $n_{v_s} = 1$. The state constraints (19d) and (19e) ensure that the strip temperatures at z_{dff}, z_{rth}, and z_{rts} remain within the permissible range, cf. (15). In (19a), $\|\boldsymbol{\theta}\|_{\mathbf{W}} = \frac{1}{2}\boldsymbol{\theta}^{\mathrm{T}}\mathbf{W}\boldsymbol{\theta}$ indicates a quadratic form. $\mathbf{W}$, $\mathbf{W}^s$, and $\mathbf{W}^u$ are positive-definite weighting matrices. J_1, $\mathbf{\Pi}$, and the state constraints (19d) and (19e) are assumed to be continuous in their arguments $\mathbf{x}$, $\mathbf{m}$, and v_s.
A (unique) solution of (19) cannot be guaranteed without further assumptions (Keerthi and Gilbert, 1985). The challenges of solving (19) are the inequality constraints (19c) – (19e). Thus, in the following, (19) will be transformed into an unconstrained optimization problem.
The box constraints (19c) can be eliminated by the nonlinear input transformation (Graichen and Petit, 2008)

$$\mathbf{u}_1 = \boldsymbol{\varphi}(\boldsymbol{v}_1) = \frac{\mathbf{u}_1^- + \mathbf{u}_1^+}{2} + \frac{\mathbf{u}_1^+ - \mathbf{u}_1^-}{2}\tanh\left(\frac{2\boldsymbol{v}_1}{\mathbf{u}_1^+ - \mathbf{u}_1^-}\right), \tag{20}$$

with the new unconstrained optimization variables $\boldsymbol{v}_1$. In (20), each mathematical operation is applied to the respective element of the vector. Equation (20) ensures strict satisfaction of (19c). To avoid singular arcs, which occur when at least one element of $\boldsymbol{v}_1$ approaches infinity, a positive definite term $\frac{1}{2}\eta\boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{v}_1$ with a small positive parameter $\eta > 0$ is added to the objective function J_1 (Graichen and Petit, 2008).
The inequality constraints (19d) and (19e) can be considered via penalty terms, which are added to the objective function J_1, i.e.,

$$\begin{aligned}
r^-(\mathbf{x}) &= \left\|\max\left(\mathbf{0}, \mathbf{y}^- - \mathbf{y}\right)\right\|_{\mathbf{W}^-} && \text{(21a)}\\
r^+(\mathbf{x}) &= \left\|\max\left(\mathbf{0}, \mathbf{y} - \mathbf{y}^+\right)\right\|_{\mathbf{W}^+}. && \text{(21b)}
\end{aligned}$$

$\mathbf{W}^-$ and $\mathbf{W}^+$ are positive-definite weighting matrices. Using the penalty terms (21), the original constraints (19d) – (19e) may be violated. Thus, a conservative design of these constraints is advisable (Nocedal and Wright, 2006; Steinboeck et al., 2013).
Using (20) and (21), the original problem (19) can be formulated as an unconstrained optimization problem

$$\begin{aligned}
\underset{\boldsymbol{v}_1\in\mathbb{R}^{N_1}}{\text{minimize}} \quad & \bar{J}_1 = \left\|\mathbf{y}-\mathbf{y}^{l,d}\right\|_{\mathbf{W}^s} + \left\|\boldsymbol{\varphi}(\boldsymbol{v}_1)-\mathbf{u}_1^d\right\|_{\mathbf{W}^u}\\
& \qquad + \frac{1}{2}\eta\boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{v}_1 + r^-(\mathbf{x}) + r^+(\mathbf{x}) && \text{(22a)}\\
\text{subject to} \quad & \mathbf{0} = \mathbf{\Pi}\left(t,\mathbf{x},\mathbf{U}(\boldsymbol{\varphi}(\boldsymbol{v}_1),s^l)\right). && \text{(22b)}
\end{aligned}$$

In contrast to (19), the formulation (22) does not contain inequality constraints and hence a solution of this optimization problem always exists.
All weighting matrices $\mathbf{W}^s$, $\mathbf{W}^u$, and $\mathbf{W}^{\pm}$ used in (22) are determined empirically. In case of maximum throughput, the strip velocity should be as high as possible. Thus, the corresponding entry in the matrix $\mathbf{W}^u$ is chosen higher than the entries related to the energy consumption, and vice versa for the case of minimum energy consumption. The weighting $\mathbf{W}^s$ is high compared to $\mathbf{W}^u$ and $\mathbf{W}^{\pm}$ to prioritize the primary control objective.

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

4.3. Trajectory generator for strip velocity

Based on process data and the optimal strip velocity v_s^l obtained from the module static optimization, the module trajectory generator for strip velocity generates the trajectory

$$v_s(t) = \begin{cases} v_s^l & t < t_s^l \\ \theta_1 + \theta_2 t + \theta_3 t^2 + \theta_4 t^3 & t_s^l \le t \le t_s^l + \Delta t_s^l \\ v_s^{l+1} & t > t_s^l + \Delta t_s^l \end{cases} \quad (23)$$

based on a third-order polynomial with the coefficients θ_i, $i = 1, 2, 3, 4$. The parameters t_s^l and Δt_s^l define the beginning and the end of the transition. At the time t_s^l and $t_s^l + \Delta t_s^l$, the trajectory $v_s(t)$ has to be continuous and continuously differentiable, i.e., $v_s(t_s^l) = v_s^l$, $v_s(t_s^l + \Delta t_s^l) = v_s^{l+1}$, $\dot{v}_s(t_s^l) = 0$, and $\dot{v}_s(t_s^l + \Delta t_s^l) = 0$. Based on these conditions, the coefficients θ_i can be determined. By means of Δt_s^l, the slope and the length of the transition can be adjusted so that the constraint (17b) is met. The maximum value of the slope $\dot{v}_s(t)$ is at $t^\star = -\theta_3/3\theta_4$. The value $\dot{v}_s(t^\star)$ has to fulfill the constraint (17b). If this condition is violated, the parameter Δt_s^l has to be adapted accordingly. In Fig. 8, an example with three different strips is illustrated.

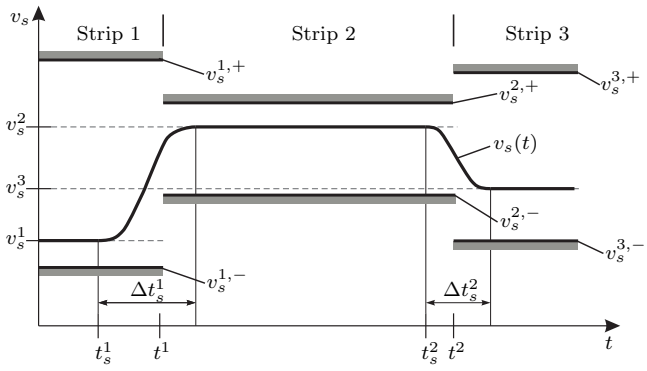

Figure 8: Design of the strip velocity trajectory.

4.4. Optimization-based trajectory planner

The module optimization-based trajectory planner (OTP) is primarily required to plan ahead for transient operating situations. The OTP determines reference trajectories $\mathbf{y}^r$ of the strip temperatures and an optimal switching time τ^l based on a long prediction horizon $[t_{k_0}, t_{k_1}]$, see Tab. 4. For realizing this task, the module OTP uses process data, the optimal switching state s^l provided by the static optimization, and the trajectory $v_{s,k}$ obtained from the trajectory generator for strip velocity. The piecewise constant trajectories $\mathbf{y}^d(t)$ of strip temperatures are also included in the process data, their values follow from the set-point temperatures $T_{s,\alpha}^{l,d}$ for each strip l, $\alpha \in \mathcal{F}$, and the position trajectory of the strip, see Sec. 3.1. The initial state of the optimization problem at the beginning of every optimization horizon is defined by the estimated current system state $\hat{\mathbf{x}}_0$ and the estimated strip emissivity $\hat{\varepsilon}_s$, see Fig. 2. These values are provided by the state observer, see Fig. 3.

Only the mass flows $\mathbf{m}_k$ of fuel of an active heating zone constitute control inputs to be manipulated. These mass flows are summarized in the vector $\mathbf{u}_{3,k}$ depending on the switching state s_k. The switching time τ^l is also an optimization variable in the OTP.

The furnace is represented by the discrete-time system (14) with the system input $\mathbf{U}_k = \mathbf{U}_k\left(\mathbf{u}_{3,k},\, s_k|_{\tau^l}\right)$. Based on the constraints (15), (16a), (17a), and (18), the nonlinear optimization problem of the OTP can be formulated as

$$\underset{(\mathbf{u}_3,\tau^l)\in\mathbb{R}^{N_3}\times\mathbb{R}}{\text{minimize}} \; J_3 = \sum_{k\in K} \left\|\mathbf{y}_k - \mathbf{y}_k^d\right\|_{\mathbf{W}_k^s} + \left\|\mathbf{u}_{3,k}\right\|_{\mathbf{W}_k^u} \quad (24a)$$

$$\text{subject to} \quad \mathbf{0} = \mathbf{\Gamma}_{k-1}\Big(\mathbf{x}_{k-1}, \mathbf{x}_k, \mathbf{U}_{k-1}\left(\mathbf{u}_{3,k-1},\, s_{k-1}|_{\tau^l}\right), \mathbf{U}_k\left(\mathbf{u}_{3,k},\, s_k|_{\tau^l}\right)\Big), \quad k \in K \quad (24b)$$

$$\mathbf{x}_{k_0} = \hat{\mathbf{x}}_0, \quad \mathbf{U}_{k_0} = \mathbf{U}_0 \quad (24c)$$

$$\mathbf{u}_{3,k}^- \le \mathbf{u}_{3,k} \le \mathbf{u}_{3,k}^+, \quad k \in K \quad (24d)$$

$$\dot{\mathbf{u}}_{3,k}^- \le \frac{\mathbf{u}_{3,k} - \mathbf{u}_{3,k-1}}{\Delta t_{k-1}} \le \dot{\mathbf{u}}_{3,k}^+, \quad k \in K \quad (24e)$$

$$\tau^{l,-} \le \tau^l \le \tau^{l,+} \quad (24f)$$

$$\mathbf{y}_k^- - \mathbf{y}_k \le \mathbf{0}, \quad k \in K \quad (24g)$$

$$\mathbf{y}_k - \mathbf{y}_k^+ \le \mathbf{0}, \quad k \in K, \quad (24h)$$

with $\mathbf{u}_3 = [\mathbf{u}_{3,k}]_{k\in K}$, the specified initial values $\mathbf{x}_{k_0}$ and $\mathbf{U}_{k_0}$, and the system input $\mathbf{U}_0 = \left[\mathbf{m}_0^{\mathrm{T}}, v_{s,0}\right]^{\mathrm{T}}$ at t_{k_0}. The constrained problem (24) possesses $N_3 = \dim(\mathbf{u}_3) + 1 = \sum_{i\in K} n_{D,i} + (k_1 - k_0)n_I + 1$ optimization variables with $(s_i, n_{D,i}) \in \{(1,4),(2,3),(3,2)\}$. Minimum energy consumption can be ensured by means of the second term in (24a). $\mathbf{W}_k^s$ and $\mathbf{W}_k^u$ are positive-definite weighting matrices. J_3, $\mathbf{\Gamma}$, and the state constraints (24g) and (24h) are assumed to be continuous in their arguments $\mathbf{x}$, $\mathbf{m}$, and τ^l.

Now, the original problem (24) is transformed into an unconstrained optimization problem similar to Sec. 4.2. The box constraints (24d) and (24f) can be eliminated by the nonlinear transformation (20). Thus, the mass flows $\mathbf{u}_{3,k}$ of fuel of an active heating zone and the switching time τ^l follow in the form $\mathbf{u}_{3,k} = \boldsymbol{\varphi}_k(\boldsymbol{\upsilon}_{3,k})$ and $\tau^l = \varphi(\tilde{\upsilon})$ with the new unconstrained optimization variables $\boldsymbol{\upsilon}_{3,k}$ and $\tilde{\upsilon}$, respectively. It is assumed that only one strip transition takes place during the prediction horizon $[t_{k_0}, t_{k_1}]$ and thus, only one switching time τ^l has to be determined. This assumption is justified because the processing time of the strips is larger than the length of the horizon of the OTP. The inequality constraints (24e), (24g), and (24h) can be considered via penalty terms. The inequality constraints of the states (24g) and (24h) can be taken into

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005

account similar to (21). The slopes of the mass flows of fuel (24e) can be incorporated in the form

$$r_k^{\dot{u}}(\mathbf{u}_{3,k-1},\mathbf{u}_{3,k}) = \|\max\Big(\mathbf{0},\mathbf{u}_{3,k-1}-\mathbf{u}_{3,k}+\Delta t_{k-1}\dot{\mathbf{u}}_{3,k}^{-},\ \mathbf{u}_{3,k}-\mathbf{u}_{3,k-1}-\Delta t_{k-1}\dot{\mathbf{u}}_{3,k}^{+}\Big)\|_{\mathbf{W}_k^{\dot{u}}}, \tag{25}$$

with the positive-definite weighting matrix $\mathbf{W}_k^{\dot{u}}$. To avoid singular arcs, again a positive definite term $\sum_{k\in K}\frac{1}{2}\eta\boldsymbol{\upsilon}_{3,k}^{\mathrm{T}}\boldsymbol{\upsilon}_{3,k}+\frac{1}{2}\eta\tilde{\upsilon}^2$ with a small positive parameter $\eta>0$ is added to the objective function J_3 in addition to the penalty terms. Thus, the original problem of the OTP (24) can be rewritten as an unconstrained optimization problem

$$\begin{aligned}\underset{(\boldsymbol{\upsilon}_3,\tilde{\upsilon})\in\mathbb{R}^{N_3}\times\mathbb{R}}{\text{minimize}}\ \bar{J}_3 = &\sum_{k\in K}\|\mathbf{y}_k-\mathbf{y}_k^d\|_{\mathbf{W}_k^s}+\|\boldsymbol{\varphi}_k(\boldsymbol{\upsilon}_{3,k})\|_{\mathbf{W}_k^u}\\ &+\frac{1}{2}\eta\boldsymbol{\upsilon}_{3,k}^{\mathrm{T}}\boldsymbol{\upsilon}_{3,k}+r_k^{-}(\mathbf{x}_k)+r_k^{+}(\mathbf{x}_k)\\ &+r_k^{\dot{u}}\left(\boldsymbol{\varphi}_{k-1}(\boldsymbol{\upsilon}_{3,k-1}),\boldsymbol{\varphi}_k(\boldsymbol{\upsilon}_{3,k})\right)\\ &+\frac{1}{2}\eta\tilde{\upsilon}^2\end{aligned} \tag{26a}$$

$$\begin{aligned}\text{subject to}\quad \mathbf{0} = &\boldsymbol{\Gamma}_{k-1}\Big(\mathbf{x}_{k-1},\mathbf{x}_k,\\ &\mathbf{U}_{k-1}\big(\boldsymbol{\varphi}_{k-1}(\boldsymbol{\upsilon}_{3,k-1}),\,s_{k-1}|_{\varphi(\tilde{\upsilon})}\big),\\ &\mathbf{U}_k\big(\boldsymbol{\varphi}_k(\boldsymbol{\upsilon}_{3,k}),\,s_k|_{\varphi(\tilde{\upsilon})}\big)\Big),\quad k\in K\end{aligned} \tag{26b}$$

$$\mathbf{0}=\mathbf{x}_{k_0}-\hat{\mathbf{x}}_0,\quad \mathbf{0}=\mathbf{U}_{k_0}-\mathbf{U}_0, \tag{26c}$$

with the assembled unconstrained input vector $\boldsymbol{\upsilon}_3=[\boldsymbol{\upsilon}_{3,k}]_{k\in K}$.

All weighting matrices $\mathbf{W}_k^s$, $\mathbf{W}_k^u$, $\mathbf{W}_k^{\dot{u}}$, and $\mathbf{W}_k^{\pm}$ used in (26) are determined empirically. Generally, the entries of the matrix $\mathbf{W}_k^s$ of the OTP are chosen to be constant apart from periods when a strip transition moves through the furnace. In this case, a modification of the entries is carried out. Here, four scenarios are distinguished as indicated in Fig. 9. In scenario I and II, the entries of the matrix $\mathbf{W}_k^s$ of the first strip decrease if a strip transition enters the furnace. The weighting of the second strip is designed to be constant and higher. Thus, a change from $T_{s,\alpha}^{1,d}$ to $T_{s,\alpha}^{2,d}$ can be realized without a violation of the bounds. In scenario III and IV, the matrix $\mathbf{W}_k^s$ corresponding to the first strip is constant to ensure that the actual strip temperature matches the desired one. A change from $T_{s,\alpha}^{1,d}$ to $T_{s,\alpha}^{2,d}$ is desired as soon as possible. Therefore, the weighting $\mathbf{W}_k^s$ of the second strip is adapted by increasing the entries as long as a strip transition appears in the optimization horizon.

4.5. Temperature regulator

The module temperature regulator (TR) is based on MPC technology and optimizes the mass flows $\mathbf{m}_k$ of fuel

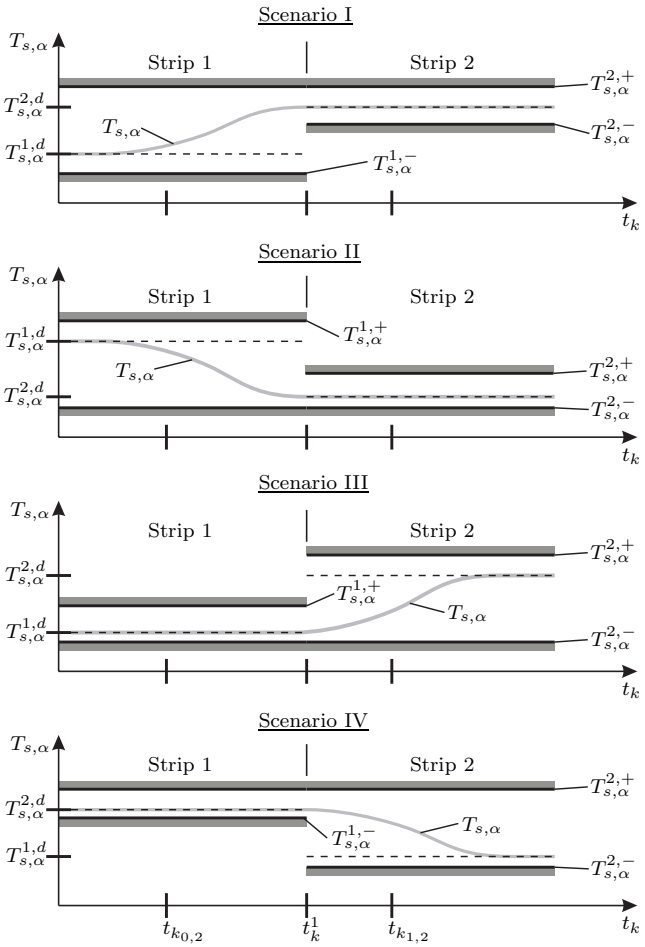

Figure 9: Four scenarios during a strip transition, $\alpha\in\mathcal{F}$.

to realize an optimal tracking control for the strip temperature based on a short time horizon $[t_{k_0},t_{k_1}]$. The TR utilizes process data, the strip velocity trajectory v_s obtained from the trajectory generator for strip velocity, the reference trajectories $\mathbf{y}^r$ of the strip temperature and the optimal switching time τ^l provided by the OTP to calculate the optimal mass flows of fuel. Moreover, the values $\hat{\mathbf{x}}_0$ and $\hat{\boldsymbol{\varepsilon}}_s$ provided by the state observer are utilized by the TR to define the initial state of the optimization problem.

The control inputs of the TR are the mass flows $\mathbf{m}$ of fuel of an active heating zone, which are summarized in the vector $\mathbf{u}_4$. The furnace is modeled by (14) with the system input $\mathbf{U}_k=\mathbf{U}_k(\mathbf{u}_{4,k},s_k)$.

The optimization problem of the TR is similar to (24), where the constraint (24f) is not taken into account. Moreover, the set-point values $\mathbf{y}^d$ in the objective function J_4 of the TR are replaced by the reference trajectories $\mathbf{y}^r$.

The constrained problem of the TR is transformed into an unconstrained problem based on the input transformation $\mathbf{u}_{4,k}=\boldsymbol{\varphi}_k(\boldsymbol{\upsilon}_{4,k})$ with the nonlinear mapping $\boldsymbol{\varphi}_k$ according to (20) and the new unconstrained optimization variables

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

$\boldsymbol{\upsilon}_{4,k}$ and the penalty terms (21) and (25). The unconstrained optimization problem of the TR can be written in the form

$$\begin{aligned}\underset{\boldsymbol{\upsilon}_4\in\mathbb{R}^{N_4}}{\text{minimize}}\quad & \bar{J}_4 = \sum_{k\in K} \|\mathbf{y}_k - \mathbf{y}_k^r\|_{\mathbf{W}_k^s} + \|\boldsymbol{\varphi}_k(\boldsymbol{\upsilon}_{4,k})\|_{\mathbf{W}_k^u} \\ & + \frac{1}{2}\eta\boldsymbol{\upsilon}_{4,k}^{\mathrm{T}}\boldsymbol{\upsilon}_{4,k} + r_k^-(\mathbf{x}_k) + r_k^+(\mathbf{x}_k) \\ & + r_k^{\dot{u}}\left(\boldsymbol{\varphi}_{k-1}(\boldsymbol{\upsilon}_{4,k-1}), \boldsymbol{\varphi}_k(\boldsymbol{\upsilon}_{4,k})\right) \qquad (27a)\\ \text{subject to}\quad & \mathbf{0} = \boldsymbol{\Gamma}_{k-1}\Big(\mathbf{x}_{k-1}, \mathbf{x}_k, \\ & \mathbf{U}_{k-1}\big(\boldsymbol{\varphi}_{k-1}(\boldsymbol{\upsilon}_{4,k-1}), s_{k-1}\big), \qquad (27b)\\ & \mathbf{U}_k\big(\boldsymbol{\varphi}_k(\boldsymbol{\upsilon}_{4,k}), s_k\big)\Big), \quad k\in K \\ & \mathbf{0} = \mathbf{x}_{k_0} - \hat{\mathbf{x}}_0, \quad \mathbf{0} = \mathbf{U}_{k_0} - \mathbf{U}_0, \qquad (27c)\end{aligned}$$

with the assembled unconstrained input vector $\boldsymbol{\upsilon}_4 = [\boldsymbol{\upsilon}_{4,k}]_{k\in K}$. The optimization problem (27) possesses $N_4 = \dim(\boldsymbol{\upsilon}_4) = \sum_{i\in K} n_{D,i} + (k_1 - k_0)n_I$ optimization variables. The second term in (27a) is used to minimize the energy consumption.

The weighting matrices $\mathbf{W}_k^s$, $\mathbf{W}_k^u$, $\mathbf{W}_k^{\dot{u}}$, and $\mathbf{W}_k^{\pm}$ in (27) are determined empirically. In contrast to the OTP, the entries of the matrix $\mathbf{W}_k^s$ are always constant. They are chosen high to prioritize the primary control objective, see Sec. 4.2.

4.6. Numerical solution of the optimization problem

For solving an unconstrained optimization problem like (22), (26), or (27) many different approaches exist, e.g., steepest descend method, Newton method, Quasi-Newton method, Gauss-Newton method (Bertsekas, 1999; Conn et al., 2000; Nocedal and Wright, 2006). In this paper, the Gauss-Newton method is employed because it features superlinear convergence. In the following, the application of the Gauss-Newton method is demonstrated for the optimization problem (26) of the OTP. The approach can be similarly applied to the static optimization and the TR, which is not further detailed in this paper.

Because (26a) is a quadratic form, (26) can be rewritten as

$$\begin{aligned}\underset{(\boldsymbol{\upsilon}_3,\tilde{\upsilon})\in\mathbb{R}^{N_3}\times\mathbb{R}}{\text{minimize}}\quad & \bar{J}_3 = \sum_{k\in K} \|\mathbf{R}_k\left(\mathbf{x}_k, \boldsymbol{\upsilon}_{3,k-1}, \boldsymbol{\upsilon}_{3,k}\right)\|_{\mathbf{W}_k} \\ & + \frac{1}{2}\eta\tilde{\upsilon}^2 \qquad (28a)\\ \text{subject to}\quad & \mathbf{0} = \boldsymbol{\Gamma}_{k-1}\Big(\mathbf{x}_{k-1}, \mathbf{x}_k, \\ & \mathbf{U}_{k-1}\big(\boldsymbol{\varphi}_{k-1}(\boldsymbol{\upsilon}_{3,k-1}), s_{k-1}|_{\varphi(\tilde{\upsilon})}\big), \qquad (28b)\\ & \mathbf{U}_k\big(\boldsymbol{\varphi}_k(\boldsymbol{\upsilon}_{3,k}), s_k|_{\varphi(\tilde{\upsilon})}\big)\Big), \quad k\in K \\ & \mathbf{0} = \mathbf{x}_{k_0} - \hat{\mathbf{x}}_0, \quad \mathbf{0} = \mathbf{U}_{k_0} - \mathbf{U}_0, \qquad (28c)\end{aligned}$$

with the weighting matrix $\mathbf{W}_k$ including the entries of $\mathbf{W}_k^s$, $\mathbf{W}_k^u$, $\mathbf{W}_k^-$, $\mathbf{W}_k^+$, $\mathbf{W}_k^{\dot{u}}$, and the scaled identity matrix $\eta\mathbf{I}$. The vector $\mathbf{R}_k(\mathbf{x}_k, \boldsymbol{\upsilon}_{3,k-1}, \boldsymbol{\upsilon}_{3,k})$ is defined as $\mathbf{R}_k = \left[(\mathbf{R}_k^y)^{\mathrm{T}}, (\mathbf{R}_k^u)^{\mathrm{T}}, (\mathbf{R}_k^-)^{\mathrm{T}}, (\mathbf{R}_k^+)^{\mathrm{T}}, (\mathbf{R}_k^{\dot{u}})^{\mathrm{T}}, \boldsymbol{\upsilon}_{3,k}^{\mathrm{T}}\right]^{\mathrm{T}}$ with the deviations of the strip temperatures $\mathbf{R}_k^y = \mathbf{y}_k - \mathbf{y}_k^d$, the deviations of the mass flows of fuel $\mathbf{R}_k^u = \boldsymbol{\varphi}_k(\boldsymbol{\upsilon}_{3,k})$, the penalty terms of the states $\mathbf{R}_k^- = \max\left(\mathbf{0}, \mathbf{y}_k^- - \mathbf{y}_k\right)$ and $\mathbf{R}_k^+ = \max\left(\mathbf{0}, \mathbf{y}_k - \mathbf{y}_k^+\right)$, and the penalty term of the slope of the mass flows of fuel $\mathbf{R}_k^{\dot{u}} = \max\Big(\mathbf{0}, \mathbf{u}_{3,k-1} - \mathbf{u}_{3,k} + \Delta t_{k-1}\dot{\mathbf{u}}_{3,k}^-, \mathbf{u}_{3,k} - \mathbf{u}_{3,k-1} - \Delta t_{k-1}\dot{\mathbf{u}}_{3,k}^+\Big)$. The Gauss-Newton method is a gradient-based method, which requires the gradient $\mathbf{g}$ and an approximate Hessian $\mathbf{H}$ of the objective function $\bar{J}_3$, i.e.,

$$\mathbf{g} = \left(\frac{\mathrm{d}\bar{J}_3}{\mathrm{d}\boldsymbol{\upsilon}}\right)^{\mathrm{T}} = \mathbf{R}^{\mathrm{T}}\mathbf{W}\frac{\mathrm{d}\mathbf{R}}{\mathrm{d}\boldsymbol{\upsilon}} \qquad (29a)$$

$$\mathbf{H} = \frac{\mathrm{d}^2\bar{J}_3}{\mathrm{d}\boldsymbol{\upsilon}^2} = \left(\frac{\mathrm{d}\mathbf{R}}{\mathrm{d}\boldsymbol{\upsilon}}\right)^{\mathrm{T}}\mathbf{W}\frac{\mathrm{d}\mathbf{R}}{\mathrm{d}\boldsymbol{\upsilon}} + \underbrace{\mathbf{R}^{\mathrm{T}}\mathbf{W}\frac{\mathrm{d}^2\mathbf{R}}{\mathrm{d}\boldsymbol{\upsilon}^2}}_{\approx\mathbf{0}}, \qquad (29b)$$

with $\boldsymbol{\upsilon} = \left[\boldsymbol{\upsilon}_3^{\mathrm{T}}, \tilde{\upsilon}\right]^{\mathrm{T}}$, the vector $\mathbf{R} = \left[\left([\mathbf{R}_k]_{k\in K}\right)^{\mathrm{T}}, \tilde{\upsilon}\right]^{\mathrm{T}}$, and the weighting matrix $\mathbf{W}$, which includes the entries of $\mathbf{W}_k$ and η. The second term in (29b) is assumed to be small compared to the first term and is thus neglected in the Gauss-Newton method (Nocedal and Wright, 2006). Therefore, only $\mathrm{d}\mathbf{R}/\mathrm{d}\boldsymbol{\upsilon}$ is needed to compute the gradient and the approximate Hessian. In this work, $\mathrm{d}\mathbf{R}/\mathrm{d}\boldsymbol{\upsilon}_3$ is analytically computed and $\mathrm{d}\mathbf{R}/\mathrm{d}\tilde{\upsilon}$ is numerically calculated. Considering the vector-valued Lagrange function (Sawaragi et al., 1985)

$$\begin{aligned}\mathbf{L}_k = {} & \mathbf{R}_k\left(\mathbf{x}_k, \boldsymbol{\upsilon}_{3,k-1}, \boldsymbol{\upsilon}_{3,k}\right) + \\ & \sum_{j=k_0+1}^{k} \boldsymbol{\Lambda}_{k,j}\boldsymbol{\Gamma}_{j-1}\Big(\mathbf{x}_{j-1}, \mathbf{x}_j, \\ & \mathbf{U}_{j-1}\big(\boldsymbol{\varphi}_{j-1}(\boldsymbol{\upsilon}_{3,j-1}), s_{j-1}|_{\varphi(\tilde{\upsilon})}\big), \\ & \mathbf{U}_j\big(\boldsymbol{\varphi}_j(\boldsymbol{\upsilon}_{3,j}), s_j|_{\varphi(\tilde{\upsilon})}\big)\Big)\end{aligned} \qquad (30)$$

with the Lagrange multipliers $\boldsymbol{\Lambda}_{k,j}$ and $k\in K$, the derivative of $\mathbf{R}_k$ with respect to the optimization variables $\boldsymbol{\upsilon}_3$ follows in the form

$$\frac{\mathrm{d}\mathbf{R}_k}{\mathrm{d}\boldsymbol{\upsilon}_3} = \frac{\partial\mathbf{L}_k}{\partial\boldsymbol{\upsilon}_3} = \left[\mathbf{R}'_{k,k_0+1}, \ldots, \mathbf{R}'_{k,k_1}\right] \qquad (31a)$$

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005

with

$$\mathbf{R}'_{k,j} = \begin{cases} \frac{\partial \mathbf{R}_k}{\partial \boldsymbol{\upsilon}_{3,j}} + \boldsymbol{\Lambda}_{k,j} \left(\frac{\partial \boldsymbol{\Gamma}_{j-1}}{\partial \boldsymbol{\upsilon}_{3,j}} \right) + \boldsymbol{\Lambda}_{k,j+1} \left(\frac{\partial \boldsymbol{\Gamma}_j}{\partial \boldsymbol{\upsilon}_{3,j}} \right) & j \in K_a \\ \frac{\partial \mathbf{R}_k}{\partial \boldsymbol{\upsilon}_{3,j}} + \boldsymbol{\Lambda}_{k,j} \left(\frac{\partial \boldsymbol{\Gamma}_{j-1}}{\partial \boldsymbol{\upsilon}_{3,j}} \right) & j = k \\ \mathbf{0} & j \in K_c \end{cases} \tag{31b}$$

$$\mathbf{0} = \boldsymbol{\Gamma}_{j-1}\Big(\mathbf{x}_{j-1}, \mathbf{x}_j, \mathbf{U}_{j-1}\big(\boldsymbol{\varphi}_{j-1}(\boldsymbol{\upsilon}_{3,j-1}), s_{j-1}|_{\varphi(\tilde{v})}\big), \mathbf{U}_j\big(\boldsymbol{\varphi}_j(\boldsymbol{\upsilon}_{3,j}), s_j|_{\varphi(\tilde{v})}\big)\Big), \quad j \in K_b \tag{31c}$$

$$\mathbf{0} = \mathbf{x}_{k_0} - \hat{\mathbf{x}}_0, \quad \mathbf{0} = \mathbf{U}_{k_0} - \mathbf{U}_0 \tag{31d}$$

$$\mathbf{0} = \frac{\partial \mathbf{L}_k}{\partial \mathbf{x}_j} = \boldsymbol{\Lambda}_{k,j} \frac{\partial \boldsymbol{\Gamma}_{j-1}}{\partial \mathbf{x}_j} + \boldsymbol{\Lambda}_{k,j+1} \frac{\partial \boldsymbol{\Gamma}_j}{\partial \mathbf{x}_j}, \quad j \in K_a \tag{31e}$$

$$\mathbf{0} = \frac{\partial \mathbf{L}_k}{\partial \mathbf{x}_k} = \frac{\partial \mathbf{R}_k}{\partial \mathbf{x}_k} + \boldsymbol{\Lambda}_{k,k} \frac{\partial \boldsymbol{\Gamma}_{k-1}}{\partial \mathbf{x}_k} \tag{31f}$$

and the sets $K_a = \{k_0+1, \ldots, k-1\}$, $K_b = \{k_0+1, \ldots, k\}$, and $K_c = \{k+1, \ldots, k_1\}$.

Remark 7. The derivative $\mathrm{d}\mathbf{R}/\mathrm{d}\tilde{v}$, which corresponds to the switching time $\tau^l = \varphi(\tilde{v})$, is numerically determined because this gradient cannot be easily calculated by means of the Gauss-Newton method (Flasskamp et al., 2012; Kaya and Noakes, 2003). The derivative $\mathrm{d}\mathbf{R}/\mathrm{d}\tilde{v}$ follows from a central difference quotient with an appropriate small step size. This step size is significantly larger than the sampling time Δt_k of the system and thus, the problem can be considered as quasi-continuous.

The calculation of $\mathrm{d}\mathbf{R}_k/\mathrm{d}\boldsymbol{\upsilon}_3$ can be carried out as follows: The state $\mathbf{x}_k$ is determined by solving (31c) in forward direction with increasing time index j and the initial condition (31d). Using (31f), the value $\boldsymbol{\Lambda}_{k,k}$ can be calculated. Afterwards, (31e) can be solved in backward direction with decreasing time index j to obtain the variables $\boldsymbol{\Lambda}_{k,j}$, $j \in K_a$. Then, $\mathrm{d}\mathbf{R}_k/\mathrm{d}\boldsymbol{\upsilon}_3$ can be determined according to (31a) and (31b). This procedure has to be done for all $k \in K$. Thus, $\mathrm{d}\mathbf{R}/\mathrm{d}\boldsymbol{\upsilon}$ can be assembled in the form

$$\frac{\mathrm{d}\mathbf{R}}{\mathrm{d}\boldsymbol{\upsilon}} = \left[\left(\frac{\mathrm{d}\mathbf{R}_{k_0+1}}{\mathrm{d}\boldsymbol{\upsilon}} \right)^{\mathrm{T}}, \ldots, \left(\frac{\mathrm{d}\mathbf{R}_{k_1}}{\mathrm{d}\boldsymbol{\upsilon}} \right)^{\mathrm{T}}, \left(\frac{\mathrm{d}\tilde{v}}{\mathrm{d}\boldsymbol{\upsilon}} \right)^{\mathrm{T}} \right]^{\mathrm{T}}. \tag{32}$$

Finally, the gradient and the approximate Hessian of the unconstrained optimization problem (28) can be determined by (29).

The Gauss-Newton method is used to iteratively solve (28), which is defined by the following algorithm.

Algorithm
Gauss-Newton method

(a) Provide an initial guess for $\boldsymbol{\upsilon}$

(b) Determine $\mathrm{d}\mathbf{R}/\mathrm{d}\boldsymbol{\upsilon}$, cf. (32).

(c) Calculate the gradient $\mathbf{g}$ and the approximate Hessian $\mathbf{H}$ according to (29).

(d) Compute a new search direction $\mathbf{d} = -\mathbf{H}^{-1}\mathbf{g}$ and calculate the optimal step length

$$\bar{\kappa} = \underset{\kappa \in \mathbb{R}^+}{\text{minimize}} \left\{ \bar{J}_3(\boldsymbol{\upsilon} + \kappa \mathbf{d}) \right\} \tag{33a}$$

$$\text{subject to} \quad \mathbf{0} = \boldsymbol{\Gamma}_{k-1}\left(\mathbf{x}_{k-1}, \mathbf{x}_k, \mathbf{U}_{k-1}, \mathbf{U}_k\right), \quad k \in K \tag{33b}$$

$$\mathbf{0} = \mathbf{x}_{k_0} - \hat{\mathbf{x}}_0, \quad \mathbf{0} = \mathbf{U}_{k_0} - \mathbf{U}_0 \tag{33c}$$

and execute the update $\boldsymbol{\upsilon} \leftarrow \boldsymbol{\upsilon} + \bar{\kappa}\mathbf{d}$.

(e) Stop if a suitable solution $\boldsymbol{\upsilon}$ has been reached and apply the transformation (20) to obtain the control inputs $\mathbf{u}_3$ and the switching time τ^l. Otherwise restart at step (b).

In step (e) of this algorithm, a termination criterion is required. The algorithm can for instance be stopped if $\bar{J}_3(\boldsymbol{\upsilon})$ or a specified norm of $\mathbf{d}$ is below a desired threshold or if a certain number of iteration loops is exceeded.

In step (d) of the algorithm, the one-dimensional optimization problem (33a) (line search problem) has to be solved. An approximate solution can be obtained by the approach of Graichen and Käpernick (2012).

For solving the optimization problem (28), the Gauss-Newton method is utilized because the implementation on a real-time system can be easily realized, it is characterized by superlinear convergence, and the gradient and the Hessian can be analytically calculated.

Using an optimization-based controller, an appropriate time grid has to be specified. Approximately 15 min are required to change between two different stationary operating conditions. Thus, the OTP uses an optimization horizon of 15 min to capture the thermal inertia of the furnace and to determine suitable reference trajectories of the strip temperatures without violating any bounds, cf. Sec. 3.2. The time horizon of the TR is 3 min to ensure real-time execution. Furthermore, this time nearly corresponds to the annealing time of a strip section, i.e., during this time interval, the strip section has to be heated to its set-point temperature. The sampling times $\Delta t_k \leq 5$ s of the OTP and the TR are chosen identical to realize the same degree of accuracy.

5. Simulation results

The mathematical model of the CSAF (Niederer et al., 2015) was validated by means of measurement data from the real furnace. Therefore, this model is used as a reference to verify the proposed control concept by simulations. A simulation scenario is created based on real measurement data. The considered scenario is representative insofar as it contains standard and special operating conditions. There are several changes of the strip properties, i.e., set-point strip temperatures $\mathbf{y}^d$, lower and upper bounds $\mathbf{y}^{\pm}$, the thickness d_s and the width b_s of the strip, see Fig. 10. Other parameters that are also time-varying

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005

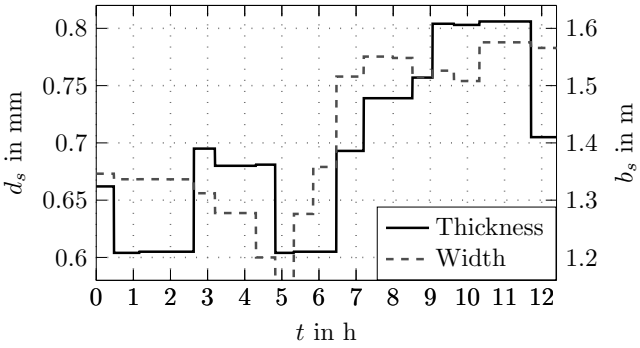

Figure 10: Thickness and width of the strip.

are the desired air-fuel equivalence ratios $\boldsymbol{\lambda}$ in the heating zones a–d and in the PCC.

In the static optimization, the weighting matrices are chosen to achieve a minimum energy consumption. In the OTP and TR, the weighting matrices are chosen primarily in view of an accurate heating of the strip, a minimum energy consumption, and acceptable slew rates of the control inputs $\mathbf{m}$.

Figure 11 shows the actual and the set-point strip temperatures as well as the corresponding lower and upper bounds at z_{dff}, z_{rth}, and z_{rts}. This demonstrates the good performance of the proposed control concept. The controller is able to fulfill all required demands of the stationary and transient furnace operations. Between 6 h and 10 h, the strip thickness varies from 0.6 mm up to 0.8 mm. Such a variation causes a significant change of the strip temperature (cf. Strommer et al. (2014a)). The set-point strip temperatures $\mathbf{y}^d$ are constant during this interval. Although the thickness changes significantly, a nearly optimal temperature tracking is achieved, see Fig. 11. Furthermore, significant changes of the set-point values $\mathbf{y}^{l,d}$ occur at approximately 0.5 h, 3.75 h, 6 h, 10.5 h, and 11.75 h, which constitute critical scenarios. However, the proposed control concept is able to handle all these different operating conditions. In Fig. 11a), the fast response characteristic of the DFF can be observed, where a nearly instantaneous heating of the strip can be realized. This aspect demonstrates the fast heating of the strip in the considered CSAF, which is in contrast to a furnace without a DFF. Generally, the strip temperatures $\mathbf{y}$ are within their admissible range.

Figure 12 shows the mass flows of fuel of the heating zones a and d as well as the switching state s. Generally, the mass flow of heating zone a is near the upper bound (base load) and the mass flow of heating zone d operates close to the lower bound (energy consumption). Moreover, acceptable slew rates of the mass flows can be observed. As mentioned in Sec. 1.1, the mass flow of fuel to the heating zones a and b depends on the width of the strip. Between 4.25 h and 5.25 h, a strip with smaller width is processed and otherwise, a strip with larger width is considered, see Fig. 10. Therefore, the boundaries of the mass flow $\dot{M}_{\alpha}^{f,\pm}$

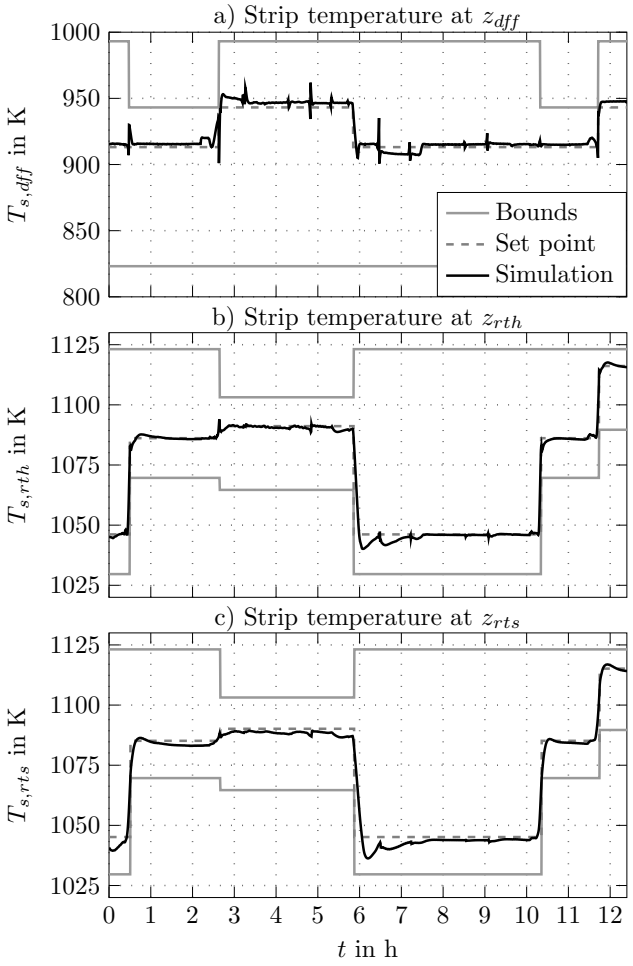

Figure 11: Strip temperature at the three pyrometers P_{dff}, P_{rth}, and P_{rts}.

vary, $\alpha \in \{hza, hzb\}$. Figure 12c) shows the switching state s. When the heating zone d is switched off, the corresponding mass flow $\dot{M}_{hzd}^{f}$ of fuel is zero, see Fig. 12b). Figure 13 shows the mass flow $\dot{M}_{rth1}^{f}$ of fuel to the zone rth 1 and the corresponding bounds. Generally, the mass flows depend on the set-point strip temperatures $\mathbf{y}^d$ and the strip properties. At 0.5 h, for instance, the thickness d_s of the strip decreases (cf. Fig. 10), the set-point strip temperature $T_{s,dff}^{d}$ remains constant (cf. Fig. 11a)), and the mass flows $\mathbf{m}_D$ of fuel of the DFF decrease. In the IFF, however, the set-point strip temperatures $T_{s,rth}^{d}$ and $T_{s,rts}^{d}$ increase and thus the mass flows $\mathbf{m}_I$ of fuel also increase. All mass flows are within their admissible range.

Figure 14 shows the strip velocity trajectory and its lower bound. As discussed in Sec. 3.2 (cf. Fig. 8), the velocity is constant during the fillet of the strip and a change occurs only at the strip transitions.

To demonstrate the necessity of the OTP, which provides the target trajectories $\mathbf{y}^r$ of the strip temperature, two simulation scenarios are considered. In the first scenario,

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

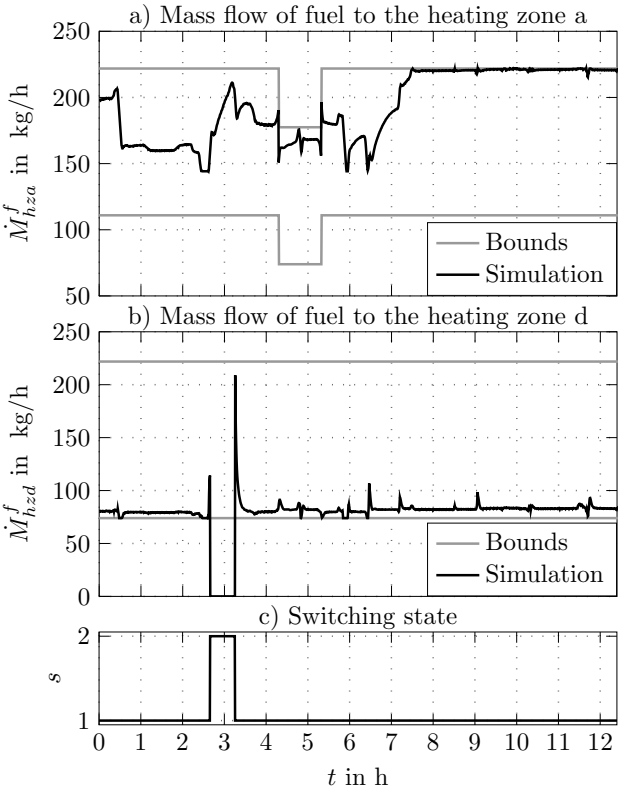

Figure 12: Mass flow of fuel of the heating zone a and d as well as the switching state.

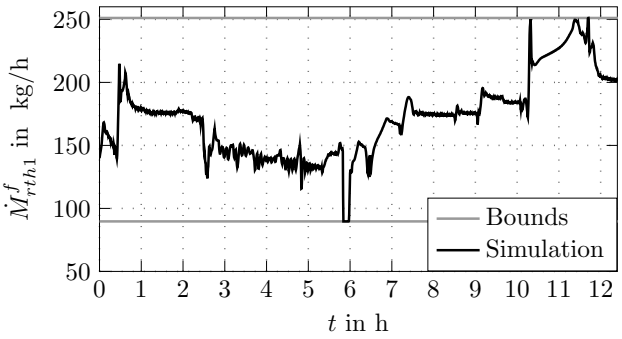

Figure 13: Mass flow of fuel of the heating zone rth 1.

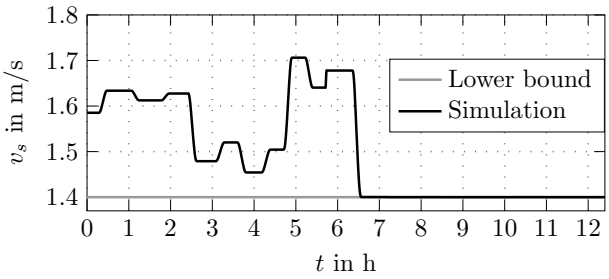

Figure 14: Strip velocity.

target trajectories are used in the TR. In the second case, they are not considered and the set-point strip temperatures $\mathbf{y}^d$ are directly fed to the TR. Figure 15 shows a comparison of these scenarios. Here, the set-point strip temperatures $\mathbf{y}^d$ and the bounds $\mathbf{y}^{\pm}$ vary at approximately 10.35 h, see Fig. 15. Using the target trajectories $\mathbf{y}^r$ in the TR, an optimal tracking of the strip temperature without a violation of the bounds is achieved. In the second scenario, a violation of the upper bound at z_{dff} and of the lower bound at z_{rts} cannot be avoided due to the short prediction horizon and the high thermal inertia of the CSAF. Figure 15a) indicates that the TR in the second scenario uses the power of the DFF for realizing the desired temperature change, whereas the strip temperature remains almost constant in the first scenario. This motivates the use of the OTP.

The stability of the control concept is not explicitly proven for the system under consideration. In (Graichen and Kugi, 2010), the stability of a suboptimal model predictive control is shown for a general nonlinear system under some mild conditions. The considered TR is also based on a suboptimal approach and issues with the numerical stability have not been observed during all the extensive tests.

A standard PC (3.6GHz, 8.0GB Ram) was used for simulating the proposed control concept. Determining one optimal solution in the OTP requires approximately 25 s CPU-time and for the TR, approximately 4 s CPU-time is required. If all heating zones are active, the OTP uses 1500 optimization variables and the TR 300 optimization variables.

6. Conclusions and outlook

A nonlinear optimization-based controller for a continuous strip annealing furnace was developed. It is based on a tailored mathematical model of the underlying process, which features both real-time capability and high accuracy. The control strategy is based on a hierarchical concept to master the challenges of mixed-integer programming and real-time execution. A static optimization is used to calculate an optimal operating point for each strip, where the switching state (on/off) of the heating zones c and d and the strip velocity are determined. A trajectory generator calculates the time evolution of the strip velocity when a new strip enters the furnace. Based on a trajectory planner, target trajectories of the strip temperature and the optimal switching times are calculated. Finally, the mass flow of fuel is optimized by a temperature regulator, which is based on a model predictive control strategy.

Based on the control objectives, the constraints, and future knowledge of material properties, a finite-time optimization problem was defined. Since the existence of a solution of such a problem generally cannot be guaranteed, it was transformed into an unconstrained problem, which is solved by means of the Gauss-Newton method.

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

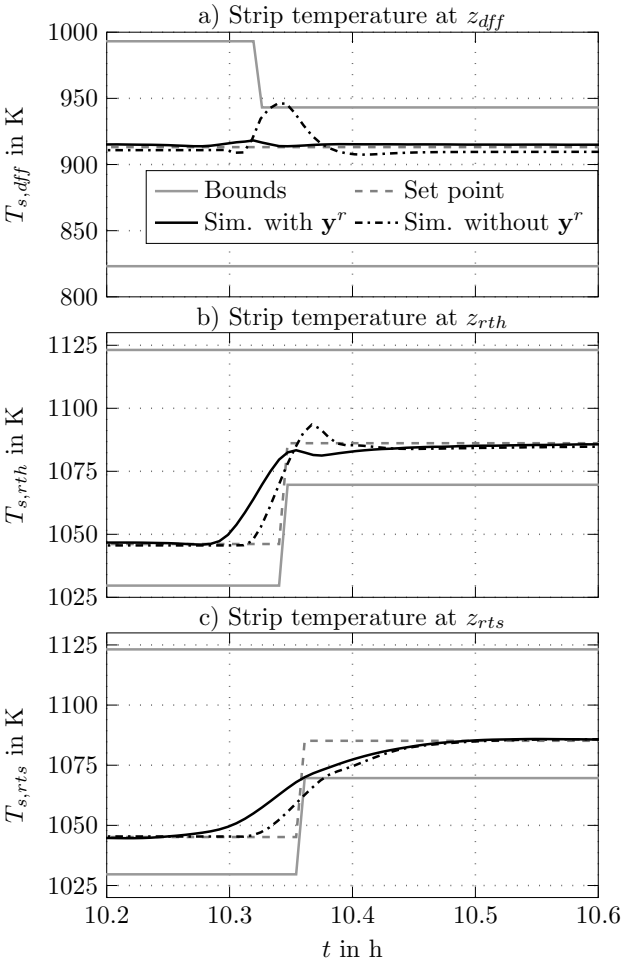

Figure 15: Strip temperature at the three pyrometers P_{dff}, P_{rth} and P_{rts} simulated with and without target trajectories $\mathbf{y}^r$.

In a simulation scenario for an experimentally validated model of a continuous strip annealing furnace at voestalpine Stahl GmbH, Linz, Austria, the good performance of the proposed control concept is demonstrated. The simulations also illustrate the need of a hierarchical control strategy.
Due to the encouraging results achieved by the proposed control concept, an implementation on the real plant is currently carried out.

Acknowledgement

This research was partially supported by the Austrian Research Promotion Agency (FFG), grant number: 834305. Moreover, the authors kindly express their gratitude to the industrial research partners voestalpine Stahl GmbH and Andritz AG.

References

Ascher, U. and Petzold, L. (1998). *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia.

Baehr, H. and Stephan, K. (2006). *Heat and Mass Transfer*. Springer-Verlag: Berlin, Heidelberg, 2^{nd} edition.

Bertsekas, D. (1999). *Nonlinear Programming*. Belmont, Massachusetts: Athena Scientific, Chengdu, China, 2^{nd} edition.

Bitschnau, L., Jakubek, S., and Kozek, M. (2010). Constrained model predictive control of a continuous annealing furnace. In *Inproceedings of the ASME 2010 Dynamic Systems and Control Conference (DSCC)*, 1–8. Cambridge, USA.

Bitschnau, L. and Kozek, M. (2009). Modeling and control of an industrial continuous furnace. In *Proceedings of the International Conference on Computational Intelligence, Modelling and Simulation (CIMSiM)*, 231–236. Brno, Czech Republic.

Buddenberg, J. and Wilke, C. (1949). Calculation of gas mixture viscosities. *Industrial & Engineering Chemistry Research*, 41(7), 1345–1347.

Conn, A., Gould, N., and Toint, P. (2000). *Trust-Region Methods*, volume 1 of *MPS/SIAM Series on Optimization*. SIAM.

de Pisón, F.J.M., Celorrio, L., de-la Parte, M.P., and Castejón, M. (2011). Optimising annealing process on hot dip galvanising line based on robust predictive models adjusted with genetic algorithms. *Ironmaking & Steelmaking*, 38(3), 218–228.

Delaunay, D. (2007). Direct fired pre-heating furnace design and control of strip annealing lines for a wide range of steel grades production. *Metallurgical Plant and Technology*, 3, 128–134.

Dunoyer, A., Burnham, K., Heeley, A., and Marcroft, S. (1998). Control of continuously operated high temperature furnaces. In *Proceedings of the International Conference on Control*, volume 1, 422–427. Swansea, UK.

Flasskamp, K., Murphey, T., and Ober-Bloebaum, S. (2012). Switching time optimization in discretized hybrid dynamical systems. In *Proceedings of the 51^{st} IEEE Conference on Decision and Control (CDC)*, 707–712. Maui, USA.

Fletcher, C. (1984). *Computational Galerkin Methods*. Springer-Verlag: Berlin, Heidelberg.

Froehlich, C., Strommer, S., Steinboeck, A., Niederer, M., and Kugi, A. (2016). Modeling of the media-supply of gas burners of an industrial furnace. *IEEE Transactions on Industry Applications*, 52(3), 2664–2672.

Graichen, K. and Käpernick, B. (2012). *Frontiers of Model Predictive Control*, chapter A Real-Time Gradient Method for Nonlinear Model Predictive Control, 10–28. InTech.

Graichen, K. and Kugi, A. (2010). Stability and incremental improvement of suboptimal MPC without terminal contraints. *IEEE Transactions on Automatic Control*, 55(11), 2576–2580.

Graichen, K. and Petit, N. (2008). Constructive methods for initialization and handling mixed state-input constraints in optimal control. *Journal of Guidance, Control, and Dynamics*, 31(5), 1334–1343.

Grossmann, I. and Kravanja, Z. (1997). *Mixed-integer nonlinear programming: A survey of algorithms and applications*, volume 93, chapter The IMA Volumes in Mathematics and its Applications, Large-Scale Optimization with Applications. Part II: Optimal Design and Control, 73–100. Springer-Verlag: New York, Berlin.

Guo, C., Zhang, Y., You, X., Chen, X., and Zhang, Y. (2009). Optimal control of continuous annealing process using PSO. In *Proceedings of the IEEE International Conference on Automation and Logistics (ICAL 09)*, 602–605. Shenyang, China.

Imose, M. (1985). Heating and cooling technology in the continuous annealing. *Transactions of the Iron and Steel Institute of Japan*, 25(9), 911–932.

Incropera, F., DeWitt, D., Bergmann, T., and Lavine, A. (2007). *Fundamentals of Heat and Mass Transfer*. John Wiley & Sons: Hoboken, New Jersey, 6^{th} edition.

Iuchi, T., Yamada, Y., Sugiura, M., and Torao, A. (2010). Chapter 4 Thermometry in steel production. *Radiometric Temperature*

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

Measurements: II. Applications. Series: Experimental Methods in the Physical Sciences, 43, 217–277.

Katsuki, M. and Hasegawa, T. (1998). The science and technology of combustion in highly preheated air. In *Proceedings of the 27th Symposium on Combustion*, volume 27, 3135–3146. Boulder, USA.

Kaya, C. and Noakes, J. (2003). Computational method for time-optimal switching control. *Journal of Optimization theory and Applications*, 117(1), 69–92.

Keerthi, S. and Gilbert, E. (1985). An existence theorem for discrete-time infinite-horizon optimal problems. *IEEE Transactions on Automatic Control*, 30(9), 907–909.

Kelly, C., Watanapongse, D., and Gaskey, K. (1988). Application of modern control to a continuous anneal line. *IEEE Control Systems Magazine*, 8(2), 32–37.

Lewis, D., Sandoz, D., Norberg, P.O., and Warwick, K. (1994). The application of predictive control to steel annealing. In *Proceedings of the International Conference on Control*, volume 1, 692–698. Coventry, UK.

Li, X.B., Liu, D., Jian, S.B., and Guo, J.X. (2004). Intelligent PID control system for vaccum annealing furnace workpieces temperature. In *Proceedings of the 3rd International Conference on Machine Learning and Cybernetics*, 934–940. Shanghai, China.

Martineau, S., Burnham, K., Haas, O., Andrews, G., and Heeley, A. (2004). Four-term bilinear PID controller applied to an industrial furnace. *Control Engineering Practice*, 12(4), 457– 464.

Michalski, L., Eckersdorf, K., Kucharski, J., and McGhee, J. (2001). *Temperature Measurement*. John Wiley & Sons: Hoboken, New Jersey, 2nd edition.

Ming, Z., Datai, Y., and Jiangang, Z. (2008). A new strip temperature control method for the heating section of continuous annealing line. In *Proceedings of the IEEE Conference on Cybernetics and Intelligent Systems*, 861–864. Chengdu, China.

Moe, J. (1962). Design of water-gas shift reactors. *Chemical Engineering Progress*, 58(3), 33–36.

Mould, P. (1982). An overview of continuous annealing technology for steel sheet products. *Journal of Minerals, Metals and Materials Society*, 34(5), 18–28.

Mullinger, P. and Jenkins, B. (2014). *Industrial and Process Furnaces: Principles, Design and Operation*. Butterworth-Heinemann, Amsterdam, 2 edition.

Niederer, M., Strommer, S., Steinboeck, A., and Kugi, A. (2014). A simple control-oriented model of an indirect-fired strip annealing furnace. *International Journal of Heat and Mass Transfer*, 78, 557–570.

Niederer, M., Strommer, S., Steinboeck, A., and Kugi, A. (2015). A mathematical model of a combined direct- and indirect-fired strip annealing furnace. In *Proceedings of the 10th International Conference on Zinc and Zinc Alloy Coated Steel (GALVATECH), Toronto, Canada*, 137–144.

Niederer, M., Strommer, S., Steinboeck, A., and Kugi, A. (2016). Nonlinear model predictive control of the strip temperature in an annealing furnace. *Journal of Process Control*, 48, 1–13.

Nocedal, J. and Wright, S. (2006). *Numerical Optimization*. Springer series in operations research. New York: Springer, 2nd edition.

Norberg, P.O. (1997). Challenges in the control of the reheating and annealing process. *Scandinavian Journal of Metallurgy*, 26, 206–2014.

Paulus, P. and Laval, P. (1985). Study of the heat buckling and shape problems in continuous heat treating lines and discussion of proposed solutions. In *Proceedings of the Conference Technology of Continuously Annealed Cold-Rolled Sheet Steel*, 419–439. Detroit, USA.

Sasaki, T., Hira, T., Abe, H., Yanagishima, F., Shimoyama, Y., and Tahara, K. (1984). Control of strip buckling and snaking in continuous annealing furnace. Technical report 9, Kawasaki Steel. 36–46.

Sawaragi, Y., Nakayama, H., and Tanino, T. (1985). *Theory of Multiobjective Optimization*. Academic Press, Orlando, Florida.

Steinboeck, A., Wild, D., and Kugi, A. (2013). Nonlinear model predictive control of a continuous slab reheating furnace. *Control Engineering Practice*, 21(4), 495–508.

Strikwerda, J. (2004). *Finite Difference Schemes and Partial Differential Equations*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2nd edition.

Strommer, S., Froehlich, C., Niederer, M., Steinboeck, A., and Kugi, A. (2017). Modeling and control of the oxygen concentration in a post combustion chamber of a gas-fired furnace. In *Proceedings of the 20th IFAC World Congress*, 14330–14335. Toulouse, France.

Strommer, S., Niederer, M., Steinboeck, A., Jadachowski, L., and Kugi, A. (2016). Nonlinear observer for temperature and emissivities in a strip annealing furnace. In *Proceedings of the 51st IEEE Industry Applications Society (IAS) Annual Meeting*, 1–8. Portland, USA.

Strommer, S., Niederer, M., Steinboeck, A., and Kugi, A. (2013). Analysis of energy comsumption in a direct-fired continuous strip annealing furnace. In *Proceedings of the 9th International and 6th European Rolling Conference*, 1–14. AIM, Italian Society for Metallurgy, Venice, Italy.

Strommer, S., Niederer, M., Steinboeck, A., and Kugi, A. (2014a). A mathematical model of a direct-fired continuous strip annealing furnace. *International Journal of Heat and Mass Transfer*, 69, 375–389.

Strommer, S., Steinboeck, A., Begle, C., Niederer, M., and Kugi, A. (2014b). Modeling and control of gas supply for burners in gas-fired industrial furnaces. In *Proceedings of the IEEE Conference on Control Applications (CCA)*, 210–215. Antibes, France.

Trinks, W., Mawhinney, M., Shannon, R., Reed, R., and Garvey, J. (2004). *Industrial Furnaces*. John Wiley & Sons.

Turns, S. (2006). *An Introduction to Combustion*. McGraw-Hill Book Company: Singapore, 2nd edition.

Ueda, I., Hosoda, M., and Taya, K. (1991). Strip temperature control for a heating section in cal. In *Proceedings of the IEEE International Conference on Industrial Electronics, Control and Instrumentation (IECON)*, volume 3, 1946–1949. Kobe, Japan.

Wu, H., Speets, R., Heeremans, F., Driss, O.B., and van Buren, R. (2014). Non-linear model predictive control of throughput and strip temperature for continuous annealing line. *Ironmaking & Steelmaking*, 42(8), 570–578.

Yahiro, M., Shigemori, H., Hirohata, K., Ooi, T., Haruna, M., and Nakanishi, K. (1993). Development of strip temperature control system for a continuous annealing line. In *Proceedings of the IEEE International Conference on Industrial Electronics, Control, and Instrumentation (IECON)*, volume 1, 481–486. Maui, USA.

Yoshitani, N. and Hasegawa, A. (1998). Model-based control of strip temperature for the heating furnace in continuous annealing. In *IEEE Transactions on Control Systems Technology*, volume 6, 146–156.

Post-print version of the article: S. Strommer, M. Niederer, A. Steinboeck, and A. Kugi, "Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace", *Control Engineering Practice*, vol. 73, pp. 40–55, 2018. DOI: 10.1016/j.conengprac.2017.12.005