Nonlinear model predictive control of the strip temperature in an annealing furnace

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

and published in *Journal of Process Control*.

Contact:

Automation and Control Institute (ACIN)
Vienna University of Technology
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

Nonlinear model predictive control of the strip temperature in an annealing furnace

M. Niederer*, S. Strommer, A. Steinboeck, A. Kugi

Automation and Control Institute, Vienna University of Technology, Gußhausstraße 27–29, 1040 Vienna, Austria

Abstract

A nonlinear model predictive controller is designed for the strip temperature in a combined direct- and indirect-fired strip annealing furnace. Based on a tailored first-principles dynamical model and the estimated current system state, the receding horizon controller selects optimal trajectories for both the fuel supply and the strip velocity so that the strip temperature is controlled to its desired target temperature. The controller additionally maximizes the throughput and minimizes the energy consumption. In the control algorithm, the dynamic optimization problem with equality constraints is numerically solved by using the Gauss-Newton method. The gradient and the approximated Hessian matrix of the objective function are analytically computed using an adjoint-based method. The capabilities of the proposed controller are demonstrated for a validated high-fidelity simulation model of an industrial annealing furnace.

Key words: steel industry, annealing furnace, direct- and indirect-fired sections, nonlinear model predictive control, receding horizon control, vector-valued Lagrangian, adjoint-based method, Gauss-Newton method

1. Introduction

1.1. Control problem

In the steel industry, strip annealing furnaces are used for the heat treatment of steel strips in order to achieve the desired metallurgical and surface properties for subsequent process steps. The annealing furnace considered in this paper, cf. Fig. 1, is part of a strip processing line of voestalpine Stahl GmbH in Linz, Austria and contains 235 m steel strip. The key parameters of the furnace are tabulated in Tab. 1. To ensure a continuous operation of the processing line, the strips are welded together to form an endless strip.

Throughput of steel	45.7 t/h
Nominal heating power	15 MW
Strip dimensions	
Thickness	0.35 – 1.2 mm
Width	800 – 1 640 mm
Strip velocities	max. 180 m/min
Number of gas burners	48
Number of radiant tubes	62
Number of heating zones	7

Table 1: Nominal parameters of the strip annealing furnace.

To meet the high demands on the quality of the final product, the temperature evolution of the strip is of importance. While the strip moves through the furnace, it has to be heated to a predefined target temperature. This strip temperature control task is a challenge mainly for the following reason: An ongoing diversification of the product portfolio essentially prevents steady-state furnace operation. The furnace can be considered as a cascade thermal system, where the strip is part of the last cascade. Since the thermal inertia of the strip is significantly lower than that of the furnace, it is difficult to consistently realize the desired target temperature in transient operational situations.

In addition to the product quality, there are further demands on the furnace operation like optimized energy consumption, material throughput, and CO_2 emissions. A furnace temperature controller that meets all these requirements is still an open research issue. This paper, therefore, explores the suitability of advanced nonlinear control and optimization methods for the considered control problem.

1.2. Existing solutions

In the literature, different concepts for strip temperature control of annealing furnaces can be found. The following summary should therefore only serve as a starting point for an in-depth exploration.

Over decades, simple PID control concepts were used for strip temperature control. PID control does not require a mathematical model and provides acceptable results for a steady-state furnace operation. However, if the furnace is not operated at steady state, this concept is no longer suitable because of the large thermal inertia of the furnace.

Occasionally, rule-based expert systems and fuzzy logic control concepts are used, see, e.g., [1, 2, 3, 4, 5]. Generally, these semi-empirical concepts use a wide range of measurements from the plant in order to characterize the operation conditions and adjust the inputs to obtain the desired temperature. Basically, they mimic the human operator but they perform the control task with greater consistency and accuracy and smaller

*Corresponding author. *Tel.:* +43 1 58801 376292, *Fax:* +43 1 58801 9376264.

Email address: niederer@acin.tuwien.ac.at (M. Niederer)

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

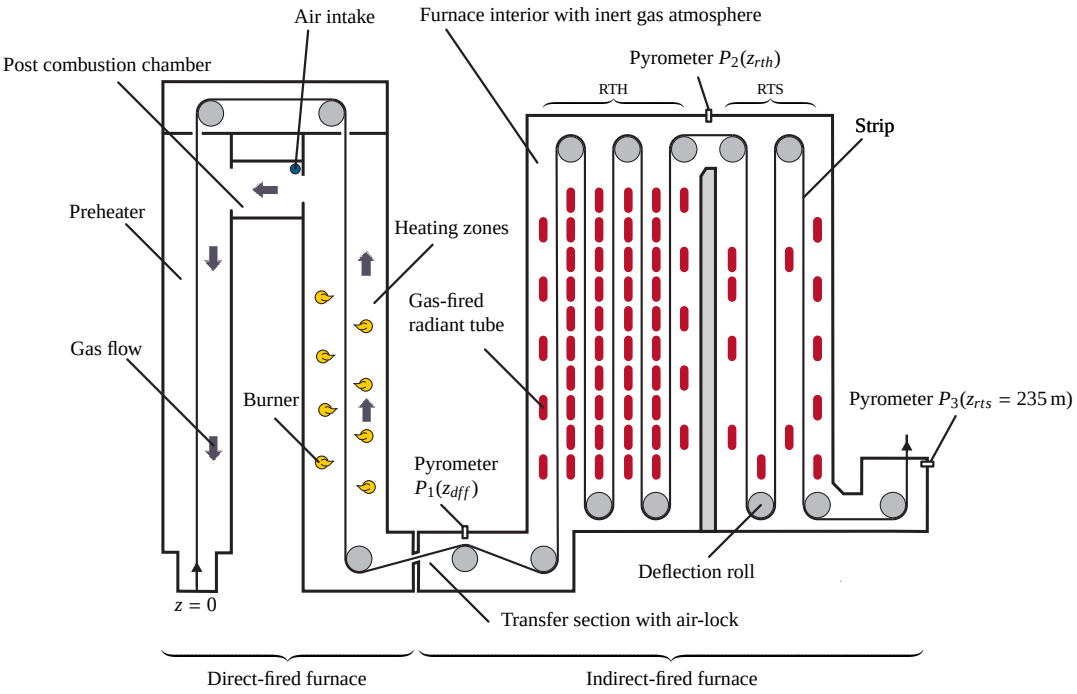

Figure 1: Combined direct- and indirect-fired strip annealing furnace.

response times than humans. Due to the complexity and wide variety of operation situations of an annealing furnace, these control concepts are, however, expensive to install, commission, and tune.

A hierarchical control concept is presented in [6]. In the topmost layer, set-point values for the line speed and the zone temperatures needed to achieve the target strip temperature are determined. In the intermediate layer, the switching times of the individual set-point values are determined using a simple model of the strip temperature and optimization based methods. In the inner-most layer, simple line speed and temperature controllers control the strip temperature and ensure that it is within defined bounds.

A hierarchically structured model-based control concept is also presented in [7]. The underlying semi-analytic furnace model is based on physical principles and measured system dynamics. In the higher control layer, reference trajectories for both the strip velocity and the strip temperature are generated. In the lower control layer, unknown parameters of the model are recursively estimated and the mass flows of fuel are determined using generalized predictive control [8, 9].

A control concept without a hierarchical structure is presented in [10]. Based on a linear model of the strip temperature, a model predictive controller for the heating power of the furnace is developed. It also takes into account input constraints. The constrained optimization problem is solved by using quadratic programming [11].

The controller presented in [12] is similar to that presented in [10] but is based on a simple nonlinear furnace model. For the utilization in a linear model predictive control concept, the furnace model is linearized. In addition to the heating power of the furnace, the controller optimizes the throughput, i.e., the line speed.

1.3. Motivation, objectives, and contributions of this work

Nonlinear model predictive control is an appropriate concept for strip temperature control of the considered annealing furnace. It is a versatile optimization-based and anticipative control method that is suitable for complex nonlinear multiple-input multiple output systems. Moreover, it allows the incorporation of various control objectives and the systematic consideration of input constraints, state constraints, and known disturbances, e.g., strip changes.

Most developed model predictive controllers for strip temperature control of annealing furnaces are based on linear system models. Though a linear model simplifies the control law, it may limit the capability and the accuracy of the obtained controller. Moreover, existing strip temperature controllers are mainly developed for indirect-fired annealing furnaces. However, the annealing furnace considered in this paper features also a direct-fired section. Since the direct- and the indirect-fired furnace sections are physically coupled by the moving steel strip, an integrated controller design for both furnace sections is recommendable. For this reason, existing control strategies cannot be directly transferred to the annealing furnace considered here. This motivates the tailored design of a nonlinear model predictive controller for non-steady-state furnace operation. It should realize the following control objectives:

- Accurate strip temperature control
- Maximization of the throughput of the steel strip
- Minimization of the fuel consumption

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

The controller has to respect several constraints:

- Bounds on the strip temperature
- Bounds on the control inputs

To satisfy these requirements, the following control inputs are available:

- Mass flows of fuel to the heating zones
- Strip velocity

However, the controller design presented in this paper can also be used when the strip velocity is prescribed by subsequent process steps and is therefore not available as a control input.

The contributions of this paper can be summarized as follows:

- Nonlinear first-principles model of the plant
- Tailored real-time capable model predictive controller for a strip annealing furnace based on the Gauss-Newton method in combination with the analytical calculation of the gradient and the approximated Hessian matrix using an adjoint-based method
- Systematic optimization of the material throughput
- The feasibility of the proposed approach is demonstrated by simulation studies for a validated high-fidelity mathematical model of an industrial annealing furnace comprising a direct- and an indirect-fired strip annealing furnace section

1.4. Contents

The paper is organized as follows: Section 2 gives an overview of the mathematical furnace model that is used as a basis for the model predictive controller design. In Section 3, the furnace temperature control system is described. A formal description of the control objectives is given in Section 4. In Section 5, the optimal control problem is first specified and then a numerical solution algorithm is proposed. An industrial application example in Section 6 demonstrates the capability of the developed controller. Finally, Section 7 gives some conclusions.

2. Mathematical furnace model

As shown in Fig. 1, the considered annealing furnace consists of a direct- and an indirect-fired section, which are physically coupled by the moving strip. In the direct-fired furnace section (DFF), there are four heating zones, each equipped with a set of burners for natural gas. To prevent the strip from oxidation, the combustion in these zones is controlled to be fuel rich. Thus, the flue gas contains unburnt products which are oxidized in a post combustion chamber by adding fresh air. The heat released in this post combustion is used to preheat the strip in the preheater. The indirect-fired furnace section features an inert gas atmosphere to prevent the strip from oxidation and is divided into a radiant tube heating section (RTH) and a radiant tube soaking section (RTS). Both sections are equipped with W-shaped gas-fired radiant tubes that are grouped into three heating zones. The strip temperature in the indirect-fired furnace section is therefore mainly controlled by the surface temperature of the radiant tubes.

A detailed first-principles model of the considered annealing furnace can be found in [13, 14, 15]. This model includes different submodels for the flue gas, the radiant tubes, the wall, the strip, and the rolls. However, the furnace model used in this paper is a reduced model in terms of complexity, dimension, and computational effort. In the following, the main differences between the furnace models are briefly outlined:

- In the reduced furnace model, a coarser spatial discretization is used throughout the furnace. This leads to a reduction of the model order.
- The modeling of the combustion process of fuel inside the heating zones of the direct-fired furnace is simplified. Hence, the complexity of the model is reduced.
- The reduced model of a radiant tube inside the indirect-fired furnace features only one temperature state. This also leads to a reduction of the model order.
- The reduced furnace model includes only temperature states of representative (reduced) radiant tubes. The temperatures of the remaining radiant tubes are determined from these representative values based on identified mappings. This leads to a further reduction of the model order.

In the following, the most fundamental equations of the reduced furnace model are summarized.

2.1. Flue gas

In the heating zones of the direct-fired furnace, natural gas (CH_4) is burnt in a fuel rich combustion process, i.e., the excess air coefficient λ satisfies $\lambda < 1$. The corresponding stationary-reaction equation reads as

$$CH_4 + 2\lambda\,(O_2 + 3.76N_2) \longrightarrow \chi^c_{CO}CO + \chi^c_{H_2}H_2 + \chi^c_{CO_2}CO_2 + \chi^c_{H_2O}H_2O + \chi^c_{N_2}N_2.$$

Here, χ^c_ν with $\nu \in \mathcal{S}_\nu = \{CO, CO_2, H_2O, H_2, N_2\}$ is the corresponding number of moles. Taking into account the water-gas-shift reaction [16, 17]

$$CO + H_2O \rightleftharpoons CO_2 + H_2,$$

which reaches a chemical equilibrium if (T^c in Kelvin)

$$K(T^c) = \exp\left(4.33 - \frac{4577.8}{T^c}\right) = \frac{\chi^c_{CO}\chi^c_{H_2O}}{\chi^c_{CO_2}\chi^c_{H_2}}$$

is satisfied, the unknowns χ^c_ν can be determined from simple mol balances. T^c denotes the adiabatic flame temperature. Let M_ν denote the molar mass of a component $\nu \in \mathcal{S}_\nu \cup \mathcal{S}_\kappa$ with $\mathcal{S}_\kappa = \{CH_4, O_2, N_2\}$. The mass flows of the combustion products $\dot{m}^c_\nu$ with $\nu \in \mathcal{S}_\nu$ read therefore as

$$\dot{m}^c_\nu = \dot{m}^b_{CH_4}\frac{\chi^c_\nu M_\nu}{M_{CH_4}}. \tag{1}$$

Moreover, the mass flows of fuel $\dot{m}^b_{CH_4}$ and the combustion air $\dot{m}^b_{O_2} + \dot{m}^b_{N_2}$ are coupled in the form

$$\dot{m}^b_\kappa = \dot{m}^b_{CH_4}\frac{\chi^b_\kappa M_\kappa}{M_{CH_4}} \tag{2}$$

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

with $\kappa \in \mathcal{S}_\kappa$ and $(\kappa, \chi_\kappa^b) = \{(CH_4, 1), (O_2, 2\lambda), (N_2, 7.52\lambda)\}$.

In the direct-fired furnace, the strip is heated by the hot flue gas that is mainly characterized by its mass and temperature. For modeling the flue gas, the furnace is discretized into N_g volume zones, where each zone is assumed to be a well-stirred reactor. The volume zones are numbered in ascending order in flow direction of the flue gas. Neglecting the flue gas dynamics [13], the mass balance for each combustion product $\nu \in \mathcal{S}_\nu$ and the enthalpy balance of an individual volume zone i reads as

$$0 = \dot{m}_{\nu,i-1} + \dot{m}^c_{\nu,i} - \dot{m}_{\nu,i} \tag{3}$$

$$0 = \sum_{\nu\in\mathcal{S}_\nu} \dot{m}_{\nu,i-1} h_\nu(T_{g,i-1}) - \sum_{\nu\in\mathcal{S}_\nu} \dot{m}_{\nu,i} h_\nu(T_{g,i}) + \sum_{\kappa\in\mathcal{S}_\kappa} \dot{m}^b_{\kappa,i} h_\kappa(T_{\kappa,i}) + \dot{Q}_{g,i}. \tag{4}$$

Here, $\dot{m}_{\nu,i-1}$ and $\dot{m}_{\nu,i}$ denote the mass flow of a component $\nu \in \mathcal{S}_\nu$ that enters or leaves the volume zone i. For calculated values χ^c_ν, the unknown mass flows $\dot{m}_{\nu,i}$ with $\nu \in \mathcal{S}_\nu$ that are required in (4) can be easily determined by solving a linear set of equations that consists of (1), (2), and (3).

In (4), $h_\nu(T)$ is the specific enthalpy of a component $\nu \in \mathcal{S}_\nu \cup \mathcal{S}_\kappa$ at the temperature T. The first two terms in (4) denote the enthalpy flows associated with the incoming and the outgoing bulk flow in the respective volume zone. The enthalpy flow of the fuel and combustion air is described by the third term. Moreover, the net heat flow $\dot{Q}_{g,i}$ into the respective zone includes all thermal interactions of the flue gas with its environment. By solving the nonlinear set of equations (4), the flue gas temperatures $T_{g,i}$ of all volume zones $i \in \{1, \dots, N_g\}$ can be determined.

2.2. W-shaped radiant tube

Inside the indirect-fired furnace section, N_r W-shaped radiant tubes are grouped into several heating zones. All radiant tubes of a single heating zone are supplied with the same amount of fuel $\dot{m}_{CH_4}$ that can be continuously adjusted between a minimum and a maximum value. As shown in [14], the heat $\dot{Q}_{c,i}$ released by the combustion process inside a single radiant tube i and transferred into the tube wall can be described by the static mapping $\dot{Q}_{c,i} = \psi_i(\dot{m}_{CH_4})$. This relation can be obtained by an energy balance model of the radiant tube and measurement data from the real plant.

Let $g_i \dot{Q}_{c,i}$, with the weighting factor $g_i > 0$, be the heat input that is allocated to the first straight pipe of the radiant tube, cf. [14]. The pipe has the surface area S_r and the wall thickness d_r. Its material properties are the specific heat capacity c_r and the mass density ρ_r. Using a simple heat balance, the differential equation of the temperature state $T_{r,i}$ of the pipe reads as

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

with the (local) heat flux $\dot{q}_{c,i} = g_i \dot{Q}_{c,i}/S_r$ at the inner pipe surface and the radiative heat flux $\dot{q}_{r,i}$ from the furnace interior to the outer pipe surface. For the calculation of the radiative heat transfer inside the furnace, the temperature states of the remaining pipes $j = 2, 3, 4$ of the radiant tube are approximated by

$$T^j_{r,i} = a^j_i T_{r,i}.$$

i.e., it is assumed that the temperature ratios of the individual radiant tube pipes are constant. The coefficients $a^j_i \in [0, 1]$ were obtained in simulation studies.

2.3. Wall

The furnace wall consists of J different layers and is discretized into N_w wall segments along the furnace. To model the heat transfer through the wall, the Galerkin weighted residual method is employed to solve the one-dimensional heat conduction equation with a Dirichlet boundary condition at the outer surface and a Neumann boundary condition at the inner surface. If the stationary solution of a multi-layered furnace wall is used as trial function, the lumped-parameter model of the (inner) surface temperature $T_{w,i}$ of a wall segment i reads as [14]

$$\frac{\mathrm{d}}{\mathrm{d}t} T_{w,i} = \frac{1}{K_1} \dot{q}_{w,i} - \frac{K_2}{K_1} \left(T_{w,i} - T_o\right), \tag{6a}$$

with the abbreviations

$$K_1 = \left(\sum_{j=1}^{J} \frac{d_{w,j}}{k_{w,j}}\right)^{-2} \sum_{j=1}^{J} \frac{\rho_{w,j} c_{w,j} k_{w,j}}{3} \left(\left(\sum_{l=j}^{J} \frac{d_{w,l}}{k_{w,l}}\right)^3 - \left(\sum_{l=j+1}^{J} \frac{d_{w,l}}{k_{w,l}}\right)^3\right) \tag{6b}$$

$$K_2 = \left(\sum_{j=1}^{J} \frac{d_{w,j}}{k_{w,j}}\right)^{-1}. \tag{6c}$$

Here, $d_{w,j}$ denotes the thickness and the material properties $c_{w,j}$, $k_{w,j}$, and $\rho_{w,j}$ denote the heat capacity, the heat conductivity, and the mass density of the layer $j \in \{1, \dots, J\}$. Moreover, $\dot{q}_{w,i}$ is the heat flux into the inner wall surface due to thermal radiation and convection. T_o is the temperature at the outer wall surface which is assumed to be equal to the ambient temperature.

2.4. Strip and rolls

Assume that the strip enters the furnace with ambient temperature T_o and moves along the direction z with a velocity $v_s > 0$. Here, z is the spatially fixed length coordinate, cf. Fig. 1. Inside the furnace, the strip is heated to a desired target temperature. Using the first law of thermodynamics, the evolution of the strip temperature $T_s(z, t)$ in an Eulerian framework reads as

$$\frac{\partial T_s(z,t)}{\partial t} = \frac{2\dot{q}_s(z,t)}{\rho_s c_s(T_s(z,t)) d_s} - v_s \frac{\partial T_s(z,t)}{\partial z}. \tag{7}$$

Here, d_s is the strip thickness, ρ_s the mass density, c_s the temperature-dependent specific heat capacity, and $\dot{q}_s$ the heat flux into the strip surface due to conduction, radiation and convection.

Let the strip be spatially discretized into N_s sections of uniform length Δz with $N_s + 1$ grid points z_i. Using the upwind

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

scheme [18] for discretizing the transport term in (7), the strip temperature $T_{s,i}(t) = T_s(z_i, t)$ at a grid point z_i reads as

$$\frac{\mathrm{d}}{\mathrm{d}t}T_{s,i} = \frac{2\dot{q}_{s,i}}{\rho_s c_s(T_{s,i}) d_s} - v_s \frac{T_{s,i} - T_{s,i-1}}{\Delta z} \tag{8}$$

with $T_{s,0} = T_o$.

The strip is guided through the furnace by means of N_d rolls that are designed as hollow cylinders with the surface area S_d and the wall thickness d_d. The relevant material parameters are the heat capacity c_d and the mass density ρ_d. Using a simple heat balance, the lumped-parameter model of the roll i with the homogeneous temperature $T_{d,i}$ reads as

$$\frac{\mathrm{d}}{\mathrm{d}t}T_{d,i} = \frac{1}{\rho_d c_d(T_{d,i}) d_d} \left(\left(1 - \frac{S^c_{d,i}}{S_d}\right) \dot{q}_{d,i} + \frac{S^c_{d,i}}{S_d} \dot{q}^c_{d,i} \right). \tag{9}$$

Here, S^c_d denotes the contact area between the roll and the strip. The heat fluxes $\dot{q}_{d,i}$ and $\dot{q}^c_{d,i}$ capture radiation and convection with the furnace interior and conduction with the strip, respectively.

2.5. Heat transfer mechanisms

The individual submodels of the considered furnace, i.e., the models of the flue gas, the radiant tubes, the wall, the strip, and the rolls, are interconnected by the heat transfer mechanisms radiation, convection, and conduction. Due to the high temperatures inside the furnace, thermal radiation is the dominant mode of heat transfer. As suggested in [13, 14], the zone and the net-radiation method [19, 20, 21, 22] are used for analyzing the heat exchange by radiation. Both methods use a spatial discretization of the furnace to obtain a linear relation between the local radiative heat flows and the fourth powers of local surface and gas temperatures. Convective heat transfer between the flue gas and its surrounding surfaces is modeled by using Newton's law of cooling and dimensionless numbers from fluid dynamics [23, 24]. Furthermore, the conductive heat exchange at the roll-strip contact is described by means of thermal resistances.

2.6. Assembled dynamical system

2.6.1. Continuous-time system

For a compact notation of the assembled furnace model, the surface temperatures of the radiant tubes, the wall, the strip, and the rolls are summarized in the vector $\mathbf{T}_\Sigma \in \mathbb{R}^{N_\Sigma}$ with $N_\Sigma = N_r + N_w + N_s + N_d = 60$ and the flue gas temperatures are summarized in the vector $\mathbf{T}_g \in \mathbb{R}^{N_g}$ with $N_g = 13$. The mass flows of fuel into the heating zones are summarized in the vector $\dot{\mathbf{m}}$. The vector $\mathbf{u} = [v_s, \dot{\mathbf{m}}^\mathrm{T}]^\mathrm{T} \in \mathbb{R}^{N_u}$ with $N_u = 8$ thus contains all control inputs. In the following, the time transformation

$$\mathrm{d}t = \frac{L}{v_s}\mathrm{d}\tau \tag{10}$$

with an arbitrary length L, the strip velocity v_s, and the normalized time τ is used. In the normalized time domain, the assembled continuous-time state-space model of the furnace reads as

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{T}_\Sigma = \mathbf{f}_1(\tau, \mathbf{T}_\Sigma, \mathbf{T}_g, \mathbf{u}) \tag{11a}$$

$$\mathbf{0} = \mathbf{f}_2(\tau, \mathbf{T}_\Sigma, \mathbf{T}_g, \mathbf{u}) \tag{11b}$$

with the initial conditions $\mathbf{T}_\Sigma(\tau_0) = \mathbf{T}_{\Sigma,0}$ and $\mathbf{T}_g(\tau_0) = \mathbf{T}_{g,0}$. Equation (11a) includes the ordinary differential equations (5), (6), (8), and (9) for the radiant tubes, the wall, the strip, and the rolls, respectively. Equation (11b) represents a set of nonlinear equations for determining the flue gas temperatures, cf. (4). In (11), $\mathbf{T}_\Sigma$ are system states and $\mathbf{T}_g$ are algebraic variables. The main benefits of using the time transformation (10) are addressed in more detail in Section 2.6.2 and Section 5.1.

2.6.2. Discrete-time system

The implementation of the furnace model on a computer system requires the discretization of the time domain. Let τ_k with $k \in \mathbb{N}_0$ be the sampling points along the (normalized) time domain. The corresponding sampling period is $\Delta\tau_k = \tau_{k+1} - \tau_k$. The discrete-time representation of the semi-explicit differential algebraic system (11) can, for instance, be obtained by applying a 1-stage half-explicit Runge-Kutta method [25]. Thus, the discrete-time system reads as

$$\mathbf{T}_{\Sigma,k+1} = \mathbf{T}_{\Sigma,k} + \Delta\tau_k \mathbf{f}_{1,k}(\mathbf{T}_{\Sigma,k}, \mathbf{T}_{g,k}, \mathbf{u}_k) \tag{12a}$$

$$\mathbf{0} = \mathbf{f}_{2,k+1}(\mathbf{T}_{\Sigma,k+1}, \mathbf{T}_{g,k+1}, \mathbf{u}_{k+1}). \tag{12b}$$

For numerical stability of (12), the necessary condition

$$L\frac{\Delta\tau_k}{\Delta z} = v_s \frac{\frac{L\Delta\tau_k}{v_s}}{\Delta z} \leq 1 \tag{13}$$

has to be satisfied. Condition (13) is the *Courant-Friedrichs-Lewy* (CFL) condition [18] because $L\Delta\tau_k/v_s$ is the time step in the original time domain, cf. (10). The CFL condition generally describes the nexus between the spatial and temporal discretization of the advection equation (7) and the strip velocity. However, (13) includes no strip velocity due to the used time transformation (10). For a computational efficient furnace model, the realized time steps should always be as large as possible, i.e., the equality sign in (13) should hold if possible. Note that a given normalized time grid which satisfies the equality of (13) does not need to be adapted if the strip velocity changes. In view of model predictive control, where the numerical effort should be minimized, this is a benefit of using a normalized time domain, especially because the strip velocity is an optimization variable.

Introducing the vector $\mathbf{T}_k = \left[\mathbf{T}^\mathrm{T}_{\Sigma,k}, \mathbf{T}^\mathrm{T}_{g,k}\right]^\mathrm{T}$, (12) can be rewritten in the form

$$\mathbf{0} = \mathbf{F}_k(\mathbf{T}_{k+1}, \mathbf{T}_k, \mathbf{u}_{k+1}, \mathbf{u}_k) \tag{14a}$$

with

$$\mathbf{F}_k = \begin{bmatrix} \mathbf{T}_{\Sigma,k+1} - \mathbf{T}_{\Sigma,k} - \Delta\tau_k \mathbf{f}_{1,k}(\mathbf{T}_{\Sigma,k}, \mathbf{T}_{g,k}, \mathbf{u}_k) \\ \mathbf{f}_{2,k+1}(\mathbf{T}_{\Sigma,k+1}, \mathbf{T}_{g,k+1}, \mathbf{u}_{k+1}) \end{bmatrix} \tag{14b}$$

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

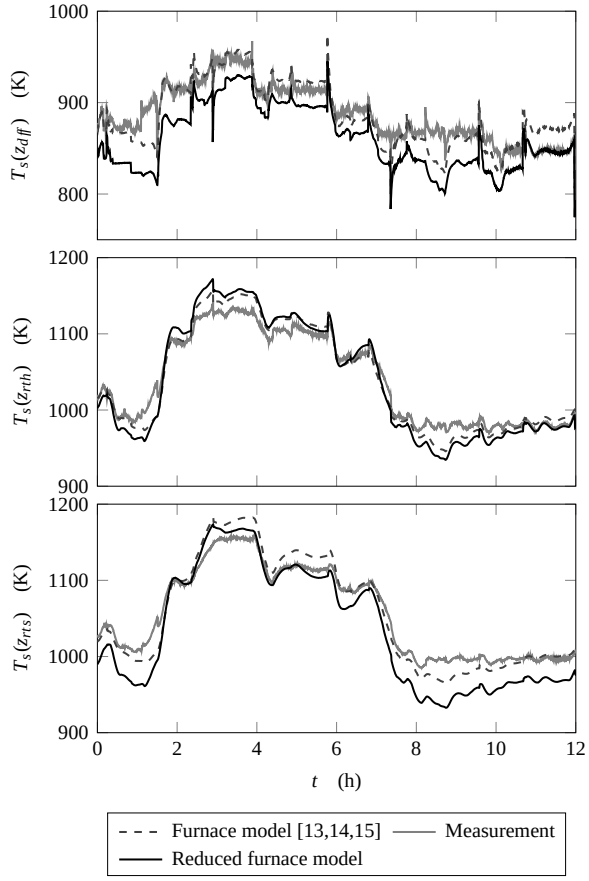

Figure 2: Strip temperatures at the pyrometer positions z_{dff}, z_{rth}, and z_{rts}.

and the initial condition $\mathbf{T}_0 = [\mathbf{T}_{\Sigma,0}^{\mathrm{T}}, \mathbf{T}_{g,0}^{\mathrm{T}}]^{\mathrm{T}}$.

The most important process quantity is the strip temperature which can be monitored by means of three radiation pyrometers, cf. Fig. 1, at the strip positions $z = z_i$ with $i \in \{dff, rth, rts\}$. The output function

$$\mathbf{h}_s(\mathbf{T}_k) \tag{15}$$

provides the simulated strip temperature at these positions. Equations (14) and (15) serve as the basis for the controller design in the following sections.

2.7. Validation of the reduced furnace model

Since the most important model output is the strip temperature at the pyrometer positions, the reduced furnace model is mainly validated by means of measurement data of these temperatures. The measurement data used for this validation stem from a measurement campaign conducted at the real plant. Most of the physical parameters used in the furnace model are at least roughly known from material handbooks or design drawings. For the unknown strip emissivity, the estimated value from [14] is used.

Fig. 2 shows the simulated and measured strip temperatures at the positions z_{dff}, z_{rth}, and z_{rts}. The simulated temperature evolution of the strip matches well the measurement data. It can be inferred from Fig. 2 that the furnace model presented in [13, 14, 15] is more accurate compared to the reduced model. However, the reduced model still captures the most important nonlinear effects and requires only 1.8 s CPU-time on a standard desktop PC (4 GHz, 16 GB RAM) for the simulation of 1 h furnace operation. This is 40 times faster than the model presented in [15]. Hence, the reduced furnace model is suitable for repeated evaluation in an optimization algorithm.

3. Furnace temperature control system

The main objective of the considered temperature control system is to ensure that the steel strips are heated to predefined target temperatures in order to achieve the desired material properties of the final product. As outlined in Fig. 3, the temperature control system consists of three hierarchical layers.

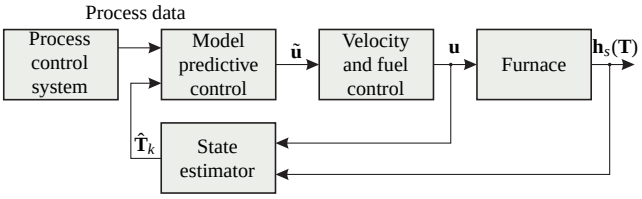

Figure 3: Hierarchical furnace temperature control system.

The topmost layer is the process control system that coordinates all tasks of the considered strip processing line. For example, it schedules the production sequence depending on priority and economic value of the individual strips. Moreover, it provides the data sets characterizing the strips (process data). These data set include target strip temperatures, maximum strip temperature limits, geometric dimensions, i.e., width, thickness, length, and material parameters like mass density and heat capacity of each strip.

In the intermediate control-layer, a model predictive controller selects reference trajectories $\tilde{\mathbf{u}}$ for the system inputs $\mathbf{u}$ so that the strip temperature is controlled to its desired target temperature. The design of the model predictive controller is the main objective of this paper and is described in detail in the following sections.

Most system states that are required as initial condition, i.e., feedback, for model predictive control are not measured. Therefore, $\mathbf{T}_k$ has to be estimated by means of a state estimator that is based on the furnace model (11). In addition to the system inputs, the estimator uses measurement data from the radiation pyrometers to improve the current state estimation $\hat{\mathbf{T}}_k$.

The lowermost layer includes decentralized control loops for realizing the desired control inputs, i.e., the strip velocity and the mass flows of fuel. In the current furnace configuration, this task is accomplished mainly by means of PI controllers. It is stressed that accurate fuel control is critical for safety reasons as well as for the product quality. Since the model predictive controller provides trajectories of future system inputs, efficient two-degrees-of-freedom control concepts may be used for fuel control, cf. [26, 27]. With such control concepts, the

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

furnace might be operated closer at its limits. For the following, it is assumed that $\mathbf{u} = \tilde{\mathbf{u}}$ holds, i.e., the inner control loops are considered ideal.

4. Control task

Starting with the strip temperature, a formal description of the control objectives and constraints is given. For the desired final product quality, the strip must be heated to its target temperature $T_{\tilde{s}}^d$ before reaching the end of the RTH section (pyrometer position z_{rth}). This temperature level should then be maintained up to the end of the RTS section (pyrometer position z_{rts}). The strip temperature at the end of the direct-fired furnace (pyrometer position z_{dff}) is of minor importance and, therefore, a target temperature does not exist.

As indicated in Fig. 4, the target temperature $T_{\tilde{s}}^d$ and the bounds $T_{\tilde{s}}^{\pm}$ on the strip temperature may vary from strip to strip and are defined using Lagrangian coordinates. The (local) Lagrangian coordinate $\tilde{z}$ is fixed for a given material point whereas the Eulerian coordinate z is spatially fixed. The mapping between Lagrangian and Eulerian coordinates is given by

$$z = \tilde{z} + \int_{t_0}^{t} v_s(\xi)\mathrm{d}\xi. \tag{16}$$

Utilizing the time transformation (10), (16) results in the linear mapping

$$z = \tilde{z} + L\,(\tau - \tau_0)\,, \tag{17}$$

which implies $z = \tilde{z}$ for $\tau = \tau_0$. The target temperature $T_{s,k}^d$ at a grid point τ_k described in Eulerian coordinates can thus be easily determined by using (17).

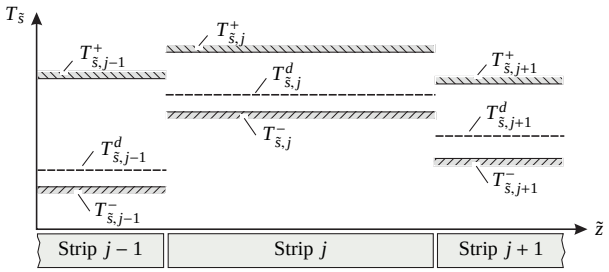

Figure 4: Target temperatures $T_{\tilde{s}}^d$ and temperature limits $T_{\tilde{s}}^{\pm}$ for different strips.

In non-steady-state furnace operation, e.g., when a welded joint that connects two strips with different target temperatures moves through the furnace, the desired target temperatures at the pyrometer positions z_{rth} and z_{rts} generally can not be realized simultaneously for both strips, at least not in the area of the welded joint. However, the temperatures of both strips must remain within certain limits, see Fig. 4. Hence, the local strip temperature should obey

$$T_{s,k}^-(z_{rth}) \le T_{s,k}(z_{rth}) \le T_{s,k}^+(z_{rth}) \quad \forall k \tag{18}$$

$$T_{s,k}^-(z_{rts}) \le T_{s,k}(z_{rts}) \le T_{s,k}^+(z_{rts}) \quad \forall k. \tag{19}$$

The values $T_{s,k}^{\pm}$ at the positions z_{rth} and z_{rts} can be determined by using (17). Moreover, the strip temperature at the pyrometer position z_{dff}, i.e., at the end of the direct-fired furnace, is constrained by

$$T_{s,k}^-(z_{dff}) \le T_{s,k}(z_{dff}) \le T_{s,k}^+(z_{dff}) \quad \forall k \tag{20}$$

in order to avoid strip damage, e.g., heat buckling, if the strip temperature is raised too fast. The limits $T_s^{\pm}(z_{dff})$ are defined similarly to the temperature limits shown in Fig. 4. For a single strip section, a typical heat-up curve with the desired target temperature T_s^d and corresponding temperature bounds $T_s^{\pm}$ is shown in Fig. 5. Introducing the vector $\mathbf{T}_s^{\pm} = [T_s^{\pm}(z_i)]_{i=dff,rth,rts}$, the constraints (18), (19), and (20) can be summarized in the form

$$\mathbf{T}_{s,k}^- \le \mathbf{h}_s(\mathbf{T}_k) \le \mathbf{T}_{s,k}^+ \quad \forall k, \tag{21}$$

with the output function $\mathbf{h}_s(\mathbf{T}_k)$ from (15).

At this point it should be noted that further control objectives can be defined at any position inside the furnace. This is because the state estimator provides the full system state, in particular the strip temperature profile along the furnace. If predefined heating trajectories are given for each strip, for example, the deviation between the desired and the estimated temperature profile can be taken into account in the control problem.

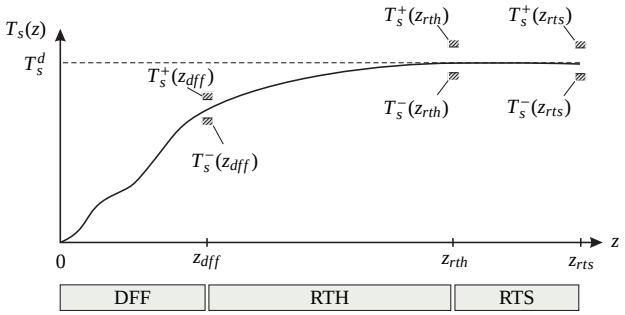

Figure 5: Temperature trajectory of a strip section moving through the furnace with target temperature T_s^d and corresponding temperature limits $T_s^{\pm}$ at the positions z_{dff}, z_{rth}, and z_{rts}.

In addition to the primary objective of accurate strip temperature control, the maximization of throughput of steel strip and the minimization of energy consumption are secondary control objectives. Generally, the maximization of throughput $\rho_s d_s b_s v_s$ is tantamount to the maximization of the strip velocity. Furthermore, the minimization of the energy consumption requires to minimize the fuel flow supplied to the heating zones of the annealing furnace. Because these three control objectives can be antagonistic, they are weighted in the optimal control problem according to their importance.

The control inputs, i.e., the mass flows of fuel and the strip velocity, are constrained in terms of both their absolute value and their slope. Hence,

$$\mathbf{u}_k^- \le \mathbf{u}_k \le \mathbf{u}_k^+, \quad \forall k \tag{22}$$

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

$$\dot{\mathbf{u}}_k^- \leq \underbrace{\frac{v_{s,k}}{L}\frac{\mathbf{u}_k - \mathbf{u}_{k-1}}{\Delta\tau_k}}_{\mathbf{h}_{u,k}(\mathbf{u}_k,\mathbf{u}_{k-1})} \leq \dot{\mathbf{u}}_k^+, \quad \forall k \tag{23}$$

must be satisfied in the discrete-time domain. Here, the slope is approximated by the difference quotient. The scaling factor $v_{s,k}/L$ is due to the time transformation (10).

5. Model predictive control

In this section, a model predictive controller for the strip temperature is developed. The basic idea of model predictive control is to recurrently solve a dynamic optimization problem defined on a finite time horizon. At the beginning of every time horizon, the current system state estimated by the state estimator is used as initial state of the optimization problem. This is the essential source of feedback used in model predictive control.

The challenge in the implementation of a model predictive controller is generally the computationally expensive solution of the dynamic optimization problem in real-time. This is the case especially for complex and high-dimensional nonlinear systems like annealing furnaces. One option to overcome this challenge is that the optimization problem is just approximately solved in order to reduce the computation time. This approach leads to a suboptimal solution of the optimization problem and may jeopardize the control accuracy. However, it facilitates a real-time implementation of the model predictive controller and is therefore used in this paper.

5.1. Constrained optimization problem

For the finite time horizon $(\tau_{k_0}, \tau_{k_1}]$ with $\tau_{k_1} - \tau_{k_0} = 1$, the control objectives are implemented by the objective function

$$\sum_{k=k_0+1}^{k_1} \|\mathbf{r}_{1,k}(\mathbf{h}_s(\mathbf{T}_k), \mathbf{u}_k)\|_{\mathbf{W}_{1,k}} + \|\mathbf{r}_{2,k}(\mathbf{u}_k)\|_{\mathbf{W}_{2,k}} \tag{24a}$$

with

$$\mathbf{r}_{1,k}(\mathbf{h}_s(\mathbf{T}_k), \mathbf{u}_k) = \sqrt{\frac{L}{v_{s,k}}} \begin{bmatrix} T^d_{s,k}(z_{rth}) - T_{s,k}(z_{rth}) \\ T^d_{s,k}(z_{rts}) - T_{s,k}(z_{rts}) \end{bmatrix} \tag{24b}$$

$$\mathbf{r}_{2,k}(\mathbf{u}_k) = \sqrt{\frac{L}{v_{s,k}}} \begin{bmatrix} \sqrt{\frac{1}{v_{s,k}}} \\ \dot{\mathbf{m}}_k \end{bmatrix} \tag{24c}$$

and the quadratic form $\|\mathbf{r}\|_{\mathbf{W}} = \mathbf{r}^{\mathrm{T}}\mathbf{W}\mathbf{r}$ where $\mathbf{W}$ is a (diagonal) positive semi-definite weighting matrix. The factor $L/v_{s,k}$ in each term of the objective function (24) is inherited from the time transformation (10) applied to an original formulation of the objective function in the real continuous-time domain.

The first term of the objective function (24) ensures strip temperature tracking by penalizing the error between the strip temperature and its target temperature at the pyrometer positions z_{rth} and z_{rts}, cf. Fig. 1. Moreover, the second term ensures high material throughput and low energy consumption by penalizing low strip velocities and high fuel consumption. The relative importance of the respective control objectives can be adjusted for individual strip sections by the weighting matrices $\mathbf{W}_{1,k}$ and $\mathbf{W}_{2,k}$.

To highlight a major benefit of using the time transformation (10), consider for a moment that the optimization problem is formulated without using this transformation. If the control objectives were then prioritized differently for individual strips, the weighting matrices $\mathbf{W}_{1,k}$ and $\mathbf{W}_{2,k}$ would depend on the strip velocity (an optimization variable). This fact would complicate the calculation of the gradient of the objective function. However, if the time transformation is used, each grid point of the normalized time horizon corresponds to a strip section and the weighting matrices are therefore independent of the strip velocity. The strip length considered in the prediction horizon $(\tau_{k_0}, \tau_{k_1}]$ is equal to the value L used in the time transformation (10), i.e., the normalized control problem is not formulated in terms of a fixed (real) time horizon but in terms of a fixed strip length.

With the objective function (24), the discrete-time system dynamics (14), and the constraints (21), (22) and (23), the furnace control problem reads as

$$\underset{\substack{\mathbf{u}_k \in \mathbb{R}^{N_u} \\ \forall k \in \bar{K}}}{\text{minimize}} \quad \sum_{k\in\bar{K}} \|\mathbf{r}_{1,k}(\mathbf{h}_s(\mathbf{T}_k), \mathbf{u}_k)\|_{\mathbf{W}_{1,k}} + \|\mathbf{r}_{2,k}(\mathbf{u}_k)\|_{\mathbf{W}_{2,k}} \tag{25a}$$

$$\text{subject to} \quad \mathbf{0} = \mathbf{F}_{k-1}(\mathbf{T}_k, \mathbf{T}_{k-1}, \mathbf{u}_k, \mathbf{u}_{k-1}), \ \forall k \in \bar{K}, \tag{25b}$$

$$\mathbf{T}_{k_0} = \mathbf{T}_0, \mathbf{u}_{k_0} = \mathbf{u}_0, \tag{25c}$$

$$\mathbf{T}^-_{s,k} \leq \mathbf{h}_s(\mathbf{T}_k) \leq \mathbf{T}^+_{s,k}, \forall k \in \bar{K}, \tag{25d}$$

$$\mathbf{u}_k^- \leq \mathbf{u}_k \leq \mathbf{u}_k^+, \forall k \in \bar{K} \tag{25e}$$

$$\dot{\mathbf{u}}_k^- \leq \mathbf{h}_{u,k}(\mathbf{u}_k, \mathbf{u}_{k-1}) \leq \dot{\mathbf{u}}_k^+, \forall k \in \bar{K}, \tag{25f}$$

with the abbreviation $\bar{K} = \{k_0 + 1, \ldots, k_1\}$ and $N_k = (k_1 - k_0)N_u$ optimization variables.

A challenge in solving the control problem (25) are the inequality constraints (25d). In literature, numerous approaches for solving this kind of problem can be found, see, e.g., [11, 28]. However, if the inequality constraints are implemented as they are, i.e., as hard constraints, the constrained optimization problem might be infeasible and its solution is generally computationally expensive. This is mainly because in every iteration of a numerical solution algorithm it has to be checked whether the inequaltiy constraints are active or not. To circumvent this issue, in particular in view of a real-time implementation, the constrained control problem (25) is transformed into an alternative representation with equality constraints only by using a nonlinear input transformation and additional penalty terms, i.e., soft constraints, in the objective function.

5.2. Approximation of inequality constraints

Since it is important to keep the number of optimization variables at a moderate level, the number of optimization variables is reduced, i.e., the solution of the optimization problem is confined to a lower dimensional subspace of $\mathbb{R}^{N_k}$. Let $\mathbf{U}_l \in \mathbb{R}^{N_u}$ be the input vector at a grid point T_l of the coarse time grid $(\mathrm{T}_{l_0}, \mathrm{T}_{l_0+1}, \ldots, \mathrm{T}_{l_1-1}, \mathrm{T}_{l_1})$ with $\mathrm{T}_{l_0} = \tau_{k_0}$ and $\mathrm{T}_{l_1} = \tau_{k_1}$.

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012

The number of optimization variables is thus reduced to $N_l = (l_1 - l_0)N_u < N_k$. The input vector $\mathbf{u}_k$ at the nodal point τ_k of the fine time grid can be computed by linear interpolation of the grid values $\mathbf{U}_l$ in the form

$$\mathbf{u}_k = \sum_{l=l_0}^{l_1} \Phi_l(\tau_k)\mathbf{U}_l \tag{26a}$$

with triangular functions

$$\Phi_l(\tau_k) = \begin{cases} \frac{\tau_k - \mathrm{T}_{l-1}}{\mathrm{T}_l - \mathrm{T}_{l-1}} & \text{if } \tau_k \in [\mathrm{T}_{l-1}, \mathrm{T}_l] \\ \frac{\mathrm{T}_{l+1} - \tau_k}{\mathrm{T}_{l+1} - \mathrm{T}_l} & \text{if } \tau_k \in (\mathrm{T}_l, \mathrm{T}_{l+1}] \\ 0 & \text{else.} \end{cases} \tag{26b}$$

The reduced control performance due to this alternative input parameterization is usually rewarded by less computational effort for solving the optimization problem.

The coarse time grid for the control inputs suggests that the second term of the objective function (24) is evaluated at the time instances T_l with the values $\mathbf{U}_l$, where $l = l_0 + 1, \ldots, l_1$. Moreover, the constraints (25e) and (25f) are replaced by

$$\mathbf{U}_l^- \le \mathbf{U}_l \le \mathbf{U}_l^+, \qquad l \in \bar{L} \tag{27}$$

$$\dot{\mathbf{U}}_l^- \le \mathbf{h}_{U,l}(\mathbf{U}_l, \mathbf{U}_{l-1}) \le \dot{\mathbf{U}}_l^+, \qquad l \in \bar{L} \tag{28}$$

with the associated function $\mathbf{h}_{U,l}$, the associated limits $\mathbf{U}_l^\pm$ and $\dot{\mathbf{U}}_l^\pm$, and the abbreviation $\bar{L} = \{l_0 + 1, \ldots, l_1\}$.

The constraints (27) are taken into account by the input transformation

$$\mathbf{U}_l = \boldsymbol{\phi}_l(\mathbf{V}_l) = \mathrm{diag}\{\mathbf{U}_l^+ - \mathbf{U}_l^-\}\boldsymbol{\varphi}(\mathbf{V}_l) + \mathbf{U}_l^- \tag{29a}$$

with the nonlinear mapping

$$\boldsymbol{\varphi}(\mathbf{V}_l) = \frac{1}{2}\left(\mathbf{1} + \tanh\left(2\mathbf{V}_l\right)\right) \tag{29b}$$

and the new unconstrained input $\mathbf{V}_l$. Here, $\mathrm{diag}(\mathbf{U})$ denotes a diagonal matrix containing the elements of $\mathbf{U}$. The mathematical operation $\tanh(\cdot)$ is meant to be individually applied to the respective vector element. Moreover, $\mathbf{1} \in \mathbb{R}^{N_u}$ represents a vector with unity elements only. A scalar representation of the mapping (29b) is shown in Fig. 6. Note that the constraints (27) are automatically satisfied if $\mathbf{U}_l$ is computed according to (29a). Clearly, $\boldsymbol{\varphi}(\mathbf{V}_l) \in [\mathbf{0}, \mathbf{1}]$ from (29b) represents a normalized input vector. The mapping (29b) requires the additional term

$$\varepsilon \sum\nolimits_{l=l_0+1}^{l_1} \mathbf{V}_l^\mathrm{T}\mathbf{V}_l \tag{30}$$

with an arbitrary value $\varepsilon > 0$ to be added to the objective function (24) in order to avoid singular arcs, cf. [29]. For a concise notation, the function, cf. (26) and (29),

$$\boldsymbol{\psi}_k(\mathbf{V}) = \sum_{l=l_0}^{l_1} \Phi_l(\tau_k)\boldsymbol{\phi}_l(\mathbf{V}_l)$$

is introduced, i.e., $\mathbf{u}_k = \boldsymbol{\psi}_k(\mathbf{V})$.

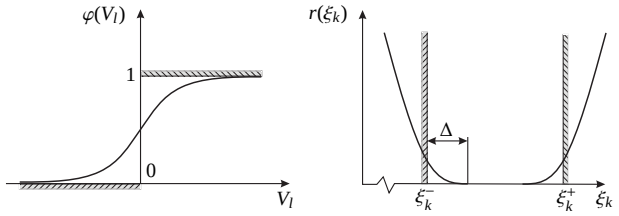

Figure 6: Input transformation $\varphi(V_l)$ and penalty term $r(\xi_k) = \left\|\max(0, \xi_k^- + \Delta - \xi_k, \xi_k - \xi_k^+ + \Delta)\right\|_\mathrm{w}$ (shown for scalar quantities only).

The remaining inequalities (25d) and (28) can be replaced, for instance, by classical penalty- or barrier-methods [30, 11]. However, for the considered control problem, the additional term

$$\sum_{k=k_0+1}^{k_1} \left\|\mathbf{r}_{3,k}(\mathbf{h}_\mathrm{s}(\mathbf{T}_k))\right\|_{\mathbf{W}_{3,k}} + \sum_{l=l_0+1}^{l_1} \left\|\mathbf{r}_{4,l}(\mathbf{V})\right\|_{\mathbf{W}_{4,k}}, \tag{31a}$$

with

$$\mathbf{r}_{3,k}(\mathbf{h}_\mathrm{s}(\mathbf{T}_k)) = \max\left(\mathbf{0}, \mathbf{T}_{s,k}^- + \boldsymbol{\Delta}_s - \mathbf{h}_\mathrm{s}(\mathbf{T}_k), \mathbf{h}_\mathrm{s}(\mathbf{T}_k) - \mathbf{T}_{s,k}^+ + \boldsymbol{\Delta}_s\right) \tag{31b}$$

$$\mathbf{r}_{4,l}(\mathbf{V}) = \max\Big(\mathbf{0}, \dot{\mathbf{U}}_l^- + \boldsymbol{\Delta}_r - \mathbf{h}_{U,l}(\boldsymbol{\phi}_{l-1}(\mathbf{V}_{l-1}), \boldsymbol{\phi}_l(\mathbf{V}_l)), \mathbf{h}_{U,l}(\boldsymbol{\phi}_{l-1}(\mathbf{V}_{l-1}), \boldsymbol{\phi}_l(\mathbf{V}_l)) - \dot{\mathbf{U}}_l^+ + \boldsymbol{\Delta}_r\Big), \tag{31c}$$

the assembled input vector $\mathbf{V} = [\mathbf{V}_l]_{l=l_0+1,\ldots,l_1}$, and $\boldsymbol{\phi}_{l_0}(\mathbf{V}_{l_0}) = \mathbf{U}_{l_0} = \mathbf{u}_0$, is added to the objective function (24). An example of the chosen approach is shown in Fig. 6 for a general quantity ξ_k constrained to $[\xi_k^-, \xi_k^+]$. How exactly the original constraints are adhered to, can be controlled to a certain extent by the parameters $\boldsymbol{\Delta}_i > 0$ with $i \in \{s, r\}$ and the weighting matrices $\mathbf{W}_i$ with $i \in \{3, 4\}$. Note that using this approach may cause minor violations of the original restrictions (25d) and (28). However, for the considered control problem, this approach is perfectly acceptable.

The objective function finally includes the terms (24), (30), and (31a). Let $\mathbf{T} = [\mathbf{T}_k]_{k=k_{0+1},\ldots,k_1}$ be the assembled temperature vector. Introducing

$$\mathbf{c}_1(\mathbf{T}, \mathbf{V}) = [\mathbf{c}_{1,k}(\mathbf{T}_k, \mathbf{V})]_{k=k_0+1,\ldots,k_1} = \begin{bmatrix} \mathbf{r}_{1,k}(\mathbf{h}_\mathrm{s}(\mathbf{T}_k), \boldsymbol{\psi}_k(\mathbf{V})) \\ \mathbf{r}_{3,k}(\mathbf{h}_\mathrm{s}(\mathbf{T}_k)) \end{bmatrix}_{k=k_0+1,\ldots,k_1}$$

$$\mathbf{c}_2(\mathbf{V}) = \begin{bmatrix} \mathbf{r}_{2,l}(\mathbf{V}_l) \\ \mathbf{r}_{4,l}(\mathbf{V}) \\ \mathbf{V}_l \end{bmatrix}_{l=l_0+1,\ldots,l_1}$$

the final objective function reads as

$$C(\mathbf{V}, \mathbf{T}) = C_1(\mathbf{V}, \mathbf{T}) + C_2(\mathbf{V}) = \|\mathbf{c}_1(\mathbf{V}, \mathbf{T})\|_{\mathbf{W}_{\mathbf{c}_1}} + \|\mathbf{c}_2(\mathbf{V})\|_{\mathbf{W}_{\mathbf{c}_2}}$$

with diagonal weighting matrices $\mathbf{W}_{\mathbf{c}_1}$ and $\mathbf{W}_{\mathbf{c}_2}$ that includes the (diagonal) entries of the matrices $\mathbf{W}_i$, $i = \{1, 3\}$, and $\mathbf{W}_i$, $i = \{2, 4\}$, as well as ε from (30), respectively.

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

Based on (26), (29), and (31a), the optimization problem (25) is transformed into the optimization problem

$$\underset{\mathbf{V}\in\mathbb{R}^{N_l}}{\text{minimize}} \quad C(\mathbf{V},\mathbf{T}) \tag{32a}$$

$$\text{subject to} \quad \mathbf{0} = \mathbf{F}(\mathbf{T},\mathbf{V}) \tag{32b}$$

with $\mathbf{F}(\mathbf{T},\mathbf{V}) = [\mathbf{F}_{k-1}(\mathbf{T}_k,\mathbf{T}_{k-1},\boldsymbol{\psi}_k(\mathbf{V}),\boldsymbol{\psi}_{k-1}(\mathbf{V}))]_{k=k_0+1,\ldots,k_1}$ where $\mathbf{T}_{k_0} = \mathbf{T}_0$. Compared to the original formulation (25), the optimization problem (32) does not include inequality constraints and a solution always exists.

5.3. Numerical solution of the optimization problem

The exact solution of an optimization problem is typically approximated by a sufficiently accurate numerical solution generated by a suitable optimization algorithm. For the optimization problem (32), many standard numerical solution methods can be found, e.g., the steepest descent method [31], the conjugate gradient method [32], the quasi-Newton method [33], the Gauss-Newton method [11], or the Newton method [11]. Since the solution of the optimization problem (32) is used in a model predictive control concept, an efficient solution method in terms of computational effort and convergence rate is essential. The quasi-Newton method features a superlinear convergence rate and is often used for problems like (32), cf. [34, 35]. This method requires only the computation of the objective function and its gradient with respect to $\mathbf{V}$. Moreover, the Hessian matrix can be estimated with little computational effort by applying, for instance, the BFGS formula [36]. However, a crucial point for the convergence rate of the quasi-Newton method is the initial choice of the estimated Hessian matrix.

The Gauss-Newton method does not have this problem because it uses a more accurate approximate calculation of the Hessian matrix. For this reason, the Gauss-Newton method is used in this paper. It provides also a superlinear convergence rate and requires an objective function that is made up of a sum of squared function values. Compared to the Quasi-Newton method, the main difference is that the Hessian matrix is approximated by exploiting the special structure of the objective function. Given that (32b) holds, the gradient vector $\mathbf{g}$ and the approximated Hessian matrix $\mathbf{H}$ read as

$$\mathbf{g} = \left(\frac{\mathrm{d}C}{\mathrm{d}\mathbf{V}}\right)^{\mathrm{T}} = 2\left(\frac{\mathrm{d}\mathbf{c}_1}{\mathrm{d}\mathbf{V}}\right)^{\mathrm{T}}\mathbf{W}_{\mathbf{c}_1}\mathbf{c}_1 + \left(\frac{\mathrm{d}C_2}{\mathrm{d}\mathbf{V}}\right)^{\mathrm{T}}$$

$$\mathbf{H} = 2\left(\frac{\mathrm{d}\mathbf{c}_1}{\mathrm{d}\mathbf{V}}\right)^{\mathrm{T}}\mathbf{W}_{\mathbf{c}_1}\left(\frac{\mathrm{d}\mathbf{c}_1}{\mathrm{d}\mathbf{V}}\right) + \left(\frac{\mathrm{d}^2C_2}{\mathrm{d}\mathbf{V}^2}\right)^{\mathrm{T}}.$$

Since $C_2(\mathbf{V})$ is independent of $\mathbf{T}$, the analytical computation of $\mathrm{d}C_2(\mathbf{V})/\mathrm{d}\mathbf{V}$ and $\mathrm{d}^2C_2(\mathbf{V})/\mathrm{d}\mathbf{V}^2$ is fairly straightforward and computationally undemanding. For the computation of $\mathrm{d}\mathbf{c}_1(\mathbf{T},\mathbf{V})/\mathrm{d}\mathbf{V}$, the vector-valued Lagrangian

$$\mathbf{L} = \mathbf{c}_1(\mathbf{T},\mathbf{V}) + \boldsymbol{\Lambda}\mathbf{F}(\mathbf{T},\mathbf{V})$$

with a lower-triangular block-structured matrix $\boldsymbol{\Lambda}$ of Lagrangian multipliers is introduced. Vector-valued Lagrangians are commonly used in multi-objective optimization [37]. In this paper, this formulation is borrowed for the analytical computation of $\mathrm{d}\mathbf{c}_1(\mathbf{T},\mathbf{V})/\mathrm{d}\mathbf{V}$, which follows in the form

$$\frac{\mathrm{d}\mathbf{c}_1(\mathbf{T},\mathbf{V})}{\mathrm{d}\mathbf{V}} = \frac{\partial\mathbf{L}}{\partial\mathbf{V}} = \frac{\partial\mathbf{c}_1(\mathbf{T},\mathbf{V})}{\partial\mathbf{V}} + \boldsymbol{\Lambda}\frac{\partial\mathbf{F}(\mathbf{T},\mathbf{V})}{\partial\mathbf{V}} \tag{33a}$$

with

$$\mathbf{0} = \mathbf{F}(\mathbf{T},\mathbf{V}) \tag{33b}$$

$$\mathbf{0} = \frac{\partial\mathbf{L}}{\partial\mathbf{T}} = \frac{\partial\mathbf{c}_1(\mathbf{T},\mathbf{V})}{\partial\mathbf{T}} + \boldsymbol{\Lambda}\frac{\partial\mathbf{F}(\mathbf{T},\mathbf{V})}{\partial\mathbf{T}}. \tag{33c}$$

In principle, $\mathrm{d}\mathbf{c}_1(\mathbf{T},\mathbf{V})/\mathrm{d}\mathbf{V}$ can be determined from (33a) by means of $\mathbf{T}$ and $\boldsymbol{\Lambda}$ that can be calculated in a straightforward way from (33b) and (33c). However, this approach is computationally expensive and memory-intensive. A numerically efficient computation of $\mathrm{d}\mathbf{c}_1(\mathbf{T},\mathbf{V})/\mathrm{d}\mathbf{V}$ according to (33) proceeds as follows: First, the state trajectory $\mathbf{T}_k$ is solved for the ascending time indices $k_0+1,\ldots,k_1$ (forward direction) by using $\mathbf{F}_{k-1}(\mathbf{T}_k,\mathbf{T}_{k-1},\boldsymbol{\psi}_k(\mathbf{V}),\boldsymbol{\psi}_{k-1}(\mathbf{V})) = \mathbf{0}$ with $\mathbf{T}_{k_0} = \mathbf{T}_0$ and $\boldsymbol{\psi}_{k_0}(\mathbf{V}) = \mathbf{u}_0$, i.e., (33b) is solved. It is easy to see that the matrix $\partial\mathbf{c}_1/\partial\mathbf{T}$ has a block-diagonal structure. Similarly, the matrix $\partial\mathbf{F}/\partial\mathbf{T}$ has a sparse block structure with non-zero matrices only on the main diagonal and the subdiagonal, see Fig. 7. Note that the entries of $\partial\mathbf{c}_1/\partial\mathbf{T}$ and $\partial\mathbf{F}/\partial\mathbf{T}$ can be analytically evaluated. Because of the sparse structure of $\partial\mathbf{c}_1/\partial\mathbf{T}$ and $\partial\mathbf{F}/\partial\mathbf{T}$, $\boldsymbol{\Lambda}$ has a lower-triangular block structure. Clearly, it is not recommendable to compute $\boldsymbol{\Lambda}$ by direct inversion of $\partial\mathbf{F}/\partial\mathbf{T}$, i.e., in the form $\boldsymbol{\Lambda} = -(\partial\mathbf{c}_1/\partial\mathbf{T})/(\partial\mathbf{F}/\partial\mathbf{T})^{-1}$. Instead, the column blocks of $\boldsymbol{\Lambda}$ should be successively computed with descending time indices $k_1,\ldots,k_0+1$ (backward direction) using the calculated state trajectory $\mathbf{T}_k$. In the same loop, the column blocks of $\mathrm{d}\mathbf{c}_1(\mathbf{T},\mathbf{V})/\mathrm{d}\mathbf{V}$ can be simultaneously computed. This is because $\partial\mathbf{F}/\partial\mathbf{V}$ has the same sparse block structure as $\partial\mathbf{F}/\partial\mathbf{T}$. Note that this strategy features modest memory requirements because the large matrix $\boldsymbol{\Lambda}$ does not need to be stored at any time. Instead, only the currently considered column block is of interest.

The iterative numerical optimization algorithm can therefore be summarized as follows:

Step 0: Provide an initial guess $\mathbf{V}$

Step 1: Compute the search direction $\mathbf{d} = -\mathbf{H}^{-1}\mathbf{g}$

Step 2: Perform a line search, i.e., solve

$$\underset{\alpha\geq 0}{\text{minimize}} \quad C(\mathbf{V}+\alpha\mathbf{d},\mathbf{T})$$
$$\text{subject to} \quad \mathbf{0} = \mathbf{F}(\mathbf{T},\mathbf{V}+\alpha\mathbf{d})$$

and compute the update $\mathbf{V} \leftarrow \mathbf{V} + \alpha\mathbf{d}$.

Step 3: Check if any termination criterion (maximum number of iterations, convergence) is fulfilled. If yes, stop here.

Step 4: Start again at Step 1.

In step 2, a scalar optimization problem for the step size α is solved by means of a local quadratic approximation

$$C(\mathbf{V}+\alpha\mathbf{d},\mathbf{T}) \approx q(\alpha) = a_0 + a_1\alpha + a_2\alpha^2$$

and an underlying interval adaption [38]. If the objective func-

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

$$\mathbf{0} = \begin{bmatrix} \frac{\partial \mathbf{c}_{1,k_0+1}}{\partial \mathbf{T}_{k_0+1}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \frac{\partial \mathbf{c}_{1,k_0+2}}{\partial \mathbf{T}_{k_0+2}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \frac{\partial \mathbf{c}_{1,k_1}}{\partial \mathbf{T}_{k_1}} \end{bmatrix} + \begin{bmatrix} \mathbf{\Lambda}_{k_0,1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{\Lambda}_{k_0+1,1} & \mathbf{\Lambda}_{k_0+1,2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{\Lambda}_{k_1-1,1} & \mathbf{\Lambda}_{k_1-1,2} & \cdots & \mathbf{\Lambda}_{k_1-1,k_1-k_0} \end{bmatrix} \begin{bmatrix} \frac{\partial \mathbf{F}_{k_0}}{\partial \mathbf{T}_{k_0+1}} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \frac{\partial \mathbf{F}_{k_0+1}}{\partial \mathbf{T}_{k_0+1}} & \frac{\partial \mathbf{F}_{k_0+1}}{\partial \mathbf{T}_{k_0+2}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \frac{\partial \mathbf{F}_{k_0+2}}{\partial \mathbf{T}_{k_0+2}} & \frac{\partial \mathbf{F}_{k_0+2}}{\partial \mathbf{T}_{k_0+3}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \frac{\partial \mathbf{F}_{k_1-1}}{\partial \mathbf{T}_{k_1-1}} & \frac{\partial \mathbf{F}_{k_1-1}}{\partial \mathbf{T}_{k_1}} \end{bmatrix}$$

Figure 7: Detailed structure of (33c) for determining $\mathrm{d}\mathbf{c}_1(\mathbf{T}, \mathbf{V})/\mathrm{d}\mathbf{V}$.

tion is evaluated at three sample points $\alpha_1 < \alpha_2 < \alpha_3$ with $\alpha_2 = (\alpha_1 + \alpha_3)/2$, the polynomial coefficients a_0, a_1, and a_2 can be determined by solving the set of equations

$$C(\mathbf{V} + \alpha_i \mathbf{d}, \mathbf{T}) = q(\alpha_i), \qquad i = 1, 2, 3.$$

If $a_2 > 0$, the step size that minimizes $q(\alpha)$ takes the form

$$-\frac{1}{2}\frac{a_1}{a_2}.$$

If this step size is additionally in the interval $[\alpha_1, \alpha_3]$, then $\alpha = -a_1/(2a_2)$ is the approximate solution of the line search problem. Otherwise, α is set to one of the interval bounds α_1 or α_3. In this case, the interval bounds are shifted accordingly in order to track the minimum of the line search problem over the subsequent Gauss-Newton iterations [38].

In step 3, the numerical solution algorithm is terminated if a fixed number N_c of iterations is achieved, i.e., a suboptimal solution for the input trajectories is used. This termination criterion ensures that a defined sampling rate of the controller can be realized. Note that the proposed numerical solution algorithm always satisfies (32b), even if the algorithm is used in a suboptimal way.

5.4. Convergence behavior of the numerical solution algorithm

A numerical solution algorithm is mainly characterized by its rate of convergence, its numerical effort per iteration, and its numerical stability. In the following, the convergence properties of the Gauss-Newton formulation proposed in the previous section is compared for a single prediction horizon with the convergence properties of the quasi-Newton method. Here, the quasi-Newton method is implemented with the BFGS formula and the identity matrix as initial condition for the estimated Hessian matrix. For comparison purposes, the line-search problem is identically realized in both algorithms.

Because of the analytical calculation of the Hessian matrix, the computation time, i.e., the computational effort, per iteration of the Gauss-Newton method is 1.06 times higher than that of the quasi-Newton method. Starting from an arbitrary initial guess $\mathbf{V}$ and an initial value $C = 209$ of the objective function, Fig. 8 shows the values C for every iteration of both solution algorithms. The Gauss-Newton method reaches the almost optimal value $C = 2.66$ after just 5 iterations whereas the quasi-Newton method requires many more iterations, i.e., the rates

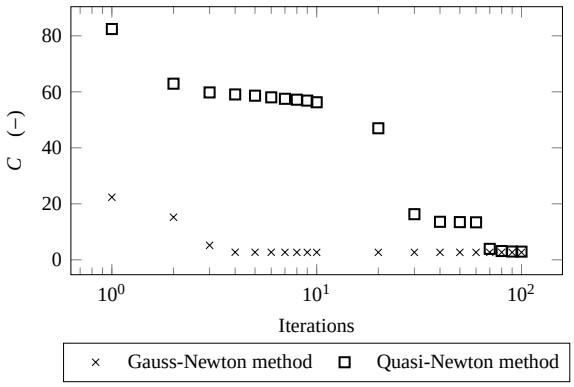

Figure 8: Decrease of the objective function C as a function of the iterations of the solution algorithm.

of convergence differ significantly. Even if both algorithms are used in a suboptimal way, the performance of the Gauss-Newton method is significantly better. This result justifies the choice of the Gauss-Newton method for the efficient solution of the optimization problem (32), although it has a slightly higher numerical effort per iteration than the quasi-Newton method.

6. Industrial application example

For the verification of the developed model predictive controller, a simulation study is carried out. In this simulation study, the real furnace (plant), cf. Fig. 3, is replaced by the validated high-fidelity furnace model presented in [13, 14, 15]. This furnace model has been extensively verified by means of measurement data from the considered combined direct- and indirect-fired strip annealing furnace of voestalpine Stahl GmbH in Linz, Austria, and is therefore suitable for the emulation of the real furnace operation. The system state that is required for the initial condition for the controller, i.e., the feedback, is provided by an ad-hoc state estimator [39]. Additionally, the estimator provides an estimation of the strip emissivity. The state estimator is based on the furnace model (11) and uses the system inputs and the noisy pyrometer measurement $\mathbf{h}_s(\mathbf{T}_s) + \mathbf{v}$ with the measurement noise $\mathbf{v} \in \mathbb{R}^3$ to improve the estimation. The entries of $\mathbf{v}$ are chosen to be uniformly distributed in the range $[-5, 5]$ K.

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

The scenario for the simulation study is taken from a real production process at the considered annealing furnace. The process data defining the scenario include geometric dimensions, material properties, target temperatures, and corresponding temperature constraints of 21 different strips with a total length of 55 km. The width b_s and thickness d_s of the strips are shown in Fig. 9. They are in the range $b_s \in [1.3, 1.5]$ m and $d_s \in [0.6, 0.8]$ mm, respectively. The absolute values and the maximum slopes of the control inputs, i.e., the input constraints, are chosen in accordance with their real counterparts.

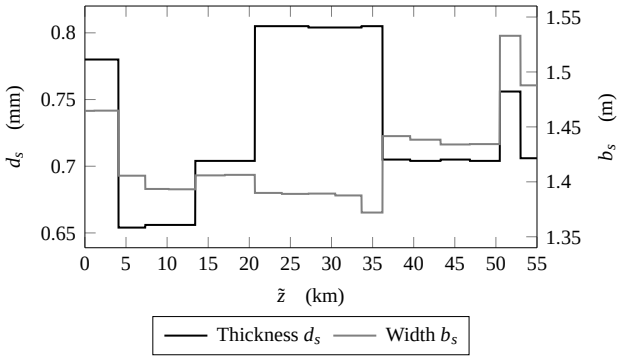

Figure 9: Geometric dimensions of the considered strips.

For the simulation, the controller is implemented in MATLAB with a sampling time $\delta_c = 2$ s. On a standard desktop PC (4 GHz, 16 GB RAM), the time required for the solution of the optimal control problem is in the range of 1.7 s. A new optimal control input is thus provided at the time instants $k\delta_c$ with $k \in \mathbb{N}$. The (constant) strip length considered in each prediction horizon $[\mathrm{T}_{l_0}, \mathrm{T}_{l_1}]$ with $\mathrm{T}_{l_1} - \mathrm{T}_{l_0}$ is $L = 1000$ m, which is more than 4 times the length of the strip inside the furnace. The optimal control problem takes into account $N_l = (l_1 - l_0)N_u = 592$ optimization variables and the numerical solution algorithm terminates after $N_c = 1$ iteration. In order to compensate the computational delay [40], the numerical solution algorithm that is started at the time instant $k\delta_c$ is initialized with an estimation of the system state at the time instant $(k+1)\delta_c$, cf. [41]. This future system state is predicted based on the estimated system state provided by the state estimator at the time instant $k\delta_c$ and by using the furnace model (12) and the open-loop inputs that are applied to the plant in the time interval $[k\delta_c, (k+1)\delta_c]$. Thus, at the time instant $(k+1)\delta_c$, an optimal control input is provided which systematically takes into account the computational delay. Moreover, the suboptimal solution of the previous optimization problem is incorporated into the initial guess of the current optimization problem. Because the previous and the current prediction horizons overlap, it is intuitive to initialize the corresponding optimization variables with the associated solution of the previous optimization problem. The remaining optimization variables at the end of the current prediction horizon are initialized with their final value at the end of the previous horizon.

In steady-state operation, the controller is tuned to satisfy both the primary and the secondary control objectives. However, in non-steady-state operation, e.g., when a welded joint that connects two strips with different target temperatures is processed, the secondary control objectives are of minor importance. To prioritize the primary control objective in non-steady-state operation, the matrix $\mathbf{W}_{\mathbf{c}_2}$ is adapted by reducing the entries as long as a welded joint appears in the prediction horizon. This adaptation may vary for different strip transitions.

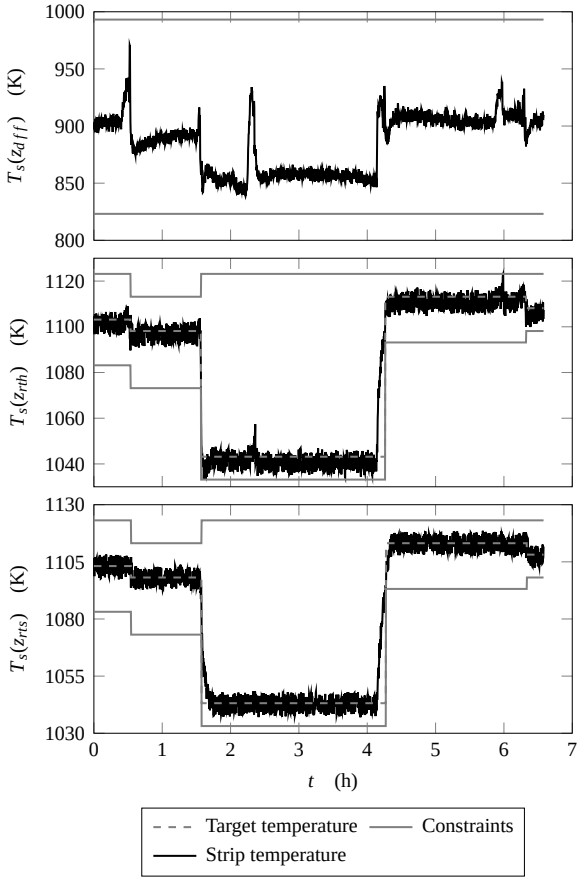

Figure 10: Simulated strip temperature, corresponding target temperatures, and the temperature constraints at the pyrometer positions z_{dff}, z_{rth}, and z_{rts}.

Because the main objective is accurate strip temperature control, the proposed controller is mainly verified by means of the strip temperatures at the two pyrometer positions at the end of the RTH and the RTS section, cf. Fig. 1. Figure 10 shows that the noisy strip temperature at these two positions accurately follows the respective target temperature and is always within the tolerance range. This also holds true in transient operational situations, e.g., when a welded joint that connects two different strips moves through the furnace. Figure 10 also shows that the strip temperature at the pyrometer position z_{dff} is within its temperature limits.

The proposed model predictive controller takes into account known future disturbances, e.g., strip transitions. This anticipative behavior is shown in Fig. 11 which is a detail of Fig. 10. In order to meet all temperature limits, the strip temperature rises before the strip transition with the corresponding change in the

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

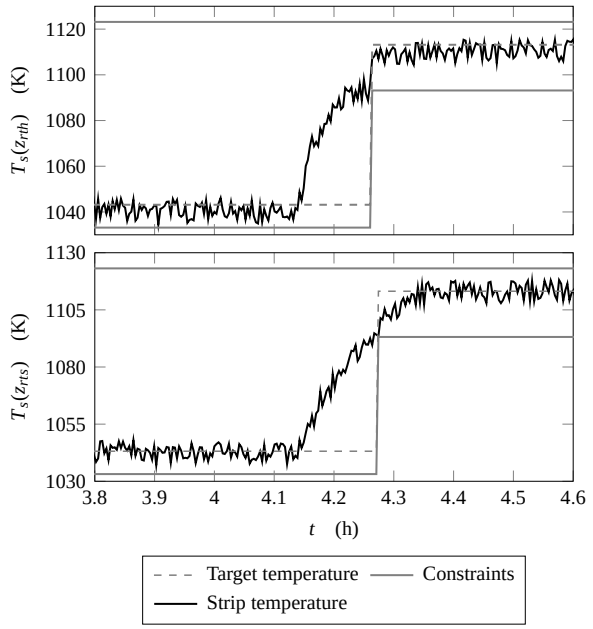

Figure 11: Detail of simulated strip temperature, corresponding target temperatures, and the temperature constraints at the pyrometer positions z_{rth}, and z_{rts}, cf. Fig. 10.

target temperature appears at the respective pyrometer measurement. A control concept that is not anticipative would react only after a change of the target temperature becomes visible in the control error at the respective position. In this case, the risk of violating the temperature limits would be much higher.

Examples of the associated control inputs, i.e., the mass flows of fuel in a typical heating zone of both the direct-fired and the indirect-fired furnace section and the strip velocity, are shown in Fig. 12. These quantities are within their constraints. The mass flows of fuel to the other heating zones look similar to those of Fig. 12.

Figures 10 and 12 show that the controller ensures both accurate strip temperature control and a maximization of the throughput. This is because the heating zones are widely operated at maximum capacity and a further increase of the strip velocity would thus be at the expense of the strip temperature accuracy.

In the considered annealing furnace, strip temperature control is currently realized by a standard PI control concept. In order to assess whether the control accuracy for the considered production scenario increases with the proposed model predictive controller, the performance indicator

$$\Delta T_s = \sqrt{\frac{1}{t_\Sigma}\int_0^{t_\Sigma} \mathbf{e}^\mathrm{T}\mathbf{e}\,\mathrm{d}t} \tag{34a}$$

with

$$\mathbf{e} = \begin{bmatrix} T_s^d(z_{rth},t) - T_s(z_{rth},t) \\ T_s^d(z_{rts},t) - T_s(z_{rts},t) \end{bmatrix} \tag{34b}$$

is used, where t_Σ is the total time that is needed for the heat

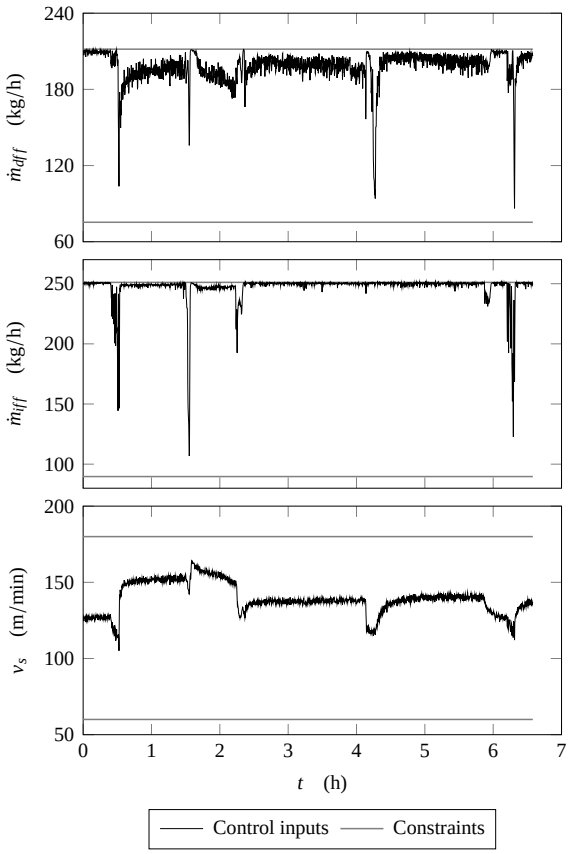

Figure 12: Simulated system inputs, i.e., mass flow of fuel to a heating zone of the direct- and the indirect-fired furnace and strip velocity, as well as corresponding constraints.

treatment of the considered 21 strip products. With the PI control concept, the heat treatment of the 21 strips required $t_\Sigma^{meas} = 8.6\,\mathrm{h}$. Using the corresponding measurement data, (34) results in $\Delta T_s^{meas} = 20.13\,\mathrm{K}$. As it can be inferred from Fig. 10, the proposed temperature controller requires only $t_\Sigma^{sim} = 6.8\,\mathrm{h}$ for processing the considered strip products. This means that an optimized strip velocity would lead to a reduction of the production time by more than 20%. Of course, this reduction is equivalent to a maximization of the material throughput. If the simulated strip temperatures $T_s(z_{rth},t)$ and $T_s(z_{rts},t)$ from Fig. 10 are used, (34) results in $\Delta T_s^{sim} = 9.15\,\mathrm{K}$, which indicates that the proposed controller would also significantly increase the control accuracy.

In addition to the control accuracy and the production time, the energy consumption of the annealing furnace is another important control performance parameter. During the time period t_Σ^{meas}, the heat treatment of the considered strip products by means of the PI control concept requires a total amount of fuel of $m_{CH_4}^{meas} = 8.95\,\mathrm{t}$. In contrast, $m_{CH_4}^{sim} = 8.92\,\mathrm{t}$ is needed in the simulation study. This means that the fuel consumption per ton strip product is slightly reduced. These obtained results indicate the great potential of the proposed controller to increase the control accuracy, to maximize the material throughput and

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

to minimize the energy consumption at the real plant.

7. Conclusions

A nonlinear model predictive controller for the strip temperature of a combined direct- and indirect-fired strip annealing furnace was proposed. Here, both the mass flows of fuel to the heating zones and the strip velocity are used as control inputs. The controller additionally maximizes the throughput and minimizes the energy consumption.

The basis of the proposed controller is a tractable first-principles model of the furnace, which is computationally inexpensive and captures the most import nonlinear effects. The model includes submodels of the flue gas, the radiant tubes, the wall, the strip, and the rolls. These submodels are interconnected by the heat transfer mechanisms radiation, convection, and conduction.

For the model predictive controller, a tailored dynamic optimization problem is solved by a numerical solution algorithm. The time domain of the furnace model is first normalized using a time transformation that depends on the strip velocity. This ensures that the strip length considered in each prediction horizon is constant. To meet the different control objectives, a suitable objective function that depends on the system inputs and states is defined for a finite time horizon. Constraints of inputs and system states are taken into account by using sigmoid functions and additional penalty terms in the objective function. The discrete-time optimization problem that finally consists of the objective function and the furnace model as an equality constraint is iteratively solved in a suboptimal way by means of the Gauss-Newton method. The required gradient and the approximate Hessian matrix of the objective function can be analytically computed using an adjoint-based method. The suboptimal solution is used for the initial guess for the next optimization problem.

The capability of the proposed controller was demonstrated by a validated high-fidelity model of an industrial annealing furnace with process data from the real production process. The controller achieves accurate strip temperature control in both steady-state and non-steady-state furnace operation and ensures compliance with all constraints. Furthermore, the controller is currently implemented by voestalpine Stahl GmbH at the considered strip processing line.

Achnowledgements

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

References

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

[2] E. Kayacan, M. Khanesar, Fuzzy Neural Networks for Real Time Control Applications, Butterworth-Heinemann, Waltham, USA, 2016.

[3] P. Mullinger, B. Jenkins, Industrial and Process Furnaces, 2nd Edition, Butterworth-Heinemann, Oxford, 2008.

[4] J. Li, H. Feng, S. Li, Wavelet prediction fuzzy neuronal network of the annealing furnace temperature control, in: Proceedings of International Conference on Control Engineering (ICEICE), Wuhan, China, 2011, pp. 940–943.

[5] S. Li, Q. Chen, G. Huang, Dynamic temperature modeling of continuous annealing furnace using GGAP-RBF neural network, Neurocomputing 69 (2006) 523–536.

[6] I. Ueda, M. Hosoda, K. Taya, Strip temperature control for a heating section in CAL, in: Proceedings of the International Conference on Industrial Electronics, Control and Instrumentation (IECON), Kobe, Japan, 1991, pp. 1946–1949.

[7] N. Yoshitani, A. Hasegawa, Model-based control of strip temperature for the heating furnace in continuous annealing, IEEE Transactions on Control Systems Technology 6 (2) (1998) 146–156.

[8] D. Clarke, C. Mohtadi, P. Tuffs, Generalized predictive control – Part I. The basic algorithm, Automatica 23 (2) (1987) 137 – 148.

[9] D. Clarke, C. Mohtadi, P. Tuffs, Generalized predictive control – Part II. Extensions and interpretations, Automatica 23 (2) (1987) 149 – 160.

[10] L. Bitschnau, S. Jakubek, M. Kozek, Constrained model predictive control of a continuous annealing furnace, in: Proceedings of the ASME 2010 Dynamic Systems and Control Conference (DSCC), Cambridge, USA, 2010, pp. 285–292.

[11] J. Nocedal, S. Wright, Numerical Optimization, 2nd Edition, Springer, New York, USA, 2006.

[12] H. Wu, R. Speets, F. Heeremans, O. Driss, R. van Buren, Non-linear model predictive control of throughput and strip temperature for continuous annealing line, Iron and Steelmaking 42 (8) (2015) 570–578.

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

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

[15] M. Niederer, S. Strommer, A. Steinboeck, A. Kugi, M. Fein, M. Boeck-Schnepps, G. Helekal, 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 Sheet co-located with the 5th International Conference on Hot Sheet Metal Forming of High-Performance Steel (GALVATECH), Toronto, Canada, 2015, pp. 137–144.

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

[17] S. Turns, An Introduction to Combustion: Concepts and Applications, 2nd Edition, McGraw-Hill, New York, 2006.

[18] M. Holmes, Introduction to Numerical Methods in Differential Equations, Springer, New York, 2007.

[19] H. Hottel, E. Cohen, Radiant heat exchange in a gas filled enclosure: Allowance for non-uniformity of gas temperature, American Institute of Chemical Engineering Journal 4 (1958) 3–14.

[20] H. Hottel, A. Sarofim, Radiative Transfer, McGraw-Hill, New York, 1967.

[21] R. Siegel, J. Howell, Thermal Radiation Heat Transfer, 4th Edition, Taylor and Francis, New York, 2002.

[22] J. Rhine, R. Tucker, Modelling of Gas-Fired Furnaces and Boilers and Other Industrial Heating Processes, McGraw-Hill, London, UK, 1991.

[23] F. Incropera, D. DeWitt, T. Bergman, A. Lavine, Fundamentals of Heat and Mass Transfer, 6th Edition, John Wiley & Sons, Hoboken, New Jersey, 2007.

[24] H. Baehr, K. Stephan, Heat and Mass Transfer, 2nd Edition, Springer, Berlin, Heidelberg, 2006.

[25] U. Ascher, L. Petzold, Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations, SIAM, Philadelphia, USA, 1998.

[26] C. Froehlich, S. Strommer, A. Steinboeck, M. Niederer, A. Kugi, Modeling of the Media-Supply of Gas Burners of an Industrial Furnace, in: Proceedings of the IEEE Industry Applications Society Annual Meeting, Addison, USA, 2015, pp. 1–9.

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.

[27] S. Strommer, A. Steinboeck, C. Begle, M. Niederer, A. Kugi, Modeling and control of gas supply for burners in gas-fired industrial furnaces, in: Proceedings of the IEEE Conference on Control Applications (CCA), Antibes, France, 2014, pp. 210–215.
[28] D. Bertsekas, Nonlinear Programming, 2nd Edition, Athena Scientific, Massachusetts, USA, 1999.
[29] K. Graichen, N. Petit, Constructive methods for initialization and handling mixed state-input constraints in optimal control, Journal of Guidance, Control, and Dynamics 31 (5) (2008) 1334–1343.
[30] I. Griva, S. Nash, A. Sofer, Linear and Nonlinear Optimization, 2nd Edition, SIAM, 2008.
[31] H. Curry, The method of steepest descent for non-linear minimization problems, Quarterly of Applied Mathematics 2 (1944) 258–261.
[32] L. Lasdon, S. Mitter, A. Waren, The conjugate gradient method for optimal control problems, Transactions on Automatic Control 12 (2) (1967) 133–138.
[33] C. G. Broyden, The convergence of a class of couble-rank minimization algorithms: General considerations, Journal of Applied Mathematics 12 (2) (1970) 76 – 90.
[34] A. Steinboeck, D. Wild, A. Kugi, Nonlinear model predictive control of a continuous slab reheating furnace, Control Engineering Practice 21 (4) (2013) 495–508.
[35] F. Schausberger, A. Steinboeck, A. Kugi, Optimization-based estimator for the contour and movement of heavy plates in hot rolling, Journal of Process Control 29 (2015) 23 – 32.
[36] R. Fletcher, Practical methods of optimization, John Wiley & Sons, Chichester, New York, 1980.
[37] Y. Sawaragi, H. Nakayama, T. Tanino, Theory of Multiobjective Optimization, Academic Press, Orlando, Florida, 1985.
[38] K. Graichen, B. Kaepernick, A real-time gradient method for nonlinear model predictive control, Frontiers of Model Predictive Control (2012) 9 – 28.
[39] S. Strommer, M. Niederer, A. Steinboeck, L. Jadachowski, A. Kugi, Nonlinear observer for temperatures and emissivities in a strip annealing furnace, in: accepted for presentation at the 2016 IEEE Industry Applications Society (IAS) Annual Meeting, Portland, USA, 2016.
[40] J. Sánchez, J. Rodellar, Adaptive Predictive Control: From the Concepts to Plant Optimization, Prentice Hall Internatioal Series in Systems and Control Engineering, Prentice Hall, London, New York, 1996.
[41] R. Findeisen, Allgöwer, Computational delay in nonlinear model predictive control, in: Proceedings of the International Symposium on Advanced Control of Chemical Processes (ADCHEM), Hong Kong,, 2003, pp. 427–432.

Post-print version of the article: M. Niederer, S. Strommer, A. Steinboeck, and A. Kugi, "Nonlinear model predictive control of the strip temperature in an annealing furnace", *Journal of process control*, vol. 48, pp. 1–13, 2016. DOI: 10.1016/j.jprocont.2016.09.012
The content of this post-print version is identical to the published paper but without the publisher's final layout or copy editing.