Modeling, Identification and Control, Vol. 47, No. 2, 2026, pp. 71–84, ISSN 1890–1328

Optimal Aluminium Extrusion Planning Under Minimum-Taper Constraints

T.A. Gabrielsen [a] L.S. Imsland [a] P. Singstad [b] A.H. Koren [b] F. Gjertsen [b] T. Aukrust [c] O.J. Emmerhoff [d]

[a] *Department of Engineering Cybernetics, NTNU, Trondheim, Norway. E-mail:* {tagabrie,lars.imsland}@ntnu.no

[b] *Cybernetica AS, Trondheim, Norway. E-mail:* {peter.singstad,anne.koren,fredrik.gjertsen}@cybernetica.no

[c] *SINTEF Industry, Oslo, Norway. E-mail:* trond.aukrust@sintef.no

[d] *Hydro Aluminium, Norway. E-mail:* ole.johannes.emmerhoff@hydro.com

Abstract

The minimum allowable initial billet taper affects the thermal margin available during aluminium extrusion and can therefore constrain the attainable production rate. This paper uses full-horizon dynamic optimization to quantify this trade-off while jointly optimizing the initial billet temperature distribution, ram-speed trajectory, and die-cooling trajectory. The formulation provides a means of predicting targets for the billet heating phase that facilitate minimum-time extrusion while accounting for peak-temperature and other process constraints. The problem is posed as a nonlinear program with extrusion time as its primary objective and with various process constraints, such as peak outlet temperature. We solve the problem with varying degrees of control freedom (ram-speed, cooling, initial taper) to quantify how the available process degrees of freedom affect the attainable extrusion time. In the considered case, under the assumption that extrusion-pressure limits remain inactive within the prescribed operating bounds, the minimum-taper constraint is active at the optimum. Over the investigated range, the corresponding minimum extrusion time increases approximately affinely by about 3.7 s per additional 10 °C/m of axial taper. An industrial closed-loop simulator exhibits the same qualitative relationship, with a corresponding increase of about 2.5 s per 10 °C/m.

Keywords: Aluminium extrusion, process control, optimal control, nonlinear programming, direct collocation

1 Introduction

In aluminium extrusion, both productivity and product quality are clear objectives in the design of the extrusion process. The exit *(peak)* temperature in aluminium extrusion is a key quality-related process variable, which usually needs to be in a narrow range of temperatures to achieve acceptable quality. On the other hand, productivity is strongly linked to the ram-speed trajectory. However, through aluminium deformation and associated heat generation, the ram speed is strongly linked to the peak temperature as well as the total extrusion time, often making the two objectives mutually constraining.

In addition to the ram speed, the initial billet temperature distribution *(taper)* and coolant flow through channels in the die can be used to compensate the heat generation. The peak temperature can be passively reduced by selecting an appropriate initial billet taper, preventing overheating and allowing higher ram speeds. Similarly, adjustments to the coolant flow can be used to actively manage the peak temperature dur-

doi:10.4173/mic.2026.2.3
© 2026 Norwegian Society of Automatic Control

ing extrusion, thereby permitting higher ram speeds. The tight coupling between initial taper, ram speed, and die cooling motivates their joint consideration in the process design.

Previous studies have already established die cooling as an effective means of thermal management. Hölker et al. (2013, 2015) numerically and experimentally investigate the effects of conformal cooling channels, and Reggiani and Donati (2019) develop numerical models predicting the effect of liquid-nitrogen cooling.

Other studies have established the use of ram speeds and initial taper as important control variables in attaining isothermal extrusion. Chanda et al. (2001) use 3D Finite Element Method (FEM) simulations to compare constant and stepwise-decreasing ram speed trajectories, showing that varying the ram speed can reduce peak temperature variations and improve productivity. Li et al. (2003) take this further by using FEM in the design of a predetermined nonlinear initial taper to suppress excess heat and stabilize the peak temperature. FEM simulations are also used by Zhou et al. (2004) to predetermine a continuously varying ram-speed trajectory that both achieves isothermal extrusion and shortens the extrusion time. A comprehensive sampling-based study by Farjad Bastani et al. (2011) covers isothermal extrusion by joint consideration of ram speed, linear initial tapers, and *container* cooling, where isothermal maps are interpolated between sample points. Lela et al. (2014) use Functional Data Analysis (FDA) and industrial data to construct a predictive model which they use to compute an appropriate ram-speed curve, and Barella et al. (2021) develop a simple model to estimate the required initial taper for isothermal extrusion under specified operating conditions. Together, these studies establish the effectiveness of ram-speed variation and initial taper selection for controlling the peak temperature. However, they predominantly prescribe trajectories via simulation studies or predictive rules rather than formal, simultaneous optimization.

The following works put a stronger emphasis on productivity: Guo et al. (2023) use FEM simulations to generate data that is used to develop a model suited for PID updates to the ram speed at prescribed points along the ram stroke during extrusion, with a maximum speed bound. Danilin et al. (2024) study the initial taper via FEM simulations with varying ram speeds for high-strength alloys, and Iddberg et al. (2024) develop a fast steady-state model that predicts the maximum allowable extrusion speed before *hot tearing*. Although productivity is a more prominent part of their motivation, these studies determine feasible or improved ram speeds through feedback adjustments, parametric simulation studies, and steady-state ram-speed limits, respectively, rather than via full-horizon numerical trajectory optimization.

Formal numerical optimization has been used in several previous works, in different ways and with varying objectives. An early occurrence is the study by Tibbetts and Wen (1998), in which they use optimization methods to design a ram-speed trajectory for maintaining high strain rates, then solve for an initial taper that stabilizes the peak temperature throughout the extrusion phase. A limitation of this sequential approach is that the taper and ram-speed trajectory are not optimized jointly. Cuéllar Matamoros (1999) instead considers simultaneous optimization of the initial taper and ram speed, where they develop computationally tractable extrusion models and formulate open-loop optimization problems for isothermal extrusion. Their thermal tracking objective is deliberately normalized by the final extrusion time, such that reducing the extrusion time does not by itself impact the tracking objective, and they restrict the initial taper to be linear. The work also separately formulates an NMPC scheme. Kathirgamanathan and Neitzert (2008) formulate a nonlinear least-squares problem connected to an ABAQUS simulator in order to optimize process parameters such as die geometry, billet temperature and ram speed, with the objective of achieving desired micro-structure characteristics. Then, recently, Zhou et al. (2023) couple an ABAQUS simulator to a Python-implemented MPC scheme to adjust the ram speed for thermal tracking. The above studies employ numerical optimization methods, but do not directly minimize extrusion time in a full-horizon dynamic optimization problem with the discussed process decision variables.

The most direct predecessors to this work are Kulås (2022) and Gabrielsen and Imsland (2023). The former considers a full-horizon open-loop dynamic optimization problem, and uses a ram-speed-based productivity objective that rewards high ram speeds at individual points along the ram stroke. However, this is not equivalent to minimizing extrusion time. Furthermore, the initial taper is optimized indirectly by including the heating phase as part of the problem and directly optimizing the heater temperature setpoints. This has the advantage of naturally incorporating the limitations of the heater process, but results in a harder, multi-stage problem. In the latter work, Gabrielsen and Imsland (2023) introduce a model transformation that instead uses extrusion length as the progress variable, enabling an efficient implementation of a direct collocation formulation of the full-horizon extrusion problem. They formulate, implement and solve a dynamic optimization problem that optimizes over the ram speed and explicitly uses minimum extrusion time as their pri-

mary objective. However, they do not include the initial taper or coolant as decision variables, nor do they study the effect of varying the available degrees of control freedom.

We have not identified a full-horizon dynamic optimization problem formulation that explicitly minimizes the extrusion time subject to peak-temperature and process constraints, while jointly optimizing the initial billet taper, the ram-speed trajectory and the die-cooling trajectory. In this paper, we formulate and implement such a problem.

To clarify the role and importance of each of the three main control degrees of freedom, we solve the problem for different configurations, where different subsets of the possible control degrees of freedom are fixed, and compare the overall improvement to a baseline simulation.

An important part of designing the extrusion process is to set a minimum required initial billet taper *(minimum taper)* to prevent air entrapment during upsetting. The choice of the minimum taper and its impact on extrusion time are not well covered in the literature. We therefore show how the proposed optimization framework can support this design process and provide results quantifying its effect on the optimal extrusion time in the particular model and problem considered here.

The contributions can be summarized as:

- Full-horizon joint optimization of initial taper, ram speed, and cooling for minimum extrusion time.
- A configuration study quantifying the value of the available process degrees of freedom.
- A parametric study showing how the imposed minimum axial taper affects the attainable optimum, with comparison against an industrial simulator.

2 Aluminium Extrusion

The basic principle of extrusion is that a metal cylinder, called a 'billet', is placed into a container and pressed through an opening in a component called the 'die', as illustrated in Fig. 1.

The *billet*, shown as the light gray section of Fig. 1, is the source metal for the extrudate. The billet can be of varying alloy composition and geometry and is usually preheated to specified temperatures. All these factors have to be considered for each individual extrusion cycle to obtain various desired process and product properties, and are of particular interest when optimizing the extrusion cycle.

The *container* is a component that encases the billet during extrusion. The container affects the extrusion process mainly through heat conduction and billet-container friction, and it may be pre-heated to reduce heat loss. The container also has a maximum pressure capacity, imposing an upper bound on the possible extrusion pressure, limiting the maximum extrusion speed.

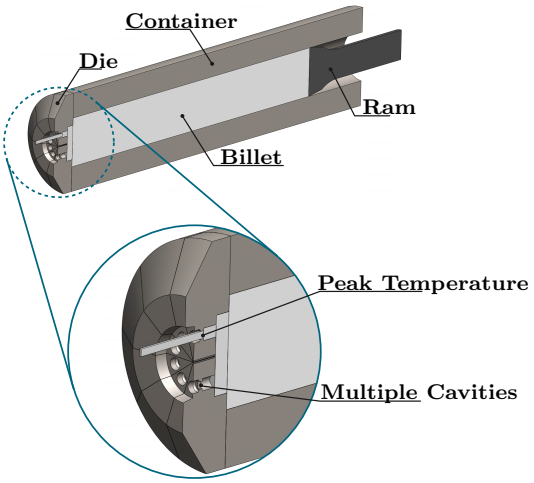

Figure 1: A sketch of a direct extrusion press, illustrating its main components and where the peak temperature of the aluminium is located.

The *die* is the component through which the billet is forced, giving shape to the outgoing aluminium profile. When force is applied to the billet, it deforms as it passes through the die and assumes its intended profile. This is a complex part of extrusion due to the complex die geometry. The work required to deform the aluminium, and the work done by friction, generate significant thermal energy, increasing the temperature of the flowing aluminium rapidly as it eventually reaches its maximum—'peak'—temperature at the die exit. To more closely control the temperature of the aluminium in the die, cooling channels may be embedded to allow the use of coolant that can be controlled by adjusting a valve opening. Hölker et al. (2013) and Hölker et al. (2015) investigate the effects of cooling channels in the die. The generation of thermal energy in the die is affected by the extrusion speed, extrusion pressure, die geometry and die cooling.

The *ram* is the hydraulically actuated piston that applies pressure to the billet, forcing it through the die. The ram force is typically controlled by a low-level controller (such as a PID controller), such that the ram speed closely tracks a reference speed, which is traditionally constant or a predetermined speed curve.

Figure 1 summarizes the main components of the extrusion press, illustrates the common axial symmetry in extrusion, and coarsely depicts some of the important and complex die design parameters. Further-

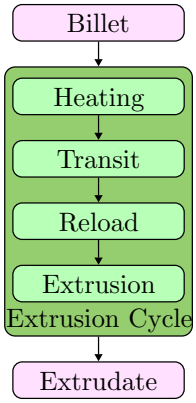

Figure 2: Illustration of the extrusion cycle of a billet, showing the four phases of the extrusion process.

more, it highlights various important die-design parameters, including feeder and port dimensions and bearing length, represented by successive reductions in cross-section towards the extrusion exit. The figure also shows how several 'cavities' are used to produce multiple profiles simultaneously.

2.1 The Extrusion Cycle

The preceding description only concerns the extrusion phase, though the complete extrusion process consists of several phases: the heating phase, the transit phase, a reload phase, and the extrusion phase, which make up the *'extrusion cycle'*. The extrusion cycle is depicted in Fig. 2.

In the *heating phase*, a billet is placed into an induction/furnace heater to be heated to a prespecified temperature distribution—*'temperature taper'*. Note that the radial temperature variations are also part of the temperature taper, though radial diffusion evens out radial variations relatively quickly. During the *transit phase*, the billet is transported from the heater to the extrusion press. In this paper, we assume that the transit phase has a fixed duration and no control inputs. Similarly, the *reload phase* is when the butt of the previous billet is sheared off and the new billet is inserted into the extrusion press. This phase is also uncontrolled, and is sometimes omitted and absorbed into the transit phase. An important distinction, however, is that the transit phase can happen in parallel to the extrusion phase of another cycle, whereas the reload phase may not. In the *extrusion phase*, the aluminium is pressed through the die and assumes the intended profile. The temperature taper of the billet at the start of the extrusion phase is called the *'initial taper'*. We define the time frame of an extrusion cycle in Fig. 3, where t_h is referred to as the 'heating time', t_t is the 'transit time', t_{RL} is 'reload time', and t_f is the 'final time'/'extrusion time'. The total cycle time is then

$$t_c = t_h + t_t + t_{RL} + t_f. \tag{1}$$

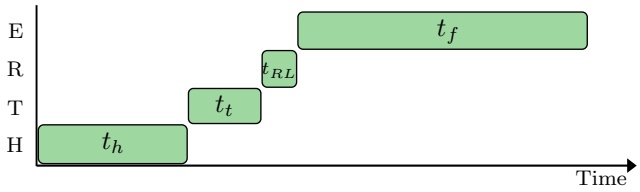

Figure 3: A visual representation of the timeline of an extrusion cycle. Phases are represented on the vertical axis, where H denotes the heating phase, T is the transit phase, R is the reload phase and E is the extrusion phase. The duration of each phase is given inside the green regions: t_h, t_t, t_{RL}, t_f.

In practice the extrusion cycles will overlap, as illustrated in Fig. 4. Therefore, the time between the start of two consecutive cycles—the extrusion period: $\mathcal{T}_{ext}$—is not necessarily the same as the total cycle time, and it is usually more accurate to assume that the extrusion period is closer to the extrusion time plus reload time; $\mathcal{T}_{ext} \approx t_f + t_{RL}$. Since the reload time is constant, the production rate is determined by the extrusion time t_f.

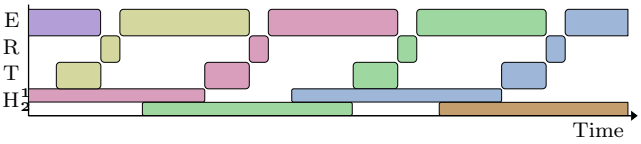

Figure 4: A visualization of consecutive cycles. Each color represents the extrusion cycle of an individual billet from heating to extrusion. Notice that multiple heating phases may take place simultaneously if there are multiple heaters available. The figure illustrates two heaters, denoted 'H1' and 'H2'.

2.2 Process Constraints

There are several constraints that must be imposed on the extrusion phase, which we present here. A natural, yet important, constraint is that the ram cannot stop nor retract during extrusion due to safety concerns. The ram speed must therefore be strictly positive, bounded away from zero. Furthermore, in practice, there is usually a constraint on the maximum pressure inside the extrusion press. Computing

the pressure accurately can be rather involved (Sheppard, 1999), and it is usually monotonically increasing in the ram speed and decreasing in the temperature. Therefore, in this work, we assume that the pressure limits are not exceeded given a sufficiently conservative lower bound on the temperature and upper bound on ram speed. Consequently, the results are conditional on the pressure limits of the relevant extrusion press and on the ram speed and temperature constraints being sufficiently conservative to prevent pressure-limit violations.

Ultimately, we have

$$v_{ram}(t) \in [\underline{v}_{ram}, \overline{v}_{ram}] \quad \forall t \in \langle 0, t_f], \tag{2}$$

where $v_{ram}(t)$ is the ram speed, and $\underline{v}_{ram}$ and $\overline{v}_{ram}$ are the minimum and maximum permissible ram speeds, respectively. We also use a universal lower bound on temperature: $\underline{T}$, and impose an upper bound $\overline{T}$ to avoid melting. The valve opening governing the flow of coolant is constrained to

$$z_{valve}(t) \in [0, \overline{z}_{valve}] \subseteq [0, 1] \quad \forall t \in [0, t_f], \tag{3}$$

where $z_{valve}(t)$ is the valve opening, with 1 being fully open, and 0 being fully closed, and $\overline{z}_{valve} \in [0, 1]$ is the maximum permissible valve opening, usually limited by either the valid regime of the cooling model or a physical limitation of the cooling system.

Often, the taper of the billet is required to strictly descend from the front (die-side) to the back (ram-side) at the start of the press. This is to prevent entrapment of air during the *'upsetting'*—engaging the ram—which may cause critical defects, or even dangerous conditions for operators. We denote the minimum required temperature gradient of the initial taper along the axial length of the billet, simply *'minimum taper'*, by $\underline{\tau}_x$ [°C/m]. In this paper, we first allow any non-increasing tapers, then study the effects of tightening the taper requirement in Section 5.4.

The peak temperature should be within a narrow range of temperatures and is therefore softly constrained to:

$$T_{peak}(t) \in [\underline{T}_{peak} - \epsilon(t), \overline{T}_{peak} + \epsilon(t)] \quad \forall t \in [0, t_f], \tag{4}$$

where $T_{peak}(t)$ is the peak temperature, and ϵ is an auxiliary/slack variable that should be close to zero at all times.

3 Extrusion Model

The modeling work in this paper is based on work by Cuéllar Matamoros (1999); Aukrust and LaZghab (1999). The extrusion phase model utilizes the axial symmetry of the extrusion press, and is therefore defined on the reduced cylindrical space. We now regard the two-dimensional spatial coordinates $(r, x) \in \mathbb{R}_+ \times \mathbb{R}$, where r represents the radial coordinate, and x represents the axial coordinate. The extrusion press is divided into several sections, each representing a different essential part of the press; the billet, the container, the feeder (first contraction of aluminium towards the die exit), the bridge (second contraction, and first cavity-specific section), the port (a continuation of the bridge section), the peak (where the aluminium reaches its peak temperature, immediately before exiting the die), the plate (the part of the die that is in contact with the peak section, and is affected by cooling), and the die (the remaining part of the die). The model sections are depicted in Fig. 5.

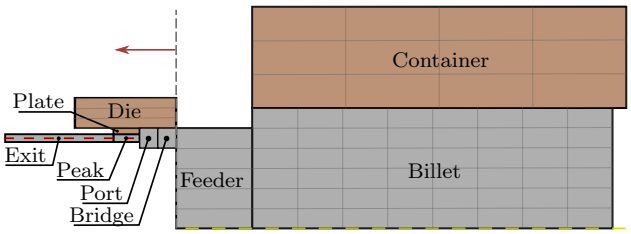

Figure 5: Cross-section of the extrusion model geometry, showing the various model sections. Steel sections are shown in brown, and aluminium sections are shown in gray. The red arrow indicates the direction of flow, or 'extrusion direction'. The vertical dashed gray line separates the sections that are common for the whole system, and the sections that represent a specific cavity. The yellow and red dashed lines represent the main and secondary center axes respectively. Grid lines inside sections depict individual control volumes. Notice that the exit-section is not modeled.

The extrusion temperature dynamics are described by partial differential equations (PDEs) in continuous time as

$$\begin{aligned}\frac{\partial T(r,x;t)}{\partial t} &= v(r,x;t)\frac{\partial T(r,x;t)}{\partial \bar{x}} \\ &+ \alpha\left(\frac{1}{\bar{r}}\frac{\partial}{\partial \bar{r}}\left(\bar{r}\frac{\partial T(r,x;t)}{\partial \bar{r}}\right) + \frac{\partial^2 T(r,x;t)}{\partial \bar{x}^2}\right) \\ &+ \tilde{\Phi}(r,x;T,\bar{v}) + \tilde{\Psi}(r,x;T,\bar{v}),\end{aligned} \tag{5}$$

where $T(r,x;t)$ is the temperature at spatial coordinate $(r,x) \in \mathbb{R}^2$ (r is the radial coordinate, and x is

the axial coordinate), at time t. The first part represents the advection of the aluminium as it is displaced by the ram, with $v(\cdot)$ being the aluminium flow speed in the extrusion direction, the second part is the heat equation in reduced cylindrical coordinates, and $\tilde{\Phi}$ and $\tilde{\Psi}$ represent the temperature contributions by viscous dissipation, respectively within each section and between sections. The latter is computed via reduction of cross-sectional area between the sections, with $\bar{v}$ being the average velocity within the given section. The PDEs are then discretized in space using finite differences (Strikwerda, 2004) to produce a set of ordinary differential equations (ODEs) describing the heat flow between a set of annular-cylindrical control volumes, which are depicted as grid lines in Fig. 5. The temperature in each cell is denoted by $T_{\zeta,r,x}$, where ζ is the name of the section, and $r, x \in \mathbb{N}^2$ (as subscripts rather than arguments) are the radial and axial cell numbers. The viscous dissipation and reduction terms, $\tilde{\Phi}$ [K/s] and $\tilde{\Psi}$ [K/s], are computed based on equations derived by Aukrust and LaZghab (1999). The viscous dissipation energy is

$$\Phi(r, x, \bar{v}) = e^{\frac{Q_f}{n_f R_f T_{r,x}}} \left(\frac{\bar{v}\bar{r}^{n_f}}{\sqrt{3}} \frac{(n_f+3)}{R^{n_f+1}}\right)^{\frac{1+n_f}{n_f}} \left(\alpha_f \sqrt{3} A_f^{\frac{1}{n_f}}\right)^{-1} \quad (6)$$

where Φ [W] is the generated thermal energy from viscous dissipation in the aluminium, $\bar{v}$ is the average velocity of the aluminium in the extrusion direction in the given section, and A_f, n_f, Q_f, R_f, and α_f are flow related parameters as discussed by Aukrust and LaZghab (1999). The reduction work is found by

$$v_{A|B} = 2\bar{v}_A \frac{R_A^2}{R_A^2 + R_B^2}, \quad R_{A|B} = \sqrt{\frac{R_A^2 + R_B^2}{2}}, \quad (7a)$$

$$\sigma_{A|B} = \frac{2}{3\alpha_f} \left(\frac{(n_f+3) v_{A|B} e^{\frac{Q_f}{R_f T_{r,x}}}}{\sqrt{3} A_f R_{A|B}}\right)^{\frac{1}{n_f}}, \quad (7b)$$

$$F_{A|B} = \pi R_A^2 K_{A|B} \sigma_{A|B} \log_e\left(\frac{R_A^2}{R_B^2}\right), \quad (7c)$$

$$\Psi_{A|B}(\bar{v}_A) = F_{A|B} v_{A|B}. \quad (7d)$$

where R_A and R_B are the radii of the section prior to the reduction and after the reduction respectively, and $K_{A|B}$ is a material-specific constant. We assume high strain rates in the peak section, making the heat generation terms different for the peak section:

$$\Phi_B = -\frac{A_f}{n_f 2^{n_f} \alpha_f} \left(\frac{r}{\lambda}\right) e^{\left(\frac{r}{\lambda} - \frac{Q_f}{R_f T(r,x)}\right)}, \quad (8)$$

$$A_{A|B} = \frac{\pi R_{billet}^2}{E_{ratio}}, \quad \sigma_{A|B} = \frac{h_B}{2\lambda \alpha_f n_f}, \quad (9)$$

$$v_{A|B} = 2\bar{v}_A \frac{\pi R_A^2}{\pi R_A^2 + A_{A|B}}, \quad (10)$$

$$F_{A|B} = \pi R_A^2 \left(K_{A|B} \sigma_{A|B} \log_e\left(\frac{\pi R_A^2}{A_{A|B}}\right)\right), \quad (11)$$

where Eq. (7d) is used to find the reduction work, with A and B representing 'port' and 'peak' respectively, E_{ratio} being the extrusion ratio (total area reduction), h_{peak} being the height of the peak section, and λ being the 'characteristic length' defined by

$$\left(\frac{2\lambda}{h_{peak}}\right)^2 \left(e^{\frac{h_{peak}}{2\lambda}} \left(\frac{h_{peak}}{2\lambda} - 1\right) + 1\right) \equiv \frac{2^{n_f+1} \bar{v}_{peak} e^{\frac{Q_f}{R_f T_{peak,r,x}}}}{\sqrt{3} A_f h_{peak}}. \quad (12)$$

Additionally, the effect of coolant is modeled as

$$\Xi(z_{valve}) = -\frac{\eta k_{max} \Delta h}{c_p \Delta V} \frac{\rho_c}{\rho_d} z_{valve}^s, \quad (13)$$

where η is an efficiency factor, k_{max} is the maximum volumetric flow, Δh is the vaporization heat of the coolant, ρ_c is the coolant mass density, ΔV is the volume of the plate section, and c_p and ρ_d are the heat capacity and mass density of the die material respectively. The nonlinear relationship between the valve opening z_{valve} and its contribution to the change in temperature is captured by the dimensionless exponent $s > 0$. The heat contribution Ξ is then added to the plate section dynamics.

3.1 Diminishing Billet

During extrusion, the billet is continuously displaced in the extrusion direction. Because we use control volumes ('cells') that are fixed in space, the aluminium is slowly leaving the cells from the back towards the front side. This is illustrated in Fig. 6, which shows the partitioning of the billet into cells and how they change as the ram extension/*'extrusion length'* increases.

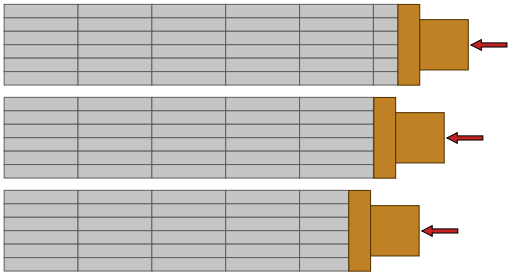

Figure 6: Illustration of the backmost cells disappearing during extrusion because of the gradual displacement of the billet.

At control-dependent times, aluminium completely leaves the backmost cells, causing instantaneous changes in the model equations as the boundary conditions change. These times are not known a priori because they depend on the ram speed trajectory, which is part of the decision space. The progressor transformation presented by Gabrielsen and Imsland (2023) is used to circumvent this problem.

4 Control Problem

In this section, we first present a dynamic optimization problem that minimizes the extrusion time subject to peak temperature requirements and other process constraints. We then briefly discuss how this optimization problem can fit into a larger control hierarchy.

The control objectives for the extrusion phase are (i) to maintain the desired peak temperature according to Eq. (4), and (ii) to minimize the extrusion time t_f, with the overall goal of maximizing production. The available decision variables are the ram speed trajectory, $v_{ram}(t)$, the coolant valve opening trajectory, $z_{valve}(t)$, and the initial taper.

4.1 Implementation

The extrusion phase control problem is solved using a nonlinear program (NLP), based on *direct collocation* (Diehl and Gros, 2024; Biegler, 2010; Rawlings et al., 2017). By employing the progressor transformation described by Gabrielsen and Imsland (2023), where the extrusion dynamics is 'progressed' by extrusion length instead of time, the collocation equations can be implemented as:

$$\begin{aligned} v_{ram,k} \cdot \boldsymbol{p}'(\tau_1, \mathcal{C}_k) - \Delta L \cdot \boldsymbol{f}_{k,1}(\boldsymbol{p}(\tau_1, \mathcal{C}_k), \boldsymbol{u}_k) &= \mathbf{0} \\ v_{ram,k} \cdot \boldsymbol{p}'(\tau_2, \mathcal{C}_k) - \Delta L \cdot \boldsymbol{f}_{k,2}(\boldsymbol{p}(\tau_2, \mathcal{C}_k), \boldsymbol{u}_k) &= \mathbf{0} \\ \vdots \\ v_{ram,k} \cdot \boldsymbol{p}'(\tau_d, \mathcal{C}_k) - \Delta L \cdot \boldsymbol{f}_{k,d}(\boldsymbol{p}(\tau_d, \mathcal{C}_k), \boldsymbol{u}_k) &= \mathbf{0} \\ \boldsymbol{T}_{k+1} - \boldsymbol{p}(1, \mathcal{C}_k) &= \mathbf{0}, \end{aligned} \tag{14}$$

for all discrete time/extrusion length points $k \in \mathbb{Z}_{0,N-1}$[1], where $\boldsymbol{p}$ is the collocation polynomial based on Lagrange interpolation polynomials (Biegler, 2010; Diehl and Gros, 2024; Rawlings et al., 2017), ΔL is the integration step in extrusion length, $\tau_{1,\dots,d}$ are the collocation points, and $v_{ram,k}$, $\boldsymbol{f}_{k,i}$ and $\mathcal{C}_k$ are the ram speed, the dynamical model and the collocation polynomial coefficients at extrusion length $L_{k,i} \triangleq L_k + \tau_i \Delta L$. Lastly, $\boldsymbol{T}_{k+1}$ is the temperature state at extrusion length L_{k+1}, for the relevant cell. The system trajectory is discretized on a uniform grid in extrusion length:

$$N \triangleq n_L \cdot n_x, \quad L_k \triangleq L_0 + k\Delta L \quad \forall k \in \mathbb{Z}_{1,N} \tag{15a}$$

$$\Delta L \triangleq (L_f - L_0)N^{-1}, \quad L_0 = 0m \tag{15b}$$

where N is the number of integration intervals, determined by n_L and n_x which are the number of integration intervals per billet cell in the axial dimension and the number of billet cells in the axial dimension respectively, and L_f is the final length. We define the following costs:

$$C_{\hat{z}} \triangleq \gamma_{\hat{z}} \sum_{k=0}^{N} \hat{z}_k, \quad C_v \triangleq \gamma_v \sum_{k=0}^{N} \frac{1}{v_k}, \quad C_\epsilon \triangleq \gamma_\epsilon \sum_{k=1}^{N} \epsilon_k \tag{16a}$$

$$C_{\Delta\hat{z}} \triangleq \gamma_{\Delta\hat{z}} \sum_{k=0}^{N-1} (\hat{z}_{k+1} - \hat{z}_k)^2, \tag{16b}$$

$$C_{\Delta v} \triangleq \gamma_{\Delta\hat{v}} \sum_{k=0}^{N-1} (\hat{v}_{k+1} - \hat{v}_k)^2 \tag{16c}$$

where v denotes v_{ram}, and all γ are weights. The valve opening variable is represented by $\hat{z}_k$, via the variable substitution $\hat{z} \triangleq z^s_{valve}$ to promote linearity, and ϵ_k represents the auxiliary variable in Eq. (4) in discrete time k. The objective function also includes a regularizing term, C_{reg}, detailed in Section 4.2.

[1] $\mathbb{Z}_{a,b}$ - Shorthand notation for the set $\{x \in \mathbb{Z} | a \leq x \leq b\}$

The complete NLP formulation is:

$$\min_{\mathcal{W}} \quad C_v + C_{\Delta v} + C_{\hat{z}} + C_{\Delta \hat{z}} + C_\epsilon + \gamma_{reg} C_{reg} \tag{17a}$$

s.t.

$$\boldsymbol{p}'(\tau_i, V_{\zeta,k}) = \frac{\Delta L}{v_{ram,k}} \boldsymbol{f}_{\zeta,k,i}(\boldsymbol{p}(\tau_i, V_{\zeta,k}), \boldsymbol{u}_k)\ 1), \tag{17b}$$

$$\boldsymbol{q}_{k+1} = \boldsymbol{p}(1, V_k) \quad 2), \tag{17c}$$

$$\boldsymbol{g}_{k,i}(\mathcal{W}) = \boldsymbol{0} \quad 3), \tag{17d}$$

$$\boldsymbol{T}_{init,r,x} \geq \boldsymbol{T}_{init,r+1,x} + \underline{\tau_r} R_b / n_r \quad 4), \tag{17e}$$

$$\boldsymbol{T}_{init,r,x} \leq \boldsymbol{T}_{init,r+1,x} + \overline{\tau_r} R_b / n_r \quad 4), \tag{17f}$$

$$\boldsymbol{T}_{init,r,x} \geq \boldsymbol{T}_{init,r,x+1} + \underline{\tau_x} L_b / n_x \quad 5), \tag{17g}$$

$$\boldsymbol{T}_{init,r,x} \leq \boldsymbol{T}_{init,r,x+1} + \overline{\tau_x} L_b / n_x \quad 5), \tag{17h}$$

$$T_{peak,k} \in [\underline{T}_{peak} - \epsilon_k, \overline{T}_{peak} + \epsilon_k] \quad 6), \tag{17i}$$

$$\epsilon_k \geq 0 \quad 6), \tag{17j}$$

$$\boldsymbol{T}_{\zeta,k} \in [\underline{T}_\zeta, \overline{T}_\zeta] \quad 7), \tag{17k}$$

$$v_{ram,k} \in [\underline{v}_{ram}, \overline{v}_{ram}] \quad 3), \tag{17l}$$

$$\hat{z}_k \in [0, \overline{z}^s_{valve}] \quad 3) \tag{17m}$$

with 1): $\forall i \in \mathbb{Z}_{0,d}, \forall k \in \mathbb{Z}_{0,N-1}, \forall \zeta \in \mathcal{S}^d$ 2): $\forall k \in \mathbb{Z}_{0,N-1}$ 3): $\forall k \in \mathbb{Z}_{0,N}$ 4): $\forall (r,x) \in \mathbb{Z}_{1,n_r-1} \times \mathbb{Z}_{1,n_x}$ 5): $\forall (r,x) \in \mathbb{Z}_{1,n_r} \times \mathbb{Z}_{1,n_x-1}$ 6): $\forall k \in \mathbb{Z}_{1,N}$ 7): $\forall k \in \mathbb{Z}_{0,N}, \forall \zeta \in \mathcal{S}$ where

$$\mathcal{W} = \mathcal{V} \cup \mathcal{Z} \cup \mathcal{D} \cup \mathcal{P} \tag{18a}$$

$$\mathcal{V} = \{v_{ram,0}, v_{ram,1}, \ldots, v_{ram,N}\} \tag{18b}$$

$$\mathcal{Z} = \{\hat{z}_0, \hat{z}_1, \ldots, \hat{z}_N\} \tag{18c}$$

$$\mathcal{D} = \bigcup_{\zeta \in \mathcal{S}^d} V_\zeta \cup \bigcup_{\zeta \in \mathcal{S}^a} \mathcal{T}_\zeta \tag{18d}$$

$$V_\zeta = \{V_{\zeta,0}, V_{\zeta,1}, \ldots, V_{\zeta,N-1}\} \tag{18e}$$

$$V_{\zeta,k} = \{\boldsymbol{T}_{\zeta,k}, \boldsymbol{v}_{\zeta,k,1}, \ldots, \boldsymbol{v}_{\zeta,k,d}\} \tag{18f}$$

$$\mathcal{T}_\zeta = \{\boldsymbol{T}_{\zeta,0}, \boldsymbol{T}_{\zeta,1}, \ldots, \boldsymbol{T}_{\zeta,N}\} \tag{18g}$$

$$\mathcal{S}^d = \{\mathit{billet}, \mathit{container}, \mathit{feeder}, \mathit{plate}, \mathit{die}\} \tag{18h}$$

$$\mathcal{S}^a = \{\mathit{bridge}, \mathit{port}, \mathit{peak}\} \tag{18i}$$

$$\mathcal{P}_\lambda = \{\lambda_0, \lambda_1, \ldots, \lambda_N\} \tag{18j}$$

$$\mathcal{P}_\epsilon = \{\epsilon_1, \epsilon_2, \ldots, \epsilon_N\} \tag{18k}$$

$$\mathcal{S} = \mathcal{S}^d \cup \mathcal{S}^a, \quad \mathcal{P} = \mathcal{P}_\lambda \cup \mathcal{P}_\epsilon \tag{18l}$$

where $\boldsymbol{v}_{\zeta,k,i}$ is the ith state collocation variable in integration interval k of ζ. The peak and port dynamics, Eqs. (6), (8), (7d), and (11), are represented by $\boldsymbol{g}$, and are assumed to be in steady state in Eq. (17d). The length and radius of the billet are given by L_b and R_b respectively, and constraints on the initial taper are imposed via Eqs. (17g), (17h), (17e), and (17f), with $T_{init,r,x}$ referring to billet temperature $T_{billet,r,x}$ at the initial time (initial taper). Additionally, all constraints are elementwise.

A last remark about the practical implementation of the NLP is that the billet temperature vector has no entries at $k = N$ since the remaining length of the billet at that point is $L_N = 0m$, and there are no cells left. In practice, the horizon is therefore reduced by one, $N_{reduced} = N - 1$, in the implementation of Eq. (17). This also reflects that, in practice, one stops the extrusion phase before the entirety of the billet is extruded, leaving a 'butt' that is sheared off before reloading with another billet. In reality, the extrusion time is therefore slightly lower than the computed $t_f = \sum_{k=0}^{N} \frac{\Delta L}{v_{ram,k}}$.

4.2 Regularization

To regularize the NLP, the following additional terms are added to the objective, via the C_{reg}-term:

$$\sum_{r=1}^{n_r} \sum_{x=2}^{n_x-1} \left(\boldsymbol{T}_{\zeta,0,(r,x+1)} - 2\boldsymbol{T}_{\zeta,0,(r,x)} + \boldsymbol{T}_{\zeta,0,(r,x-1)}\right)^2, \tag{19}$$

$$\sum_{r=2}^{n_r-1} \sum_{x=1}^{n_x} \left(\boldsymbol{T}_{\zeta,0,(r+1,x)} - 2\boldsymbol{T}_{\zeta,0,(r,x)} + \boldsymbol{T}_{\zeta 0,(r-1,x)}\right)^2, \tag{20}$$

where Eqs. (19) and (20) are applied with $\zeta = \mathit{billet}$ (the 'initial taper'). The subscript 0 indicates initial values. These terms are penalties on curvature in the axial and radial dimensions respectively, which favor linear temperature distributions. This is relevant because intricate, nonlinearly varying temperature distributions may be difficult to implement in practice, and it also helps the convergence of the NLP.

4.3 Extrusion Control Hierarchy

We now suggest how our optimization problem may fit into a larger control hierarchy. In Fig. 7 we depict an *Extrusion Cycle Optimization Layer (ECOL)* which receives information about the new billet and computes optimal open-loop reference trajectories for the heating and extrusion phases.

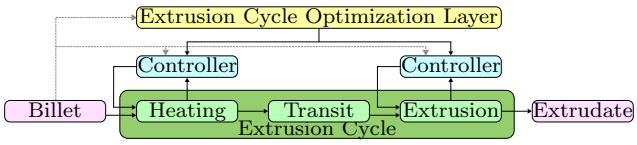

Figure 7: The proposed control hierarchy for the extrusion process, consisting of an extrusion cycle optimization layer (yellow) that plans an optimal extrusion process in open-loop and provides references to a lower layer (blue) of online feedback control loops for both the heating and extrusion phases.

The ECOL may first solve the extrusion phase problem, generating a target taper for the heating phase. Potentially, a heating phase problem may be posed with the optimal initial taper as a terminal state constraint, and the objective may be chosen independently, for example, focusing on minimizing energy consumption. The heating problem is simpler because it lacks the intricate dynamics related to flow of aluminium, since all aluminium is stationary in this phase. Figure 8 shows how the extrusion phase optimization problem, OCP E, could be incorporated into a larger control layer (ECOL). The ECOL then passes the optimal initial taper to another optimization problem that ensures that the initial taper is produced in an optimal and *feasible* way. Naturally, in this scheme, we desire

$$t_h \leq (t_{RL} + t_f) n_h, \tag{21}$$

where n_h is the number of heaters that can be used in parallel, such that the extrusion period is not delayed by the heating phase. A heating problem with a time budget can ensure this if it is feasible. If the time budget constraint is consistently infeasible, this suggests that increasing the number of heaters may increase production.

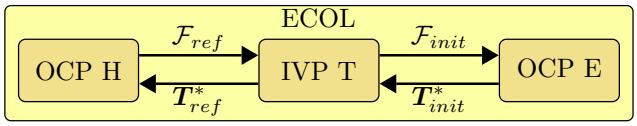

Figure 8: The figure shows a proposed structure of an extrusion cycle optimization layer. OCP E and OCP H are optimal control problems for the extrusion and heating phase respectively. IVP T is an initial value problem for the transit phase that connects the heating and extrusion phases. The right-facing arrows indicate the order in which the feasible regions $\mathcal{F}_{ref}$ (achievable tapers by heater) and $\mathcal{F}_{init}$ (achievable initial tapers) are estimated and passed on, while the left-facing arrows indicate the order in which the optimal control problems are solved and their solutions are passed to the next to be used as terminal constraints.

5 Results

In this section we present solutions to Eq. (17) for varying configurations. We start by presenting a baseline solution with a fixed ram speed curve and initial taper, whose values are taken from industrial practice. We then solve the problem for configurations where varying control inputs are free to be optimized, in order to investigate how optimization of various control inputs affects the extrusion times. Specifically, we define the configurations:

'ram': The initial taper is fixed, coolant is not used, and the ram speed trajectory is optimized.

'ram-valve': The initial taper is fixed, and both ram speed and valve opening trajectories are optimized.

'taper': The ram-speed and valve-opening trajectories are fixed at 3 mm/s and 0 %, respectively, and only the initial taper is optimized. The regularization parameter is set individually for the initial billet temperatures (*initial taper*) to $\gamma_{reg} = 100$.

'ram-taper': The valve opening trajectories are kept constant at 0%, while the ram speed trajectory and initial taper are optimized. The regularization parameter is set individually for the initial billet temperatures to $\gamma_{reg} = 100$.

'full': Ram speed, valve opening, and initial taper are all optimized. We focus particularly on the sensitivity of the extrusion time to changes in the minimum allowable axial taper, presented in Section 5.4.

5.1 Case Study and Optimization Setup

Unless otherwise specified, the following applies: Weights used for the objective are $\gamma_v = 50$, $\gamma_z = 1$, $\gamma_\epsilon = 10^7$, $\gamma_{\Delta v} = 10$, and $\gamma_{\Delta \hat{z}} = 10$. These weights are found by trials and error. The regularizing terms are weighted $\gamma_{reg} = 1$. Furthermore, the order of the collocation polynomials is $d = 2$, with Gauss-Legendre collocation points, and the number of integration steps per cell is $n_L = 1$. The desired range for the peak temperature is $[600\,^\circ\text{C}, 610\,^\circ\text{C}]$, the minimum and maximum ram speeds are $\underline{v}_{ram} = 2.5\ mm/s$ and $\overline{v}_{ram} = 15\ mm/s$, and the maximum coolant valve opening is $\overline{z}_{valve} = 35$ %. The maximum and minimum tapers are respectively $\overline{\tau}_x = \overline{\tau}_r = 200\ ^\circ\text{C}/m$ and $\underline{\tau}_x = \underline{\tau}_r = 0\ ^\circ\text{C}/m$, and the billet length is $L = 0.9\ m$. The initial billet temperature is constrained to $\boldsymbol{T}_{billet,0} \geq 400\ ^\circ\text{C}$, and the *plate* section, which is affected directly by cooling, is constrained to $T_{plate} \geq 300\ ^\circ\text{C}$.

For all simulations, the initial section temperatures are based on observed initial temperatures for a real press. The exact initial temperatures used in this paper are given in Table 1.

The NLP in Eq. (17) is solved in MATLAB using IPOPT (Wächter and Biegler, 2006) interfaced from CasADi v.3.5.5 (Andersson et al., 2019). The computations are performed on a *Dell Precision 5560* with the following processor: *11th Gen Intel(R) Core(TM)*

i7-11850H @ 2.50GHz, 2496Mhz, 8 Cores, 16 Logical Processors.

Table 1: The initial conditions used for all simulations, given in [°C], where the different entries represent the different control volumes/cells in the section. The initial billet and container states are radially uniform, thus only the axial distribution is shown, with the values at the top being to the left in Fig. 5. The other sections only have one cell in the axial direction, and the values shown here represent the cells that are stacked radially in Fig. 5.

$\boldsymbol{T}_{billet,0}$	$\boldsymbol{T}_{cont.,0}$	$\boldsymbol{T}_{feeder,0}$	$T_{plate,0}$	$\boldsymbol{T}_{die,0}$
482.75	444.75	463.25	480.45	487.65
477.67	440.62	456.65		488.45
472.59	436.50	450.05		489.05
467.50	432.37	443.45		489.25
462.42	428.24	436.85		489.45
457.34	424.11			489.45
452.26	419.99			
447.18	415.86			
442.10	411.73			
437.01	407.60			
431.93	403.48			
426.85	399.35			

5.2 Nominal Operation

First, we present a simulation in which all initial state values and control variables are fixed, providing a nominal extrusion scenario that serves as a baseline for comparison. The baseline is used only to compare the effects of different process degrees of freedom and is not intended to represent optimized or best industrial practice. For this simulation, cooling is not used; $\hat{z}_k \equiv 0$, the ram speed is constant at $v_{ram,k} \equiv 3\ mm/s$, and the initial taper is radially uniform and linearly decreasing from 482.8 °C to 426.9 °C, front to back. The solution is shown in Fig. 9. The baseline uses a prescribed constant ram speed and initial taper akin to what may be specified by an operator.

5.3 Varying Process Degrees of Freedom

The optimal ram speed trajectories for all the configurations are shown in Fig. 10, and they all achieve isothermal extrusion (satisfy peak temperature requirements). All configurations with free initial taper produce an initial taper consisting of a uniform temperature equal to the minimum temperature of 400 °C, except for 'taper', which exhibits a minor radial gradient and has an axial taper of 535.9 °C → 429.6 °C (front to back). In other words, apart from the 'taper' configuration, each initial cell temperature is minimized in all configurations. The 'taper' solution yields an initial taper close to the one used in the baseline, which is expected because the other control variables are unchanged and the baseline already satisfies the peak temperature requirement. In the 'ram-valve' and 'full' configurations, the valve opening remained at its upper bound: 35 % opening. Notice that the extrusion time is therefore limited by the maximum permissible valve opening, thus this amount to another important production-limiting constraint.

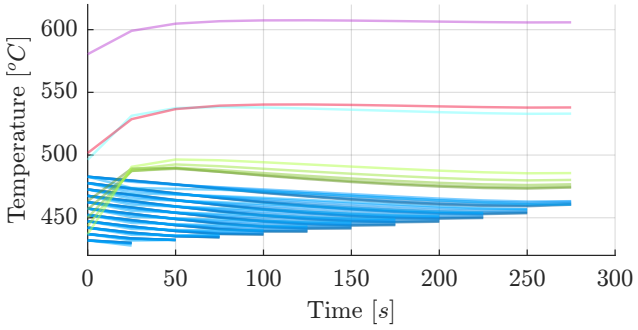

Figure 9: Result of a simulation of the extrusion phase using data-based initial conditions and control inputs given in Table 1 and Section 5.2, to act as a baseline simulation. The figure shows the aluminium temperatures over time. Colors: -peak, -port, -bridge, -feeder, -billet. Multiple graphs have the same color (varying shade) if they represent control volumes within the same section (see Fig. 5).

The computation times and number of iterations required to solve Eq. (17) for the configurations presented in this section are presented in Table 2, along with the respective extrusion times of the solutions and their improvement compared to the baseline simulation. We see that the computation times are shorter than their respective modeled extrusion times, suggesting that billet-to-billet planning may be computationally feasible. The extrusion times decrease with increasing control freedom, with the 'full' version reducing the extrusion time by 51.2 % from the baseline. By comparing the 'ram-valve' version to the 'full' version, we see that the introduction of initial taper as a control variable reduces the extrusion time by $(3\ min\ 13\ s - 2\ min\ 26\ s)/3\ min\ 13\ s = 24.3$ %. Note that these reductions reflect only the particular model, problem formulation, and chosen baseline, and should not be taken as a direct improvement on current industrial practices.

The results are also conditional on the ram-speed and temperature constraints being sufficiently conser-

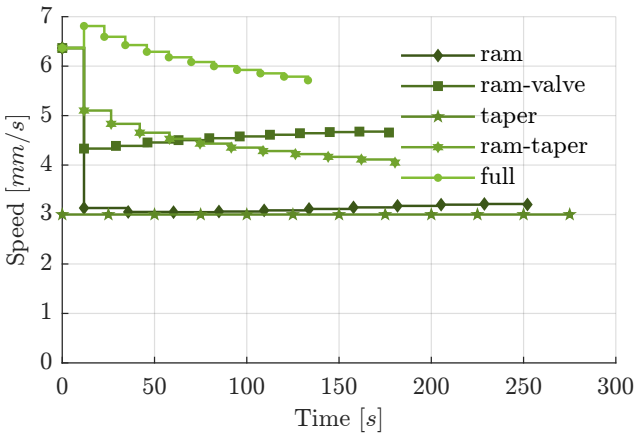

Figure 10: Optimal ram speed trajectories computed as the solutions to the configurations of Eq. (17) presented in Section 5.3, with the initial conditions from Table 1.

vative to prevent violations of the extrusion pressure limits, as discussed in Section 2.2.

Table 2: The computation times and number of iterations to solve the various NLP configurations described in Section 5.3, and the respective extrusion times, t_f, for the obtained solutions, where $\mathbf{t}_{\text{comp}}$ is the computation time, $\mathbf{N}_{\text{iter}}$ is the number of solver iterations, and "Red." is the reduction in t_f relative to the baseline: $(100\ \%) \cdot (t_{f,base} - t_{f,conf.})/t_{f,base}$.

Config.	$\mathbf{t}_{\text{comp}}$	$\mathbf{N}_{\text{iter}}$	$\mathbf{t_f}$	**Red.**
sim.	0.39 s	6	5 min	
ram	0.67 s	11	4 min 36 s	8.2 %
ram-val.	1.65 s	23	3 min 13 s	35.6 %
taper	22.4 s	165	5 min	
ram-tap.	1 min 58 s	845	3 min 19 s	33.7 %
full	13.9 s	109	2 min 26 s	51.2 %

A brief note on Fig. 10: All optimal ram speed trajectories (for which the ram speed was free) use 6.4 mm/s as their first value. We attribute this to the steady-state approximation of the peak temperature. Since the peak temperature is determined by algebraic equations rather than dynamic equations, it does not itself have an initial value, which is instead determined by the initial values of adjacent control volumes and the ram speed. Given the same initial conditions for all configurations, the initial map from ram speed to peak temperature is also the same, hence the same ram speed of 6.4 mm/s is chosen in all cases. For the subsequent sample points, the initial taper influences the optimal trajectories via the dynamic model.

5.4 Productivity Sensitivity to the Minimum Initial Taper

The results suggest that the initial taper is as 'flat' as possible, which means that the gradient of the initial taper is minimized, and its values reach the lower bound of 400 °C. However, as mentioned in Section 2.2, a practical implementation must preclude air entrapment during ram engagement by imposing a nonzero minimum taper ($\underline{\tau}_x > 0$), but it is not immediately clear how to choose an appropriate value. A larger minimum taper provides a greater safety margin, and a smaller taper facilitates short extrusion times. The relationship between the minimum-taper requirement and the minimum attainable extrusion time must therefore be evaluated on a press-to-press basis. This relationship may also vary from alloy to alloy and from profile to profile; thus, it is highly desirable to have a tool that allows one to study it in detail without the need for extensive experimental trials for each new design. We emphasize that experimentally determining the minimum extrusion time for a given initial taper would generally require multiple trials to identify the corresponding ram-speed and cooling trajectories. Our control problem simplifies this procedure by simultaneously optimizing the corresponding control trajectories. We now demonstrate that our control problem implementation can be used to efficiently investigate the sensitivity of the minimum attainable extrusion time with respect to the imposed minimum axial taper.

Multiple versions of the 'full' NLP configuration from Section 5.3 are solved with increasing minimum taper, i.e. an increasingly strict lower bound on the minimum allowable taper. The different NLP formulations constrain the initial taper to values ranging from $\underline{\tau}_x = 0$ °C/m to $\underline{\tau}_x = 150$ °C/m, and their solutions are presented in Fig. 11. For all restrictions, the optimal initial taper coincides with the minimum taper. In other words, in our setup, all optimal initial tapers are as 'flat' as permitted, such that the choice of minimum taper effectively amounts to choosing the initial taper directly. Therefore, the minimum taper directly limits the extrusion time, suggesting that the minimum taper is an important design parameter. The figure reveals an approximately affine trend, where the extrusion time increases on average by about 3.7 s per additional 10 °C/m axial taper. Note that since the heating capability is not explicitly accounted for in the optimization problem, the feasibility of a selected initial taper should be addressed separately.

We compare the results with simulations from a simulator used in industry that is based on the same un-

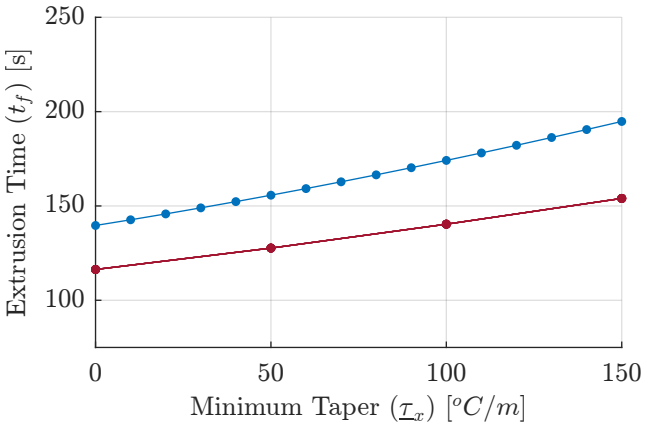

Figure 11: Blue: extrusion times versus the minimum taper imposed on the 'full' configuration from Section 5.3. Red: results from an industrial extrusion press simulator with an online NMPC optimizer. The figure shows an approximately affine map in both cases, with slopes of the same order of magnitude.

derlying extrusion press, alloy and profile, and a similar modeling approach. The simulator also uses an NMPC feedback controller, providing a closed-loop comparison to the open-loop optimization results. The simulation results are shown in red in Fig. 11, for four different initial tapers; $\tau_x \in \{0, 50, 100, 150\}$. The results show that the simulations also yield an approximately affine increase in extrusion time with increasing initial taper. We clarify that the simulator uses the same initial tapers, but does not use the corresponding optimal ram speed trajectory as control references, since doing so would largely prescribe the extrusion time. Furthermore, it uses periodic steady-state initial conditions for the remaining sections[2]. The simulations yielded shorter extrusion times than those predicted by the optimization problem solutions. Because the two approaches differ in both model and control implementation, the extrusion times are not expected to be equal. Nonetheless, the slopes of the affine relationships are of the same order of magnitude. The open-loop solutions computed by our implementation suggested an average slope of 0.37 $s \cdot m/^\circ C$, and the closed-loop simulation results suggest 0.25 $s \cdot m/^\circ C$.

[2]The remaining initial conditions were based on the terminal temperatures of the simulation, to emulate a periodic solution over multiple cycles.

6 Conclusion

The main result is that the imposed minimum axial taper is active at the optimum throughout the investigated range, in the absence of active pressure constraints, and therefore acts as a productivity-limiting design constraint in the considered extrusion model. The resulting minimum extrusion time varies approximately affinely with the taper requirement, and the same qualitative relationship is observed in an industrial closed-loop simulator.

As a concrete example, the results presented in Fig. 11 show that with a 10 °C/m reduction, the optimal extrusion time is reduced by almost 4 s for the particular extrusion model used in this paper. Accordingly, reducing the minimum taper from 100 °C/m to 50 °C/m reduces the extrusion time from 174.1 s to 155.7 s. From a complete extrusion cycle perspective, assuming a conservative reload time of $t_{RL} = 30$ s, the corresponding extrusion periods are $\mathcal{T}_{ext,(\underline{\tau}_x=100)} = 204.1$ s and $\mathcal{T}_{ext,(\underline{\tau}_x=50)} = 185.7$ s, corresponding to production rates of $3600\ s/\mathcal{T}_{ext,(\underline{\tau}_x=100)} = 17.6\ billets/hour$ and $3600\ s/\mathcal{T}_{ext,(\underline{\tau}_x=50)} = 19.4\ billets/hour$ respectively. This amounts to a 10 % increase in production rate under the aforementioned assumptions, suggesting that properly considering the safety-vs.-production trade-off is important. Dynamic optimization methods may aid in analyzing this trade-off and in the design of the initial taper.

We also compare configurations with different available process degrees of freedom, isolating the effect of the three main control channels and highlighting the importance of their joint consideration. Section 5.3 shows that by solving a dynamic optimization problem over the entirety of the extrusion phase horizon, the extrusion time progressively decreases as additional process degrees of freedom are included in the optimization, as summarized in Table 3. These results indicate that using an optimization layer to jointly consider the initial taper and planned control trajectories (ram speed and cooling) may meaningfully impact the extrusion time. However, realizable improvement in an industrial application will heavily depend on model accuracy, the choice of constraints, and assumptions about unmodeled constraints.

In all configurations where both the ram speed and initial taper are optimized, the optimal initial taper coincides with the minimum-taper constraint Eq. (17g), while its lowest temperature coincides with the minimum billet-temperature constraint ($\boldsymbol{T}_{billet} \geq 400$ °C). This is equivalent to minimizing the temperature in each individual control volume in the billet, thus minimizing the thermal energy in the billet.

Under parameter variations, such as drift or disturbances, it is desirable to re-plan often, preferably on

Table 3: Reduction in extrusion time obtained by adding the decision variable(s) in the left column to the starting configuration in the top row. IT: initial taper; RS: ram speed; VO: valve opening.

	Starting configuration			
Add	**IT**	**RS**	**RS+VO**	**RS+IT**
RS	33.7 %	–	–	–
VO	–	29.9 %	–	26.4 %
RS+VO	51.2 %	–	–	–
IT	–	27.8 %	24.3 %	–

a billet-to-billet basis. In this case, the average computation time should be smaller than the average extrusion period, preferably with some margin to optimize a subsequent heating phase problem. We see in Table 2 that the computation times are in the range $\langle 0\ s, 2\ min]$, and the extrusion times are in the range $t_f \in [2\ min, 6\ min]$. In our computational setup, these results support the computational feasibility of billet-to-billet planning. Specialized optimal control solvers and improved hardware may reduce the computation times even further.

Future work should experimentally validate the proposed optimization framework and further reconcile the extrusion model with the industrial simulator. Important extensions include incorporating explicit pressure constraints and accounting for the attainability of optimized initial tapers through the heating and transit phases.

Acknowledgments

This work was funded by the "ExtruTeC" project of the Research Council of Norway (Project No. 313919), in partnership with Hydro Extruded Solutions AS, Cybernetica AS, SINTEF Industry, and the Department of Engineering Cybernetics at the Norwegian University of Science and Technology.

References

Andersson, J. A. E., Gillis, J., Horn, G., Rawlings, J. B., and Diehl, M. CasADi: a software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*, 2019. 11(1):1–36. doi:10.1007/s12532-018-0139-4.

Aukrust, T. and LaZghab, S. Thin shear boundary layers in flow of hot aluminium. *International Journal of Plasticity*, 1999. doi:10.1016/S0749-6419(99)00047-9.

Barella, S., Gruttadauria, A., Gerosa, R., Mainetti, G., and Mainetti, T. Predictive Tools for in-Line Isothermal Extrusion of 6xxx Aluminum Alloys. In *The 1st International Electronic Conference on Metallurgy and Metals*. MDPI, page 24, 2021. doi:10.3390/IEC2M-09239.

Biegler, L. T. *Nonlinear programming: concepts, algorithms, and applications to chemical processes*. Number 10 in MOS-SIAM series on optimization. SIAM, Philadelphia, Pa, 2010. doi:10.1137/1.9780898719383.

Chanda, T., Zhou, J., and Duszczyk, J. A comparative study on iso-speed extrusion and isothermal extrusion of 6061 Al alloy using 3D FEM simulation. *Journal of Materials Processing Technology*, 2001. 114(2):145–153. doi:10.1016/S0924-0136(01)00724-5.

Cuéllar Matamoros, C. F. *Modeling and control for the isothermal extrusion of aluminium*. Ph.D. thesis, ETH Zurich, 1999. doi:10.3929/ETHZ-A-003821721.

Danilin, V. N., Aleshchenko, A. S., Danilin, A. V., and Koshmin, A. N. Simulation of taper heating and variable pressing rate to improve extrusion performance for high-strength aluminum alloys. *Modelling and Simulation in Materials Science and Engineering*, 2024. 32(6):065006. doi:10.1088/1361-651X/ad56a6.

Diehl, M. and Gros, S. Numerical Optimal Control (Draft). 2024. URL `https://www.syscop.de/files/2024ws/NOC/book-NOCSE.pdf`.

Farjad Bastani, A., Aukrust, T., and Brandal, S. Optimisation of flow balance and isothermal extrusion of aluminium using finite-element simulations. *Journal of Materials Processing Technology*, 2011. 211(4):650–667. doi:10.1016/j.jmatprotec.2010.11.021.

Gabrielsen, T. A. L. and Imsland, L. S. Progressor Transformation of Dynamical Systems and Application to Dynamic Optimization. In *2023 IEEE Conference on Control Technology and Applications (CCTA)*. IEEE, Bridgetown, Barbados, pages 496–501, 2023. doi:10.1109/CCTA54093.2023.10252981.

Guo, C., Qiao, J., Wang, W., Sun, D.-Z., and Li, Q. Optimization strategy for the indirect isothermal extrusion of high-strength aluminum alloy based on temporal fusion transformer. *The International Journal of Advanced Manufacturing Technology*, 2023. 126(11-12):5179–5189. doi:10.1007/s00170-023-11370-9.

Hölker, R., Haase, M., Khalifa, N. B., and Tekkaya, A. E. Hot Extrusion Dies with Conformal Cooling Channels Produced by Additive Manufacturing. *Materials Today: Proceedings*, 2015. 2(10):4838–4846. doi:10.1016/j.matpr.2015.10.028.

Hölker, R., Jäger, A., Ben Khalifa, N., and Tekkaya, A. E. Controlling heat balance in hot aluminum extrusion by additive manufactured extrusion dies with conformal cooling channels. *International Journal of Precision Engineering and Manufacturing*, 2013. 14(8):1487–1493. doi:10.1007/s12541-013-0200-1.

Iddberg, M., Myhr, O. R., Nesse, A., and Furu, T. A Modelling Framework for Rapid Evaluation of Speed Limitations during Extrusion of Aluminium Profiles. *Key Engineering Materials*, 2024. 988:31–46. doi:10.4028/p-G6Pyxv.

Kathirgamanathan, P. and Neitzert, T. Optimal Process Control Parameters Estimation in Aluminium Extrusion for Given Product Characteristics. 2008. URL https://www.iaeng.org/publication/WCE2008/WCE2008_pp1436-1441.pdf. ISBN:978-988-17012-3-7.

Kulås, M. *Dynamic Optimization and Parameter Estimation used on an Extrusion Press for Aluminium*. Master's thesis, Norwegian University of Science and Technology, Trondheim, 2022. URL https://jaschke.folk.ntnu.no/Masters/MasterTheses/2022/MatsKulas/.

Lela, B., Musa, A., and Zovko, O. Model-based controlling of extrusion process. *The International Journal of Advanced Manufacturing Technology*, 2014. 74(9-12):1267–1273. doi:10.1007/s00170-014-6054-6.

Li, L., Zhou, J., and Duszczyk, J. A 3D FEM simulation study on the isothermal extrusion of a 7075 aluminium billet with a predetermined non-linear temperature distribution. *Modelling and Simulation in Materials Science and Engineering*, 2003. 11(3):401–416. doi:10.1088/0965-0393/11/3/311.

Rawlings, J. B., Mayne, D. Q., and Diehl, M. M. *Model Predictive Control: Theory, Computation and Design*. Nob Hill Publishing, LLC, 2017. URL https://sites.engineering.ucsb.edu/~jbraw/mpc/.

Reggiani, B. and Donati, L. Prediction of liquid nitrogen die cooling effect on the extrusion process parameters by means of FE simulations and experimental validation. *Journal of Manufacturing Processes*, 2019. 41:231–241. doi:10.1016/j.jmapro.2019.04.002.

Sheppard, T. *Extrusion of Aluminium Alloys*. Springer US, Boston, MA, 1999. doi:10.1007/978-1-4757-3001-2.

Strikwerda, J. C. *Finite difference schemes and partial differential equations*. Society for Industrial and Applied Mathematics, Philadelphia, Pa, 2nd edition, 2004. doi:10.1137/1.9780898717938.

Tibbetts, B. and Wen, J. T.-Y. Extrusion process control: modeling, identification, and optimization. *IEEE Transactions on Control Systems Technology*, 1998. 6(2):134–145. doi:10.1109/87.664181.

Wächter, A. and Biegler, L. T. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 2006. 106(1):25–57. doi:10.1007/s10107-004-0559-y.

Zhou, J., Li, L., and Duszczyk, J. Computer simulated and experimentally verified isothermal extrusion of 7075 aluminium through continuous ram speed variation. *Journal of Materials Processing Technology*, 2004. 146(2):203–212. doi:10.1016/j.jmatprotec.2003.10.018.

Zhou, L., Liu, Y., Zhang, Y., and Yin, X. A Co-Simulation Approach Based on Close-Loop Feedback Control for Aluminum Alloys Profile Extrusion. 2023.