Simultaneous mixed-integer dynamic scheduling of processes and their energy systems

Florian Joseph Baader[a,b], André Bardow[a,c,d], Manuel Dahmen[a,*]

[a] Forschungszentrum Jülich GmbH, Institute of Energy and Climate Research, Energy Systems Engineering (IEK-10), Jülich 52425, Germany

[b] RWTH Aachen University Aachen 52062, Germany

[c] ETH Zürich, Energy & Process Systems Engineering, Zürich 8092, Switzerland

[d] RWTH Aachen University, Institute of Technical Thermodynamics, Aachen 52056, Germany

arXiv:2111.09842v1 [math.OC] 18 Nov 2021

Abstract: Increasingly volatile electricity prices make simultaneous scheduling optimization for production processes and their energy supply systems desirable. Simultaneous scheduling needs to account for both process dynamics and binary on/off-decisions in the energy system and thus leads to challenging mixed-integer dynamic optimization problems. In this contribution, we propose an efficient scheduling formulation that consists of three parts: a linear scale-bridging model for the closed-loop process output dynamics, a data-driven model for the process energy demand, and a mixed-integer linear model for the energy system. Process dynamics are discretized by collocation yielding a mixed-integer linear programming (MILP) formulation. We apply the scheduling method to a single-product reactor, with 5.6 % economic improvement compared to steady-state operation, and a multi-product reactor, with 5.2 % improvement compared to sequential scheduling. While capturing 85 % and 96 % of the improvement realized by a nonlinear optimization, the MILP formulation achieves optimization runtimes sufficiently fast for real-time scheduling.

Keywords:
Simultaneous scheduling, Demand response, Mixed-integer dynamic optimization, Mixed-integer linear programming, Integration of scheduling and control

*M. Dahmen, Forschungszentrum Jülich GmbH, Institute of Energy and Climate Research, Energy Systems Engineering (IEK-10), Jülich 52425, Germany
E-mail: m.dahmen@fz-juelich.de

1 Introduction

Current efforts to reduce greenhouse gas emissions increase the share of renewable electricity production in many countries. Due to the intermittent nature of renewable electricity production, stronger volatility in electricity prices or even electricity availability is expected (Merkert et al., 2015). This price volatility may offer economic benefits to industrial processes that can dynamically adapt their operation and thus their power consumption in so-called demand response (DR) (Mitsos et al., 2018). Ideally, demand response reacts to imbalances of electricity demand and supply and therefore also stabilizes the electricity grid (Zhang and Grossmann, 2016).

A promising way to achieve DR is to consider volatile prices in scheduling optimization (Merkert et al., 2015) that determines the process operation for a time horizon in the order of one day (Baldea and Harjunkoski, 2014; Daoutidis et al., 2018; Seborg et al., 2010). However, industrial processes are often not supplied directly by the electricity grid but by a local on-site multi-energy system. The local multi-energy system supplies all energy demanded by the process, e.g., heating, cooling, or electricity, and exchanges electricity with the grid (Voll et al., 2013). Operating local energy systems is a complex task as these systems typically consist of multiple redundant units with non-linear efficiency curves and minimum part-load constraints leading to discrete on/off-decisions (Voll et al., 2013). Thus, the electricity exchange between the energy system and the grid is not directly proportional to the process energy demand. Consequently, optimal DR scheduling must consider processes and their energy systems simultaneously. Moreover, such a simultaneous scheduling can improve the efficiency of energy system operation by shifting process energy demand in time (Bahl et al., 2017). Still, scheduling is usually carried out sequentially: The process schedule is optimized first and only then the energy system operation is optimized (Agha et al., 2010; Leenders et al., 2019).

The simultaneous scheduling of processes and their energy systems leads to computationally challenging problems. Process scheduling can already be a very demanding task on its own if nonlinear process dynamics need to be considered (Flores-Tlacuahuac and Grossmann, 2010); therefore, considering dynamics is a major research topic in process systems engineering referred to as integration of scheduling and control (Baldea and Harjunkoski, 2014; Daoutidis et al., 2018; Harjunkoski et al., 2009; Engell and Harjunkoski, 2012; Beal et al., 2017; Flores-Tlacuahuac and Grossmann, 2006). For DR problems, process dynamics are often scheduling-relevant (Mitsos et al., 2018; Baldea and Harjunkoski, 2014; Daoutidis et al., 2018; Caspari et al., 2019; Otashu and Baldea, 2019) because the time to drive the process from one steady state to another steady state is often in the same order of magnitude as the electricity-price time steps.

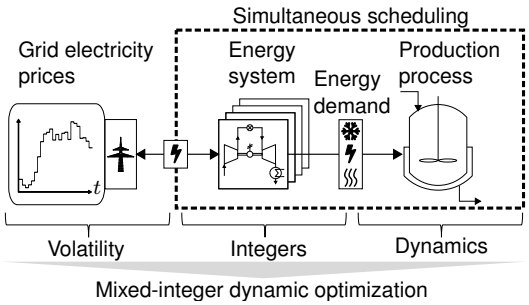

Fig. 1. Volatile grid electricity prices call for a simultaneous scheduling of production processes and their local energy supply systems. While energy systems introduce integer decision variables, processes often exhibit scheduling-relevant dynamics. Simultaneous scheduling thus results in computationally challenging mixed-integer dynamic optimization (MIDO) problems.

The desired simultaneous scheduling of processes and their energy systems is especially challenging due to the simultaneous presence of process dynamics and discrete on/off-decisions in the energy system (Figure 1). Because of the discrete decisions, standalone energy system optimization problems are preferably formulated as mixed-inter linear programs (MILPs) (Voll et al., 2013; Risbeck et al., 2017; Mitra et al., 2013; Carrion and Arroyo, 2006; Sass et al., 2020). A MILP formulation is usually applicable

because: (i) Nonlinear part-load efficiencies can be approximated reasonably well using piece-wise affine functions (Sass et al., 2020), and (ii) the dynamics of the energy system units are negligible or can be captured using ramping constraints (Sass and Mitsos, 2019).

As process dynamics are often scheduling-relevant, the simultaneous scheduling needs to be integrated with control. Even though conceptually, all approaches for the integration of scheduling and control can be used, the on/off-decisions significantly increase the computational complexity. However, scheduling must be performed online. Harjunkoski et al. (2014) state that generally optimization run times should be between 5 and 20 minutes.

In this work, we present a formulation for simultaneous scheduling of processes and their energy systems that aims at real-time-applicable runtimes. We rely on two promising approaches from the integration of process scheduling and control: (i) dynamic scale-bridging models (Du et al., 2015; Baldea et al., 2015), where the controlled process output is forced to follow a linear differential equation and (ii) dynamic data-driven models (Mitsos et al., 2018; Pattison et al., 2016; Kelley et al., 2018a,b). Specifically, our formulation consists of three parts: ① a scale-bridging model considering the dynamics of the production process, ② a piece-wise affine dynamic data-driven model for the energy demand of the process, and ③ a MILP energy system model with piece-wise affine approximations of nonlinear component efficiency curves. We discretize the linear differential equations in time using a high-order collocation scheme to receive linear constraints (Biegler, 2010). Consequently, we achieve an MILP formulation for the entire scheduling problem.

A preliminary version of our approach has been presented in a conference contribution (Baader et al., 2020) where we considered DR for a building energy system. In the present contribution, we describe our method in more detail and apply it to a chemical production system, i.e., a continuous-stirred-tank-reactor (CSTR) cooled by three compression chillers. Furthermore, the new method is explicitly compared against a standard sequential scheduling approach from industrial practice (Agha et al., 2010). The remainder of this paper is structured as follows: In section 2, the method is described in detail; in section 3, a first case study considering a multi-product scenario is performed; in section 4, a second case study considering a single-product scenario is performed, and section 5 concludes the work.

2 Method

In this section, we present our method for simultaneous dynamic scheduling of production processes and their energy systems. We refer to our method as *simultaneous dynamic scheduling*. The core of *simultaneous dynamic scheduling* is an efficient scheduling model consisting of three parts: ① the production process, ② the energy demand, and ③ the energy system (Figure 2). Model ① determines the controlled process output y_{cv}, e.g., the concentration in a reactor. We use a scale-bridging model (SBM) proposed by Baldea and co-workers (Du et al., 2015; Baldea et al., 2015) that describes a linear closed-loop response and represents the slow scheduling-relevant dynamics only. A linear SBM can be incorporated in scheduling optimization much more efficiently than a nonlinear full-order process model. The SBM relies on an underlying control to enforce the desired linear closed-loop response. The closed-loop response describes the evolution of the controlled variable y_{cv} and its time derivatives depending on the set-point w_{SP}:

$$y_{cv} + \sum_{i=1}^{r} \tau_i \frac{d^i y_{cv}}{dt^i} = w_{SP} \tag{1}$$

In equation 1, r is the order of the SBM and τ_i are time constants. We discuss both order and time constants in section 2.1. To linearize the closed-loop response, we propose to place a set-point filter (Corriou, 2018) in front of the controlled plant (Figure 2). This set-point filter converts the piece-wise constant set-point w_{SP} given by the scheduling optimization to a smooth filtered set-point $w_{SP,fil}$ that can be tracked by the underlying process control such that $y_{cv} \approx w_{SP,fil}$. In essence, we assume that the linear dynamics of the set-point filter can model the process output dynamics for the scheduling-relevant time-scale. Instead of the combination of set-point filter and tracking control previous publications used exact input-output feedback linearization (Du et al., 2015; Daoutidis and Kravaris, 1992; Corriou, 2018)

or scheduling-oriented model predictive control (SO-MPC) (Baldea et al., 2015). The proposed set-point filter increases the flexibility of the scale-bridging approach as it allows to use non-model-based tracking controls, e.g., PID-control (Corriou, 2018), as well.

The scale-bridging equation 1 is more than an approximation: Whenever the actual value of y_{cv} deviates from the closed-loop response described by equation 1 the underlying control acts to bring the controlled variable y_{cv} back to the desired closed-loop trajectory. Consequently, deviations of the controlled variable from its optimized trajectories are kept small. Note that, in this paper, we only discuss the case of a single controlled variable y_{cv}. Still, the scale-bridging approach is applicable to multi-input, multi-output processes, as discussed in Du et al. (2015). Consequently, our method can be extend to multi-input, multi-output processes.

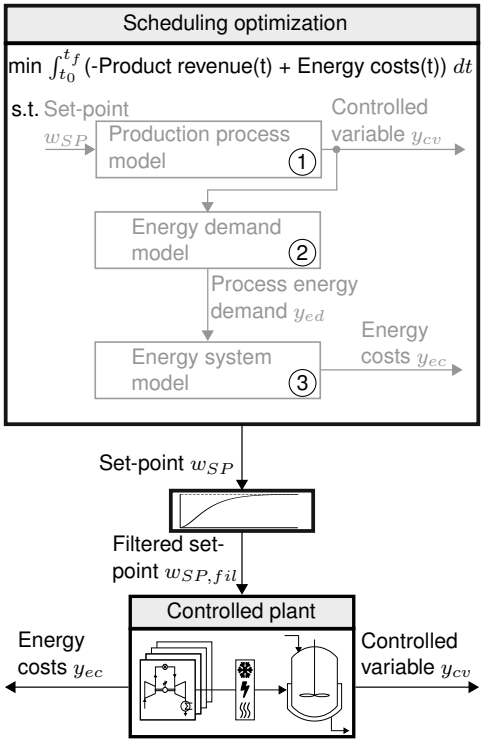

Fig. 2. Proposed simultaneous scheduling of processes and their energy systems based on our scheduling model consisting of three parts. A set-point filter converts the optimized piece-wise constant set-point w_{SP} to a smooth filtered set-point $w_{SP,fil}$, which defines the desired linear closed-loop process behavior.

Model ② is a dynamic data-driven model (Mitsos et al., 2018) that determines the process energy demand y_{ed} based on the current state of the production process, i.e., the controlled variable y_{cv} and its time derivatives. In principle, a wide range of data-driven models derived from recorded data, or mechanistic models can be used here. Examples of data-driven models being applied successfully in dynamic demand response optimization can be found in Pattison et al. (2016); Kelley et al. (2018a,b); Tsay and Baldea (2019). Our energy demand model ② can be dynamic and mixed-integer but must be linear as we aim for an MILP formulation. In contrast to the controlled process output y_{cv}, there is no correction for deviations between the actual process energy demand y_{ed} and the model prediction. Instead, we assume that such deviations are compensated by the energy system, which is reasonable if (i) the energy system can react significantly faster than the process and (ii) the energy system has spare capacity larger than the maximum error of the data-driven model.

Model ③ is the energy system model that determines the energy costs depending on the energy demand. The structure of the energy system is modeled by energy balances that connect the energy system components with demands. Moreover, the efficiency of individual energy system components is modeled as a function of the part-load fraction. Thus, the required input power $P_{c,in}$ of an energy

system component c is a nonlinear function of the desired output power $P_{c,out}$ (Sass et al., 2020). To obtain a MILP formulation, we follow the established approach of modeling part-load efficiency curves as piece-wise affine functions (Voll et al., 2013). In general, piece-wise affine efficiency curves require binary variables. Binary variables increase the computational burden; however, they can be avoided if the input power is a convex function of the output power (Carrion and Arroyo, 2006), which is the case for many energy system components of practical relevance (Voll et al., 2013).

By combining the three models ① - ③, we receive a linear differential algebraic equation system (DAE) containing integers:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}_{SP}) = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y} + \mathbf{C}\mathbf{z} + \mathbf{D}\mathbf{w}_{SP} \tag{2}$$

$$\mathbf{0} = \mathbf{g}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}_{SP}) = \mathbf{E}\mathbf{x} + \mathbf{F}\mathbf{y} + \mathbf{G}\mathbf{z} + \mathbf{H}\mathbf{w}_{SP} \tag{3}$$

where $\mathbf{x}$ are the differential states, $\mathbf{y}$ are continuous variables, $\mathbf{z}$ are discrete variables, $\mathbf{w}_{SP}$ are set-points, t is time, $\mathbf{f}$ and $\mathbf{g}$ are functions that are linear in $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}_{SP}$, and $\mathbf{A} - \mathbf{H}$ are matrices. Note that all variables are functions of time although not stated explicitly to improve readability.

We choose a discrete-time MILP formulation for our simultaneous scheduling problem because for variable electricity prices, discrete-time formulations usually perform better than continuous-time formulations (Castro et al., 2009), as the electricity markets imposes a discrete time structures, e.g., hourly constant prices. As our model consists of linear differential equations, time discretization with collocation in discrete time leads to linear constraints (Biegler, 2010).

In the following, we discuss the scale-bridging model parameters (section 2.1) and the scheduling optimization problem (section 2.2).

2.1 Scale-bridging model parameters

For the scale-bridging model ①, we have to determine the order r and the time constants τ_i from equation 1, as well as upper and lower bounds for the set-point w_{SP}^{max}, w_{SP}^{min}, respectively. The order r reflects how many stages of inertia a change in the manipulated variable u has to overcome before changing the controlled variable y_{cv}. If a process model is available, the order r can be derived mathematically by analyzing the relative degree of the process model defined as the number of times the controlled variable y_{cv} has to be differentiated with respect to time until the manipulated variable u appears explicitly (Corriou, 2018). If no process model is available, the order r needs to be chosen based on knowledge or intuition about the main inertia of the process.

As discussed by Baldea et al. (2015), the choice of the time constants τ_i is critical for the performance of the scale-bridging approach:

On the one hand, if the time constants are too small, the scale-bridging dynamics are too fast and cannot be realized by the controlled process. On the other hand, if the time constants are too large, the scale-bridging dynamics are overly conservative and process flexibility is wasted. However, a rigorous way to tune the time constants τ_i is missing in the literature.

In this paper, we argue that the time constants τ_i need to be tuned simultaneously with the set-point bounds w_{SP}^{max} and w_{SP}^{min}. For illustration, we consider a transition of the controlled variable starting from a small value y_{cv}^{start} and ending at a new steady state with a higher value y_{cv}^{end} close to the maximum allowable value y_{cv}^{max}. To speed up the transition, scheduling optimization might choose a set-point $w_{SP,elevated}$ which is elevated above y_{cv}^{end} and even y_{cv}^{max} for a certain period of time. However, choosing an elevated set-point value can lead to dynamics that are too fast for the controlled process. In particular, if the time constants τ_i are small, the scale-bridging dynamics are already fast and an elevated set-point may drive the controlled variable to infeasible values. A trade-off arises because we want to choose small time constants in general but also want to avoid slow transitions towards the bounds of the controlled variable.

In our case studies, we tune the scale-bridging parameters using a simple heuristic relying on simulations (section 3.2.1). Alternatively, existing knowledge about the time constants of the process, or measurements can be used to calibrate the SBM.

2.2 Scheduling optimization problem

To derive a complete problem formulation based on equations 2 and 3, we add a suitable objective function, discretize time, and add inequality constraints to account for variable bounds, minimum part-load, and problem specific constraints.

The objective Φ in simultaneous DR scheduling is to maximize cumulative product revenue $\Phi_{Product}$ at final time t_f minus the cumulative energy costs Φ_{Energy} at final time:

$$\min \Phi = -\Phi_{Product}(t_f) + \Phi_{Energy}(t_f) \tag{4}$$

$$\frac{d\Phi_{Product}}{dt} = \sum_{p\in\mathbb{P}} K_p q_p \tag{5}$$

$$\frac{d\Phi_{Energy}}{dt} = \sum_{e\in\mathbb{E}} K_e P_e \tag{6}$$

$$\text{with} \quad \Phi_{Product}(t_0) = \Phi_{Energy}(t_0) = 0$$

Here, $\mathbb{P}$ is the set of products, q_p the flow rate of product p, and K_p the price of p. Similarly, $\mathbb{E}$ is the set of end-energy forms consumed, K_e is the time-dependent price of energy e, and P_e is the consumed power of energy e. t_0 denotes the initial time.

For time discretization, we use three time grids (Figure 3). Grid 1 is given by the electricity market and contains piece-wise constant electricity prices with time step Δt_{elec}, e.g., 1 h or 15 min. Grid 2 contains discrete decision variables $\mathbf{z}$ and piece-wise constant set-points $\mathbf{w}_{SP}$. The resolution of grid 2 should not be too fine as it increases the number of integer variables and thus the combinatorial complexity of the optimization problem. Still, it should be possible to alter discrete decisions $\mathbf{z}$ and set-points $\mathbf{w}_{SP}$ at least at every step change of electricity prices. Thus, the electricity price time step resolution should constitute a lower bound on the resolution of grid 2. Making grid 2 finer than grid 1 by selecting time steps $\Delta t_{dis} < \Delta t_{elec}$, gives a higher flexibility and thus might enable higher profits. We recommend to use time steps with lengths $\Delta t_{dis} = \frac{1}{n_1}\Delta t_{elec}$ with n_1 being a small natural number.

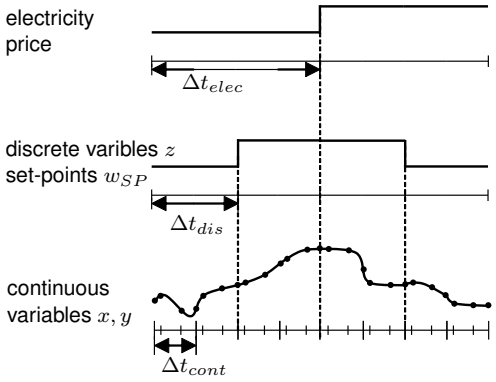

Fig. 3. Three time grids used for discretization with timesteps Δt_{elec}, Δt_{dis}, Δt_{cont}, respectively.

Grid 3 is used for continuous variables $\mathbf{x}, \mathbf{y}$. Differential states $\mathbf{x}$ are discretized using collocation. Similar to the argument above, we propose to use finite elements with length $\Delta t_{cont} = \frac{1}{n_2}\Delta t_{dis}$. The natural number n_2 is chosen to be greater than or equal to one because whenever electricity prices, discrete variables, or set-points perform step-changes, a new collocation element is necessary such that non-smoothness in differential states $\mathbf{x}$ is possible. In result, $\mathbf{x}$ are continuous at the border of collocation elements but first derivatives are allowed to perform step changes.

Within a finite element f_e of grid 3, a collocation polynomial x_{f_e} of order N_{cp} is used to discretize

differential states (Biegler, 2010; Nicholson et al., 2018):

$$x_{f_e}(\overline{\tau}) = \sum_{j=0}^{N_{cp}} l_j(\overline{\tau}) x_{f_e,j}, \qquad \overline{\tau} \in [0,1] \tag{7}$$

$$l_j = \prod_{k=0,k\neq j}^{N_{cp}} \frac{\overline{\tau} - \overline{\tau}_k}{\overline{\tau}_j - \overline{\tau}_k} \tag{8}$$

$$\left.\frac{dx}{dt}\right|_{t_{fe,k}} = \frac{1}{\Delta t_{cont}} \sum_{j=0}^{N_{cp}} x_{f_e,j} \frac{dl_j(\overline{\tau}_k)}{d\overline{\tau}} \tag{9}$$

In equations 7 and 8, the l_j are Lagrange basis polynomials, $\overline{\tau}$ is the scaled time within a finite element, and $x_{f_e,j}$ are state values at discretization points. In equation 9, $\left.\frac{dx}{dt}\right|_{t_{fe,k}}$ is the approximated time derivative at a collocation point k, which is set equal to the right hand side of the linear differential equation 2 for every time point $t_{f_e,k}$. The term $\frac{dl_j(\overline{\tau}_k)}{d\overline{\tau}}$ is a constant parameter in the optimization because it only depends on $\overline{\tau}$. Moreover, as we choose discrete time, Δt_{cont} is constant. Therefore, $x_{f_e,j}$ are the only optimization variables, and thus, discretization with equation 9 leads to linear constraints.

As inequality constraints, we consider upper and lower bounds for all variables, minimum part-load constraints for energy system components, and problem-specific constraints, e.g., minimum production targets. Minimum part-load constraints are realized with a binary variable $z_{on,c}$ that indicates if the output power $P_{c,out}$ of an energy system component c is zero or between the minimal and maximal allowed value, $P_{c,out}^{min}$ and $P_{c,out}^{max}$, respectively:

$$z_{c,on} P_{c,out}^{min} \leq P_{c,out} \leq z_{c,on} P_{c,out}^{max} \tag{10}$$

Assembling the discussed equations gives the simultaneous scheduling optimization problem for the production process and its energy system.

3 Case study 1: Multi-product reactor

In this section, we assess the computational performance of our method in a first case study considering a multi-product reactor. We benchmark the economic value of simultaneous dynamic scheduling to a standard sequential scheduling and to a nonlinear scheduling optimization with the true process model.

3.1 Setup

The setup of the case study is visualized in Figure 4. An exothermic multi-product CSTR is cooled with three compression chillers (CC). We use an exemplary reactor model from Petersen et al. (2017). In the CSTR, a component A reacts to a component B. The reactor can produce three products I, II, III, which are defined by the desired concentration of component A, C_A. We assume a small tolerance of +/- 0.01 $\frac{mol}{L}$ such that for each product, we obtain a product band. Whenever the concentration C_A is within one of the three product bands, the associated product is produced. If the concentration is outside of the three product bands, which happens necessarily during transitions, no product is produced. For illustration, we consider prices of 1, 0.75, and 0.5 money unit (MU) (Table 1) and require a minimum daily production of 5 hours and a maximum daily production of 8 hours for each product.

In the CSTR model, the rate of change for the concentration of component A is given by the material balance and the rate of change for the temperature T is given by the energy balance:

$$\frac{dC_A}{dt} = \frac{q}{V}(C_{A,feed} - C_A) - k e^{-\frac{E_A}{RT}} C_A \tag{11}$$

$$\frac{dT}{dt} = \frac{q}{V}(T_{feed} - T) - \frac{k \Delta H_r}{\varrho C_P} e^{-\frac{E_A}{RT}} C_A - \frac{Q_{cool}}{\varrho C_P V} \tag{12}$$

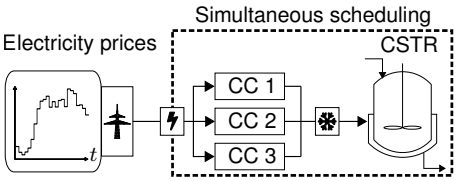

Fig. 4. Case study: Simultaneous scheduling of a continuous-stirred-tank-reactor (CSTR) cooled by three compression chillers (CC1, CC2, CC3). Time-varying electricity prices provide an economic incentive for DR.

Tab. 1. Product band $\left[C_{A,p}^{min}, C_{A,p}^{max}\right]$ in mol/L, price K_p^P in money unit (MU)/m^3, and cooling power in steady state production $Q_{cool,p}^{steady}$ in $\frac{MJ}{h}$, for products $p \in \{I, II, III\}$

p	$\left[C_{A,p}^{min}, C_{A,p}^{max}\right]$	K_p^P	$Q_{cool,p}^{steady}$
I	[0.09, 0.11]	1	6.05
II	[0.29, 0.31]	0.75	5.43
III	[0.49, 0.51]	0.5	4.65

In equations 11 and 12, q is the flow rate, V the reactor volume, $C_{A,feed}$ the feed concentration, k the reaction constant, E_A the activation energy, R the gas constant, T_{feed} the feed temperature, ΔH_r the enthalpy of reaction, ϱ the density, c_P the heat capacity, and Q_{cool} is the cooling provided to the reactor. The parameter values listed in Table 2 are exemplary values from Petersen et al. (2017), except that we varied the activation energy E_A to obtain an operating temperature where cooling with compression chillers is a realistic option.

An efficient chiller is used for base-load cooling, whereas chiller 2 has a medium coefficient of performance (COP), and chiller 3, which has a low COP, is used for peak cooling (see Table 3). We use the compression chiller model from Voll et al. (2013) with a minimum part load of 20 % and a coefficient of performance depending on nominal COP, $COP_{CC,i}^N$, cooling load $Q_{CC,i}$, and nominal cooling load $Q_{CC,i}^{max}$:

$$COP_{CC,i} = COP_{CC,i}^N(0.8615q_{CC,i}^3 - 3.5494q_{CC,i}^2 + 3.679q_{CC,i} + 0.0126), \text{with } q_{CC,i} = \frac{Q_{CC,i}}{Q_{CC,i}^{max}} \quad (13)$$

Note that our case study is meant to be an illustrative example rather than a real case. It allows us to study whether the proposed method is able to consider process dynamics and discrete on/off-decisions for energy system components simultaneously. We want to stress that even though the original nonlinear process model is a small-scale model, the resulting scale-bridging model (①) would have the same basic structure and computational complexity if a larger process model would be considered as the number of scheduling-relevant dynamics is typically small (Baldea and Harjunkoski, 2014; Baldea et al., 2015; Pattison et al., 2016).

Tab. 2. CSTR model parameters (adapted from Petersen et al. (2017))

symbol	value	unit
q	100	$\frac{m^3}{h}$
V	100	m^3
$C_{A,f}$	1	$\frac{mol}{L}$
k_0	$7.2 \cdot 10^{10}$	$\frac{1}{h}$
$\frac{E_A}{R}$	6500	K
T_f	350	K
$\frac{\Delta H_r}{R\varrho C_P}$	-209	$\frac{Km^3}{mol}$
ϱ	1000	$\frac{kg}{m^3}$
C_P	0.239	$\frac{J}{kgK}$

Tab. 3. Nominal cooling power $Q_{CC,i}^{max}$ and coefficient of performance $COP_{CC,i}^{N}$ for compression chillers

compression chiller	$Q_{CC,i}^{max}$ $[\frac{MJ}{h}]$	$COP_{CC,i}^{N}[-]$
1	4.8	6
2	2.3	4.5
3	1.5	3

We employ conventional PID control (Corriou, 2018) to track the filtered set-point for the concentration C_A by manipulating the cooling power Q_{cool}:

$$Q_{cool} = K_P * \left(e + \tau_D \frac{de}{dt} + \frac{1}{\tau_I} \int_0^t e dt \right) + Q_{cool}^{PID,0} \quad \text{, with } e = w_{SP,fil} - C_A \tag{14}$$

The controller parameters in equation 14 are: K_P, τ_D, τ_I, and $Q_{cool}^{PID,0}$. We choose $Q_{cool}^{PID,0}$ to be the steady-state cooling power of product II (Table 1) and manually tune the remaining controller parameters in a simulation such that the filtered set-point $w_{SP,fil}$ is tracked stable and accurately. The resulting parameters are: $K_P = 1000 \ \frac{MJ\ L}{h\ mol}$, $\tau_D = 0.1\ h$, and $\tau_I = 0.2\ h$. The stable and accurate set-point tracking is shown in the following (Figure 6).

3.2 Simultaneous dynamic scheduling

To apply our simultaneous dynamic scheduling method, we now set up the three parts of our model and the scheduling optimization problem as presented in Section 2.

3.2.1 Scale-bridging production process model

As discussed in section 2.1, we need to choose the order r, the time constants τ_i, and the set-point bounds w_{SP}^{min}, w_{SP}^{max} for the scale-bridging production process model ①. The natural order r in this case is 2, as can be shown from the physical process model: The manipulated variable Q_{cool} does not appear in the first derivative of the controlled variable C_A (equation 11). If equation 11 is differentiated with respect to time and the term $\frac{dT}{dt}$ is replaced using equation 12, the second time derivative $\frac{d^2C_A}{dt^2}$ appears as an explicit function of the input Q_{cool}. Thus, the system has the relative degree 2. The more descriptive explanation is that a change in the cooling power Q_{cool} has to first overcome the inertia of the temperature T and then the inertia of the concentration C_A.

Second-order systems are described in control theory by the time constant of their natural oscillation β and a damping coefficient ζ (Corriou, 2018). The two tunable time constants τ_1 and τ_2 become:

$$\tau_1 = 2\zeta\beta \tag{15}$$

$$\tau_2 = \beta^2 \tag{16}$$

Following Du et al. (2015), we choose a critically damped response, i.e., $\zeta = 1$, as we want to have fast but no oscillating dynamics.

In the following, we describe the heuristic procedure used to define the remaining time constant β simultaneously with the bounds for the set-point w_{SP}^{max} and w_{SP}^{min}. The allowed range of the set-point must at least cover the operating range of the concentration C_A which is between $C_A^{min} = 0.1\ \frac{mol}{L}$ and $C_A^{max} = 0.5\ \frac{mol}{L}$. However, as discussed in section 2.1, it is reasonable to allow elevated set-points in order to avoid overly conservative transitions towards the bounds of the concentration C_A. We introduce an elevation constant $w_{SP}^{elevation}$ and calculate the bounds of the set-point as:

$$w_{SP}^{max} = C_A^{max} + w_{SP}^{elevation} \tag{17}$$

$$w_{SP}^{min} = C_A^{min} - w_{SP}^{elevation} \tag{18}$$

We want to find a combination of β and $w_{SP}^{elevation}$ that (i) is feasible, i.e., the filtered set-point can be tracked accurately without oscillations, and (ii) allows for fast product transitions. In the following, we

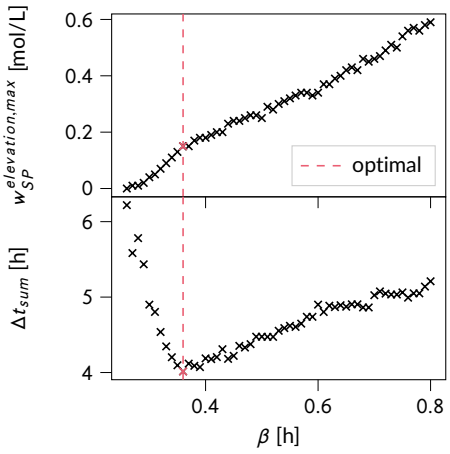

Fig. 5. Result of the parameter tuning: maximum allowable set-point elevation $w_{SP}^{elevation,max}$ (top) and sum of all 6 transition times Δt_{sum} resulting with $w_{SP}^{elevation,max}$ (bottom) for different values of the time constant β. We choose the optimal, i.e., smallest Δt_{sum} by setting $\beta = 0.36\ h$ and $w_{SP}^{elevation} = 0.15\ \frac{mol}{L}$.

first present a routine to evaluate the feasibility and speed for a given combination of β and $w_{SP}^{elevation}$ and then explain how we explore the space of possible combinations.

To evaluate a combination of β and $w_{SP}^{elevation}$, we first optimize and then simulate all 6 possible transitions between the 3 products bands. For a given combination of β and $w_{SP}^{elevation}$ and for each transition s, we perform the following four steps:

1. Optimize a sequence for the set-point w_{SP} using model ① to start from product p_{start} and reach the product band of product p_{end} as fast as possible.
2. Take the resulting set-point sequence as input to a simulation of the set-point filter, the underlying PID-control, and the nonlinear process model.
3. Based on the simulation result, check if a transition is feasible. We define a transition to be feasible if (i) the concentration reaches the product band of the desired product p_{end} and (ii) once the product band is reached the concentration stays inside the band of p_{end}.
4. Store the time needed to reach the product band Δt_s.

A combination of β and $w_{SP}^{elevation}$ is feasible if all 6 transitions are feasible. We evaluate the quality of feasible parameter combinations by the sum of all 6 transition times

$$\Delta t_{sum} = \sum_{s \in \mathbb{S}} \Delta t_s, \tag{19}$$

where $\mathbb{S}$ is the set of possible transitions. As we aim for fast transitions, we prefer feasible combinations of β and $w_{SP}^{elevation}$ with a small value of Δt_{sum}.

To explore the space of possible combinations, we first set the set-point elevation to zero, i.e., $w_{SP}^{elevation} = 0\ \frac{mol}{L}$, and search for the smallest time constant β that leads to feasible transitions. Next, we continue to increase β, i.e., we slow the dynamics down, but allow elevated set-points.

As Figure 5 shows, exploring the trade-off between set-point elevation and time constants improves the scale-bridging model performance significantly. The smallest, i.e., fastest, possible time constant $\beta^{min} = 0.26\ h$, which does not allow any set-point elevation, leads to a combined transition time of $\Delta t_{sum} = 6.22\ h$. The slightly higher time constant $\beta = 0.36\ h$ in combination with a set-point elevation of $w_{SP}^{elevation} = 0.15\ \frac{mol}{L}$ allows to reduce the transition time by 35% to the optimum of $\Delta t_{sum} = 4.02\ h$. The resulting transitions with the chosen optimal values are shown in Figure 6. Note that the set-point elevation is not strictly increasing with β due to the discretization. For example, with $\beta = 0.49\ h$, a

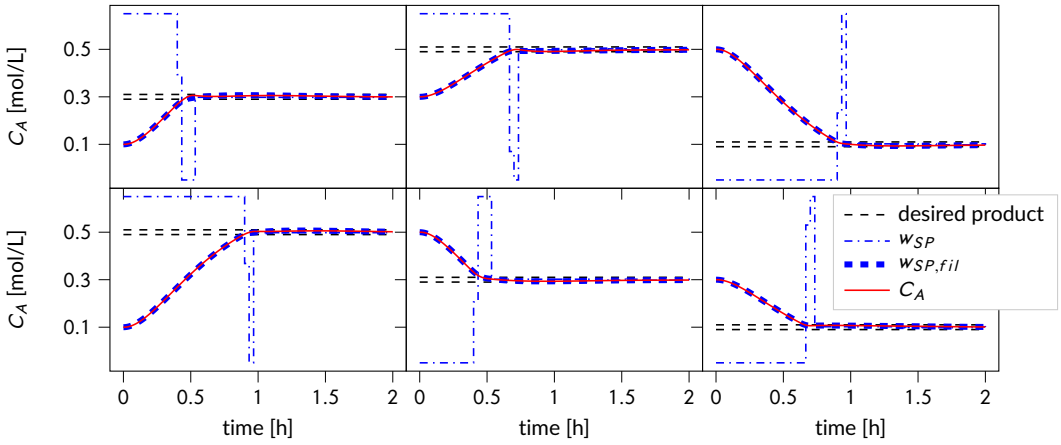

Fig. 6. Six possible transitions between the three products with the chosen time constant $\beta = 0.36\ h$ and set-point elevation $w_{SP}^{elevation} = 0.15\ \frac{mol}{L}$ (compare to Figure 5). The piece-wise constant set-point from optimization w_{SP} results in a filtered set-point $w_{SP,fil}$, which can be tracked accurately such that the actual value of the concentration C_A resulting from the controlled nonlinear process model matches the filtered set-point $w_{SP,fil}$ well.

set-point elevation of $w_{SP}^{elevation} = 0.26\ \frac{mol}{L}$ is feasible, whereas for $\beta = 0.50\ h$, $w_{SP}^{elevation} = 0.26\ \frac{mol}{L}$ is not feasible. The reason is that in one transition the set-point is at $w_{SP}^{max} = 0.76\ \frac{mol}{L}$ for one discretization time step longer with $\beta = 0.50\ h$ compared to $\beta = 0.49\ h$ leading to a slight overshoot of the concentration out of the product band.

3.2.2 Energy demand model

For part ②, the process energy demand model, we express the cooling power Q_{cool} as a function of the concentration C_A and its time derivatives. As the operation of the multi-product reactor is divided in production and transition periods and we need to model the cooling power accurately in particular during the long production periods, we split Q_{cool} into a steady-state and a dynamic part:

$$Q_{cool} = Q_{cool}^{steady} + Q_{cool}^{dynamic} \tag{20}$$

To approximate the first contribution Q_{cool}^{steady}, we assume that steady-state cooling powers are known for all three products and interpolate Q_{cool}^{steady} as a piece-wise affine function of C_A. The three steady-state operating points $C_A = \{0.1\ \frac{mol}{L}, 0.3\ \frac{mol}{L}, 0.5\ \frac{mol}{L}\}$ with corresponding cooling powers Q_{cool}^{steady} lead to two piece-wise affine segments: The first affine segment approximates Q_{cool}^{steady} for $C_A \leq 0.3\ \frac{mol}{L}$ and the second affine segment approximates Q_{cool}^{steady} for $C_A \geq 0.3\ \frac{mol}{L}$. With the binary variable z_{cool}^{steady}, we can express Q_{cool}^{steady} as

$$Q_{cool}^{steady} = Q_{cool,0.3\frac{mol}{L}}^{steady} + m_1^{steady}(1 - z_{cool}^{steady})(C_A - 0.3\ \frac{mol}{L}) + m_2^{steady} z_{cool}^{steady}(C_A - 0.3\ \frac{mol}{L}), \tag{21}$$

where $Q_{cool,0.3\frac{mol}{L}}^{steady}$ is the steady-state cooling power at $C_A = 0.3\ \frac{mol}{L}$, m_1^{steady} is the slope for $C_A \leq 0.3\ \frac{mol}{L}$, and m_2^{steady} is the slope for $C_A \geq 0.3\ \frac{mol}{L}$. The slopes m_1^{steady} and m_2^{steady} are calculated from the cooling power at steady-state operating points (Table 1). The bilinear terms $z_{cool}^{steady} C_A$ are reformulated using the Glover reformulation (Glover, 1975).

The approximation of $Q_{cool}^{dynamic}$ is fitted to simulation data. Again, we simulate all six possible transitions using the nonlinear reactor model and the underlying control. The resulting cooling power is the red curve in Figure 7. The total cooling power deviates from Q_{cool}^{steady} (dashed green curve in Figure

7) during transitions. We model the dynamic part of the cooling power $Q_{cool}^{dynamic}$ as a linear function of the derivatives of the concentration, i.e.,

$$Q_{cool}^{dynamic} = c_1 \frac{dC_A}{dt} + c_2 \frac{d^2 C_A}{dt^2}, \tag{22}$$

with the two fitting parameters c_1, c_2, whose values are determined using the normal equation method (Lewis et al., 2006). The values are: $c_1 = -2.98\ \frac{MJ}{h}$, $c_2 = 0.453\ \frac{MJ}{h}$. The resulting approximation of Q_{cool} is shown in blue in Fig 7. Note that in Figure 7, the concentration does not reach steady state after entering the product bands. Still, the fitted dynamic cooling power $Q_{cool}^{dynamic}$ is negligible for 5 of 6 production phases and only in the second production phase there is a small offset between model and actual cooling power.

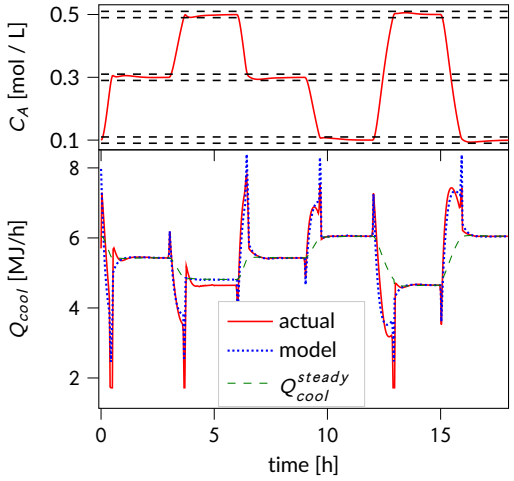

Fig. 7. Fitting results for steady-state and total cooling power Q_{cool}^{steady} and Q_{cool}, respectively, of the process energy demand model (compare to equations 20 and 21)

3.2.3 Energy system model

For the energy system model ③, we have to calculate the electric input power $P_{CC,i}$ needed for the compression chillers as a function of the required cooling power $Q_{CC,i}$. As the COP of the compression chillers depends on the part-load fraction (equation 13), $P_{CC,i}$ is a nonlinear function of $Q_{CC,i}$, which we approximate as a piece-wise affine function. We use two piece-wise affine segments per chiller with the breakpoint at 70 % part load; two segments provide a good approximation (Voll et al., 2013). The piece-wise affine curves can be modeled without introduction of additional binary variables, as the electric input power $P_{CC,i}$ is a convex function of the cooling power $Q_{CC,i}$. Using equations from Neisen et al. (2018), we introduce two continuous variables $y_{i,1}, y_{i,2}$ that cover the two affine segments:

$$Q_{CC,i} = y_{i,1} + y_{i,2} \qquad \forall i = 1,2,3 \tag{23}$$

$$P_{CC,i} = P_{CC,i}^{min} z_{on,CC,i} + y_{i,1} m_{i,1} + y_{i,2} m_{i,2} \qquad \forall i = 1,2,3 \tag{24}$$

In equation 24, $P_{CC,i}^{min}$ is the electric input power at minimum part-load of chiller i, $z_{on,CC,i}$ is a binary indicating whether chiller i is active, and $m_{i,1}, m_{i,2}$ are the slopes within the two piece-wise affine segments. Finally, we include the energy balance stating that the cooling demand of the reactor must be matched by the compression chillers:

$$\sum_{i=1}^{3} Q_{CC,i} = Q_{cool} \tag{25}$$

3.2.4 Scheduling optimization problem

In this case study, we use the objective function as introduced in the method section (equations 4 - 6). For time discretization, we have $\Delta t_{elec} = 1\ h$ as we consider hourly changing electricity prices. For discrete variables, we choose $\Delta t_{dis} = 15$ min so the optimization can differentiate between transitions that can be performed within 30, 45, or 60 minutes. This resolution is reasonable as the transition times are between 28 and 54 minutes (Figure 6). Note that though a finer discretization, e.g., 10 minutes, may lead to even higher profits, more binary variables will further increase the computational burden. For the continuous variables, we find a discretization of $\Delta t_{cont} = 15$ min with $N_{cp} = 3$ collocation points to give a sufficiently accurate time discretization. In section 3.5, we evaluate the optimization result in a simulation on the original nonlinear process model and thereby also verify the accuracy of our time discretization.

Finally, we add problem-specific constraints. Two constraints per product determine if a product p is produced, i.e, the binary variable z_p should be one whenever $C_{A,p}^{min} \leq C_A \leq C_{A,p}^{max}$ and zero in all other cases. As small variations between model ① and the nonlinear process model can occur, we add a small safety margin ε_p:

$$-(1 - z_p) \leq C_A - C_{A,p}^{min} - \varepsilon_p \qquad \forall p \tag{26}$$

$$-(1 - z_p) \leq -C_A + C_{A,p}^{max} - \varepsilon_p \qquad \forall p \tag{27}$$

We choose $\varepsilon_p = 0.003\ \frac{mol}{L}$ based on simulation results. If the process cannot be simulated, a safety margin needs to be chosen based on measurement data. It would be reasonable to first start with a conservative safety margin, evaluate the difference between model and plant data and then adjust the safety margin. Similar to the production binaries z_p, we need constraints that enforce the correct behavior of the binary z_{cool}^{steady} needed for the cooling model (compare to equation 21):

$$C_A - 0.29 \geq -(1 - z_{cool}^{steady}) \tag{28}$$

$$C_A - 0.31 \geq z_{cool}^{steady} \tag{29}$$

Lastly, we need to secure that the active chillers have more spare capacity than the maximum error of the data-driven model. As discussed in section 2, our assumption that errors in energy demand are compensated by the energy system is only valid if spare capacity is given. Consequently, we want to avoid a situation in which the scheduling optimization drives all active compression chillers to full load. If for example two chillers are active and operate at full load and the true nonlinear cooling power demand is slightly higher than the cooling power calculated from model ②, an unplanned start-up of chiller three has to be performed. However, we want to avoid excessive on/off-switches as they shorten the equipment life-time. To this end, based on the fit of model ②, we introduce a safety margin of $\varepsilon_Q = 0.1$ by setting

$$Q_{cool} \leq (1 - \varepsilon_Q) \sum_{i=1}^{3} z_{on,CC,i} Q_{CC,1}^{max}, \tag{30}$$

such that there always is a 10 % safety margin between the amount of cooling the active chillers can supply at full load and the cooling power calculated from model ②. Again, this safety margin could alternatively be determined based on measurements.

To speed up the solution time, we add the constraint that every product can only be produced once during the 24 hour horizon. Note that this additional constraint does not cut off the optimal solution: As all the transition times are between 0.5 and 1 hour, it is not reasonable to perform more transitions than necessary, because any additional transition would correspond to at least half an hour without product revenue, which clearly outweighs energy cost savings. The constraint limiting the number of production starts can be seen as a discrete time analogue to setting the number of production slots equal to the number of products in a continuous time formulation, see, e.g., Flores-Tlacuahuac and Grossmann (2006). To implement the constraint, we introduce a continuous variable $s_{p,t}$ that is equal to one, whenever production of product p starts in a timestep t, and zero otherwise. Here, we use the index t to indicate the timestep. For the sake of readability, this index is neglected in all the previous equations as the timestep is the same for all variables there. For every discrete timestep t, the following

equations from Neisen et al. (2018) are implemented such that $s_{p,t}$ is one if and only if the production binary $z_{p,t}$ is one and the production binary $z_{p,t-1}$ is zero:

$$0 \leq s_{p,t} \leq 1 \tag{31}$$
$$s_{p,t} \geq z_{p,t} - z_{p,t-1} \tag{32}$$
$$s_{p,t} \leq 1 - z_{p,t-1} \tag{33}$$
$$s_{p,t} \leq z_{p,t} \tag{34}$$
$$\sum_{t \in \mathbb{T}_{dis}} s_{p,t} \leq 1 \qquad \forall p \tag{35}$$

Note that one can easily change equation (35) to allow two or more production starts if desired.

To formulate the optimization problem, we use pyomo (Hart et al., 2017, 2011) and the extension pyomo.dae (Nicholson et al., 2018) to discretize time. We solve the problem using gurobi version 8.1.0 (Gurobi Optimization, LLC, 2021) with an optimality gap of 1.0 %. All calculations are performed on a Windows 10 machine with Intel(R) Core(TM) i5-8250U core and 24 GB RAM.

3.3 Sequential steady-state scheduling benchmark

This benchmark represents a typical sequential scheduling without DR, referred to as SEQ in the following. First, the process schedule is optimized with only the product revenue $\Phi_{Product}$ in the objective function. Second, the energy costs are minimized for fixed production decisions.

We perform a steady-state scheduling with predefined minimum transition times between products using the pre-optimized transition profiles from section 3.2.1. For the steady-state scheduling optimization, only the discrete time grid is necessary. At each point in time, the process is either in steady-state production or in transition:

$$\sum_{p \in \mathbb{P}} z_p + \sum_{s \in \mathbb{S}} z_s = 1 \tag{36}$$

In equation 36, z_s is a binary indicating if the process is in transition s with $\mathbb{S}$ being the set of possible transitions.

The cooling power is assumed to be equal to the steady-state cooling power $Q_{cool,p}^{steady}$ of a product p during production phases. During transition phases, we use the average cooling power $Q_{cool,s}^{average}$ within a transition s. This average power is calculated from the simulations shown in Figure 7 as the integral over the transition divided by the transition length. However, during transitions peak cooling demands vary significantly from average cooling demands. Consequently, the active chillers have to provide enough spare capacity to be able to supply the peak cooling demand $Q_{cool,s}^{peak}$ during a transition s. The calculation of the cooling power is given by equation 37 and the constraint to enforce spare capacity by equation 38:

$$Q_{cool} = \sum_{p \in \mathbb{P}} z_p Q_{cool,p}^{steady} + \sum_{s \in \mathbb{S}} z_s Q_{cool,s}^{average} \tag{37}$$

$$\sum_{s \in \mathbb{S}} z_s Q_{cool,s}^{peak} \leq \sum_{i=1}^{3} z_{on,CC,i} Q_{CC,i}^{max} \tag{38}$$

As the reactor cannot jump between products, a transition s must be at least as long as the minimum transition time Δt_s^{min}. For implementation, we introduce the variable Δt_s that indicates the time the process has been in a transition state s. We show the relevant equations for transition $s = I \rightarrow II$. Again, the index t is used to indicate the time step:

$$\Delta t_{I \rightarrow II,t} \leq z_{I \rightarrow II,t} \tag{39}$$
$$\Delta t_{I \rightarrow II,t} \leq \Delta t_{I \rightarrow II,t-1} + \Delta t_{dis} \tag{40}$$

The process can only be in the transition $s = I \rightarrow II$, if it already was in $s = I \rightarrow II$ in the last time step or if it was producing I:

$$z_{I\rightarrow II,t} \leq z_{I\rightarrow II,t-1} + z_{I,t-1} \tag{41}$$

The process can only produce product II, if it was producing II in last time step, or of it was in transition state $I \rightarrow II$ longer than the minimum time required for the transition ($\Delta t_{I\rightarrow II,t-1} \geq \Delta t^{min}_{I\rightarrow II}$), or if it was in transition state $III \rightarrow II$ longer than the minimum time ($\Delta t_{III\rightarrow II,t-1} \geq \Delta t^{min}_{III\rightarrow II}$). To implement this logic, we use big-M constraints (Bemporad and Morari, 1999):

$$z_{II,t} \leq 3z_{II,t-1}M + (\Delta t_{I\rightarrow II,t-1} - \Delta t^{min}_{I\rightarrow II})M + 3(1 - z_{II,t})M + 2z_{III\rightarrow II,t-1}M \tag{42}$$

$$z_{II,t} \leq 3z_{II,t-1}M + (\Delta t_{III\rightarrow II,t-1} - \Delta t^{min}_{III\rightarrow II})M + 3(1 - z_{II,t})M + 2z_{I\rightarrow II,t-1}M \tag{43}$$

In equations 42 and 43, M needs to be a sufficiently large constant. We choose $M = 2.4 \cdot 10^5$.

3.4 Scheduling with full nonlinear model

To estimate an upper bound on the economic performance of simultaneous scheduling, we optimize the nonlinear full-order system model. To this end, we replace the models ①, ②, ③ in the optimization problem with the nonlinear reactor model (equations 11 -12) and the nonlinear compression chiller efficiency (equation 13). Again, time is discretized using collocation and we receive a MINLP. We solve the MINLP optimization problem using BARON version 20.10.16 (Khajavirad and Sahinidis, 2018) in heuristic mode, i.e., the resulting solution is no rigorous bound. We refer to this benchmark as MINLP. To obtain a feasible initial point, we fix the binary variables to the values resulting from our simultaneous dynamic scheduling and solve the resulting NLP.

3.5 Results

In this section, we compare the economic profit obtained with our *simultaneous dynamic scheduling* (SDS) to the sequential scheduling (SEQ) and the full-order nonlinear scheduling (MINLP). While in case of the MINLP the profit is the objective value in the optimization, for the sequential scheduling and the *simultaneous dynamic scheduling*, the profit is derived from a simulation of the original nonlinear process model. Accordingly, the optimized set-point sequence is used as input to a simulation of the underlying controller and the nonlinear process model.

The MINLP solution improves the SEQ solution by 5.4 % (Figure 8). Our *simultaneous dynamic scheduling* gains 5.2 % compared to SEQ and thus captures 96 % of the MINLP improvement. The improved economics mainly stem from demand response, i.e., products with higher cooling demands are produced at times of lower electricity prices (Figure 9). Additionally, we notice a higher energy efficiency during transition phases such that our *simultaneous dynamic scheduling* reduces the amount of electricity consumed by 1.2 % compared to SEQ.

Figure 10 shows concentration C_A and cooling power Q_{cool} for both SDS and SEQ. In the case of SDS, the difference between the optimization model and the actual cooling power during the transitions is smaller than in the case of SEQ because the dynamics of the cooling power are modeled in SDS while SEQ only considers the average cooling power during a transition. Modeling the dynamics of the cooling power within a transition leads to better scheduling decisions regarding the on/off status of the three compression chillers. The most distinct difference occurs in the transition from product I to product III. In the sequential scheduling, this transition features a high cooling power peak, which requires to turn on chiller 3 with the worst COP. In the case of SDS, the same transition is shaped such that it is not necessary to turn on chiller 3. Moreover, SDS anticipates that chiller 2 with the medium COP can be turned off during the second half of the transition. We expect that energy efficiency improvements are even higher in cases with longer transitions. Naturally, the potential for energy efficiency improvements depends on the accuracy of the data-driven energy demand model (②). In this case study, we accounted for inaccuracies in model ② by adding a 10 % safety margin to the cooling power (equation 30) and still find significant energy efficiency improvements.

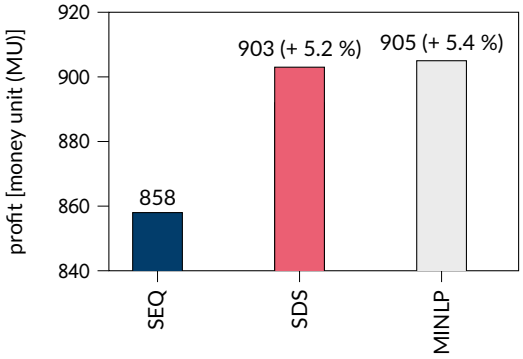

Fig. 8. Economic improvements obtained with the two demand response approaches *simultaneous dynamic scheduling* (SDS), and the full nonlinear model (MINLP) compared to sequential steady-state scheduling (SEQ) without DR.

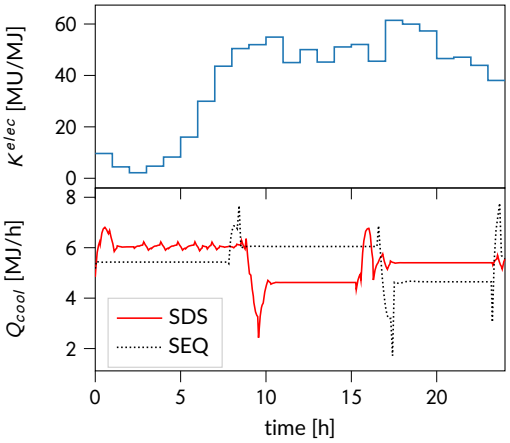

Fig. 9. Electricity price K^{elec} and simulated cooling power Q_{cool} for *simultaneous dynamic scheduling* (SDS) and sequential steady-state scheduling (SEQ). SDS performs demand response and shifts cooling power to times of favorable prices.

Note that in our illustrative example, we assume that once the energy system components are active they can react instantaneously. Moreover, we assume that the frequency of on/off-switches resulting from the scheduling optimization with 15 minute resolution for discrete variables is acceptable. In practice, it might be necessary to consider ramp limits (Sass and Mitsos, 2019), or minimum up and down times (Carrion and Arroyo, 2006). Such constraints can be added in a straight-forward manner to the formulation if needed.

A solution with a 1.0 % optimality gap is found and proven in 264 s. Such a solution time is applicable for both offline day-ahead scheduling and online re-scheduling during the day, e.g., with a sampling time of one hour. Note that in re-scheduling the solution from the last scheduling-iteration can be used for initialization to further speed up the optimization. We also observe that SDS finds feasible near-optimal solutions quickly as shown in the convergence plot (Figure 11). After 135 s a solution is found that has only 1.7 % gap to the final lower bound and already outperforms the sequential scheduling.

3.5.1 Influence of time constant β

If a nonlinear process model is not available, the time constant β cannot be chosen as in section 3.2.1 but needs to be chosen based on intuition or recorded product transitions. In result, the time constant might be suboptimal. To study the influence of the time constant choice on the profit, the time constant

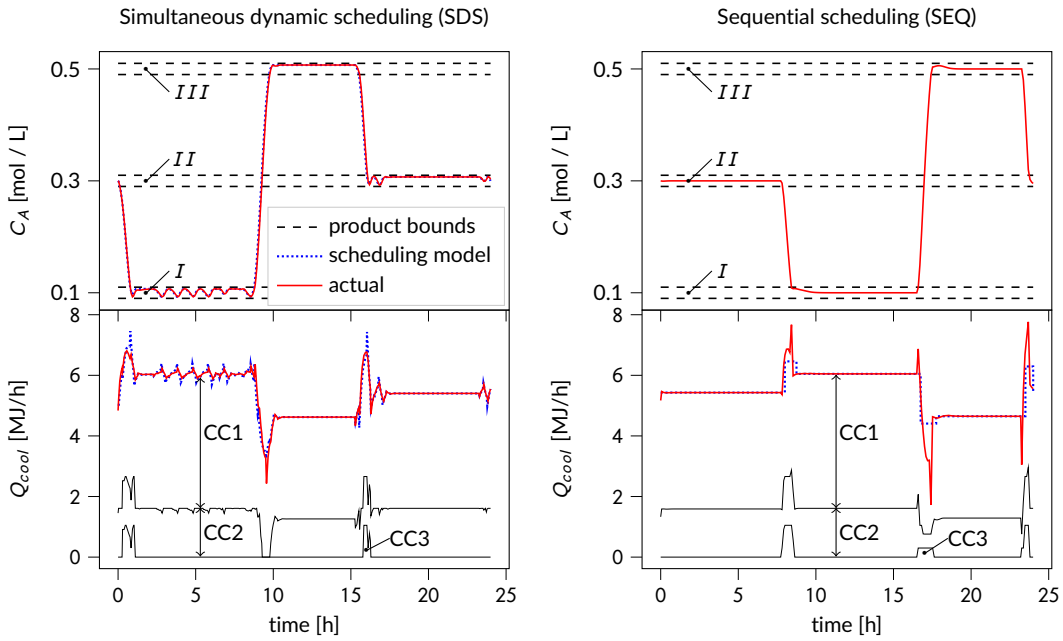

Fig. 10. Comparison of concentration C_A and cooling power Q_{cool} between simultaneous dynamic scheduling (SDS, left) and sequential scheduling (SEQ, right). We indicate the three product bands (I, II, III) and the fraction of the cooling power Q_{cool} that is supplied by the three compression chillers (CC1, CC2, CC3).

in our case study is increased from the optimal value $\beta = 0.36$ h by up to 100 % (Figure 12). Note that the profit is calculated in a simulation using the original process model which leads to small differences between scheduling optimization and process simulation and therefore the simulated objective shown in Figure 12 does not strictly increase with β.

As long as β is increased by 20 % or less the objective does not worsen more than 0.5 %. This result can be explained by the total production time which is 21.75 h for $\beta = 0.36$ h. These 21.75 hours of production are still reached for $\beta = 0.43$ h and the loss in profit is small. Generally, production time changes in 0.25 h steps corresponding to the time discretization of the binary variables (section 3.2.4). When β is increased by more than 20 % above the optimal value, production is lost and the objective substantially worsens. But, even for a 50 % increase, the objective is still better than that of the sequential solution.

4 Case Study 2: Reactor with variable concentration

While multi-product processes are one example for scheduling-relevant dynamics, single-product processes can also introduce dynamics if they can vary their controlled variable around a nominal value as long as the nominal value is reached on average over the time horizon. To demonstrate that our SDS also works for a single-product case, we present a second case study and again study the influence of the time constant. The second case study is constructed by modifying the first one.

4.1 Setup

A similar setup is used as in the previous case study with a CSTR and three compression chillers (Figure 4). Instead of a multi-product CSTR, we assume a single-product CSTR with a nominal concentration $C_A^{nom} = 0.3\ \frac{mol}{L}$. The CSTR has flexibility because we assume that the concentration is allowed to vary between $C_A^{min} = 0.09\ \frac{mol}{L}$ and $C_A^{max} = 0.51\ \frac{mol}{L}$ as long as the nominal concentration is reached on

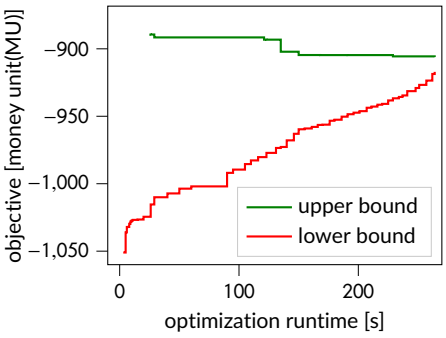

Fig. 11. Convergence plot of *simultaneous dynamic scheduling*.

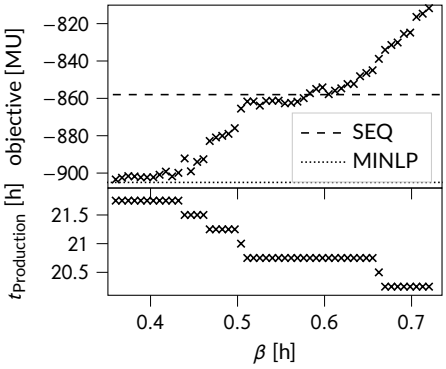

Fig. 12. Objective (in money unit (MU)) and total production time $t_{\mathrm{Production}}$ for different time constants β starting from nominal $\beta = 0.36$ h up to $\beta = 0.72$ h. The objective values resulting from the sequential approach (SEQ) and MINLP optimization with both 21.75 hours of production are shown for comparison (dashed and dotted lines, respectively).

average over the time horizon $t_f - t_0$. Accordingly, the condition

$$\int_{t_0}^{t_f} C_A dt = C_A^{nom}(t_f - t_0) \tag{44}$$

has to hold. Such setups occur when the product can be stored and is well-mixed in the storage tank. Note that an equation of this type would also occur for processes having a variable production rate as controlled variable y_{cv}. Only the concentration C_A in equation 44 would have to be replaced by the production rate.

Obviously, it is favorable to operate at concentrations with a high cooling demand at times of low electricity prices and at concentrations with low cooling demand at times of high prices. The challenge for scheduling optimization thus is to find a trajectory for the concentration that (i) can be realized by the process and (ii) reaches the nominal concentration on average. At the same time, the on/off status of the chillers has to be determined. Note that the scheduling problem has three differences compared to the previous case:

1. Only the cumulative energy costs at final time $\Phi_{Energy}(t_f)$ are considered in the objective (equation 4) since the production is fixed.

2. Equations 5, 26, 27, and 31 - 35 are removed since they are associated with the different products and the production bands.

3. Equation 44 is included such that the nominal concentration is reached on average.

As energy costs are the only objective function in the second case study, a sequential scheduling is not applicable because there is no objective for the process optimization. Thus, we benchmark our SDS against a steady-state operation of the CSTR at the nominal concentration. Again, as a second benchmark, a MINLP optimization is performed using BARON in heuristic mode. A feasible initial point is found by fixing the binary variables to the values from our *simultaneous dynamic scheduling* and solving the resulting NLP.

4.2 Results

The MINLP solution improves the steady-state solution by 6.6 %. Our *simultaneous dynamic scheduling* (SDS) reduces costs by 5.6 % compared to the steady-state benchmark and thus captures 85 % of the MINLP improvement. The optimization runtime of our SDS approach is only 68 s. Note that again the MINLP optimization with BARON does not provide a feasible point without initialization from the SDS solution.

Compared to case study 1, the choice of the time constant β has a much lower impact on the economic result (Figure 13). Note that we use the same time constant β as in case study one, because the transitions studied during tuning cover the complete range of allowed concentrations (compare to section 3.2.1). If β is doubled from 0.36 h to 0.72 h, the cost reduction still amounts to 5.2 % compared to steady-state operation (Figure 13). The operation is very similar for both time constants and the cooling power only deviates significantly in hours 1, 7, and 16-17 (Figure 14). In hours 16 and 17, the schedule with $\beta = 0.72$ h drives the reactor from minimum concentration to maximum concentration. The higher flexibility of the low time constant $\beta = 0.36$ h allows to consume more cooling in hour 16 and less in hour 17 compared to the case where $\beta = 0.72$ h. Still, the larger time constant can capture the main trend of the electricity price profile, which has a peak in the morning and another one in the afternoon. Such a price profile is typical for the German market where the main price periodicities are 24 and 12 hours (Schäfer et al., 2020). Even if β is increased by a factor of 10, the scheduling can still capture the main trend of the electricity price profile (Figure 14). Therefore, significant cost reductions can be reached even if the chosen time constants are far above the optimal value.

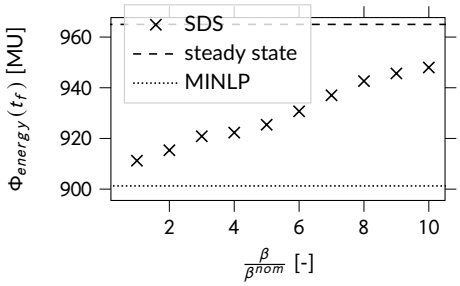

Fig. 13. Energy costs $\Phi_{energy}(t_f)$ (in money unit (MU)) in the second case study achieved with simultaneous dynamic scheduling (SDS) for different time constants β normalized to nominal value $\beta^{nom} = 0.36$ h. The energy costs resulting from steady-state operation and the MINLP benchmark are shown for comparison (dashed and dotted lines, respectively).

5 Conclusion and Discussion

For power-intensive processes, volatile electricity prices provide an opportunity to increase profit via demand response (DR). A particularly promising DR option is the simultaneous scheduling optimization of processes and their energy systems. As such an optimization must consider scheduling-relevant process dynamics as well as on/off-decisions in the energy supply system, computationally challenging nonlinear mixed-integer dynamic optimization (MIDO) problems arise. In this work, we present an efficient *simultaneous dynamic scheduling* (SDS) approach that relies on a tailored scheduling model consisting of (i) a linear scale-bridging model for the closed-loop response of the process, (ii) a data-driven model for

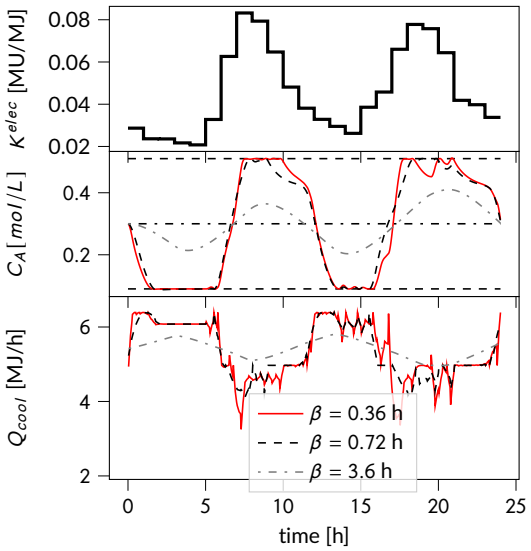

Fig. 14. Electricity price K^{elec}, concentration C_A, and cooling power Q_{cool} in the second case study for three different values of the time constant β.

the process energy demand, and (iii) a mixed-integer linear programming (MILP) model of the energy system. Using a discrete time formulation and collocation, we receive an overall MILP formulation that can be optimized in practically relevant times.

First, we apply the method to a case study of a multi-product continuous stirred tank reactor (CSTR) cooled by three compression chillers. Compared to a typical sequential scheduling, we find that the presented SDS approach improves economic profit by 5.2 %, just shy of the 5.4 % found by nonlinear scheduling optimization using the original nonlinear process model. Second, we investigate a single-product reactor with a variable concentration. Here, SDS outperforms a steady-state operation by 5.6 % while a nonlinear scheduling reaches 6.6 %. In both case studies, the optimization runtime is sufficiently fast for online optimization. As the proposed scheduling model always has the same basic structure, we expect the method to be real-time applicable in many cases.

A restriction of our method is that the scale-bridging approach imposes a single common linear closed-loop response in all operating regimes, which may cut off some of the process flexibility and thus DR potential. For example, in our first case study, we must choose the time constants of the enforced linear closed-loop response such that all six product transitions are feasible. Due to the nonlinear behavior of the CSTR, some of the transitions could in principle be performed faster, however, the critical transition, i.e., the slowest one, limits the time constants for the scale-bridging model. Moreover, it may in general be difficult to find the time constants that give the fastest possible linear closed-loop response. Finding the time constants using the heuristic presented in section 3.2.1 is straightforward if the controlled process can be simulated and the relevant transitions can be studied in numerical experiments. For multi-product processes, which are inherently dynamic, the time constants can also be chosen based on recorded transitions. Our sensitivity study shows that as long as transition times are only moderately larger than necessary, costs can still be reduced compared to a standard sequential scheduling. For processes that are currently operated in steady state without DR, no recorded transitions might be available. However, we demonstrate that, for such processes, time constant choice is less critical as even greatly suboptimal values may allow to follow slow trends in the electricity price profile.

Overall, our results demonstrate that the proposed method offers a favorable trade-off between accurate handling of dynamic flexibility and online applicable optimization run-times.

Author contributions

Florian J. Baader: Conceptualization, Methodology, Software, Investigation, Validation, Visualization, Writing - original draft. **André Bardow**: Funding acquisition, Conceptualization, Supervision, Writing – review & editing. **Manuel Dahmen**: Conceptualization, Supervision, Writing – review & editing.

Declaration of Competing Interest

We have no conflict of interest.

Acknowledgements

This work was supported by the Helmholtz Association under the Joint Initiative "Energy System 2050–A Contribution of the Research Field Energy" and under the Initiative "Energy System Integration".

Nomenclature

Abbreviations

CC	compression chiller
COP	coefficient of performance
CSTR	continuous-stirred-tank-reactor
DR	demand response
MIDO	mixed-integer dynamic optimization
MILP	mixed-integer linear programming
MINLP	mixed-integer nonlinear programming
MU	money unit
SDS	simultaneous dynamic scheduling
SEQ	sequential scheduling benchmark
SO-MPC	scheduling-oriented model predictive control

Greek symbols

β	time constant of natural oscillation
ε	safety margin
ζ	damping coefficient
ϱ	density
τ	time constant
$\overline{\tau}$	scaled time
Φ	objective

Latin symbols

$\mathbf{A} - \mathbf{H}$	matrices
C_A	concentration of component A
c	fitting coefficient
c_P	heat capacity
E_A	activation energy
f, g	functions
ΔH_r	enthalpy of reaction
K	price
k	reaction constant
l	Lagrange polynomial
M	big-M constant
m	linear slope
N_{cp}	order of collocation polynomial
n	natural number
P	power
Q	thermal power
q	flow rate
R	gas constant
r	order of differential equation
s	variable to indicate start of production
T	temperature
t	time
u	manipulated variable
V	volume
w_{SP}	set-point
x	differential state
y	continuous variable
z	discrete variable

Sets

$\mathbb{E}$	end-energy forms
$\mathbb{P}$	products
$\mathbb{S}$	set of possible transitions
$\mathbb{T}_{dis}$	timepoints on discrete grid

Subscripts

0	initial
c	component
cv	controlled variable
$cont$	continuous
$cool$	cooling
dis	discrete
e	end-energy form
ec	energy costs
ed	energy demand
$elec$	electricity
f	final
f_e	finite element
fil	filtered
in	input
on	on-off status
out	output
p	product
s	transition
sum	summed value

Superscripts

end	final value
max	maximum value
min	minimum value
N	nominal
$start$	starting value
$steady$	steady-state value

Bibliography

Agha, M. H., Thery, R., Hetreux, G., Hait, A., and Le Lann, J. M. (2010). Integrated production and utility system approach for optimizing industrial unit operations. *Energy*, 35(2):611–627.

Baader, F. J., Mork, M., Xhonneux, A., Müller, D., Bardow, A., and Dahmen, M. (2020). Mixed-integer dynamic scheduling optimization for demand side management. In Pierucci, S., Manenti, F., Bozzano, G. L., and Manca, D., editors, *30th European Symposium on Computer Aided Process Engineering*, volume 48 of *Computer Aided Chemical Engineering*, pages 1405–1410. Elsevier.

Bahl, B., Lampe, M., Voll, P., and Bardow, A. (2017). Optimization-based identification and quantification of demand-side management potential for distributed energy supply systems. *Energy*, 135:889–899.

Baldea, M., Du, J., Park, J., and Harjunkoski, I. (2015). Integrated production scheduling and model predictive control of continuous processes. *AIChE Journal*, 61(12):4179–4190.

Baldea, M. and Harjunkoski, I. (2014). Integrated production scheduling and process control: A systematic review. *Computers & Chemical Engineering*, 71:377–390.

Beal, L., Petersen, D., Pila, G., Davis, B., Warnick, S., and Hedengren, J. (2017). Economic benefit from progressive integration of scheduling and control for continuous chemical processes. *Processes*, 5(4):84.

Bemporad, A. and Morari, M. (1999). Control of systems integrating logic, dynamics, and constraints. *Automatica*, 35(3):407–427.

Biegler, L. T. (2010). *Nonlinear Programming: Concepts, Algorithms, and Applications to Chemical Processes*. Society for Industrial and Applied Mathematics.

Carrion, M. and Arroyo, J. M. (2006). A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem. *IEEE Transactions on Power Systems*, 21(3):1371–1378.

Caspari, A., Offermanns, C., Schäfer, P., Mhamdi, A., and Mitsos, A. (2019). A flexible air separation process: 2. optimal operation using economic model predictive control. *AIChE Journal*, pages 45–4393.

Castro, P. M., Harjunkoski, I., and Grossmann, I. E. (2009). New continuous-time scheduling formulation for continuous plants under variable electricity cost. *Industrial & Engineering Chemistry Research*, 48(14):6701–6714.

Corriou, J.-P. (2018). *Process control: Theory and applications*. Springer, Cham, Switzerland, second edition edition.

Daoutidis, P. and Kravaris, C. (1992). Dynamic output feedback control of mimimum-phase nonlinear processes. *Chemical Engineering Science*, 47(4):837–849.

Daoutidis, P., Lee, J. H., Harjunkoski, I., Skogestad, S., Baldea, M., and Georgakis, C. (2018). Integrating operations and control: A perspective and roadmap for future research. *Computers & Chemical Engineering*, 115:179–184.

Du, J., Park, J., Harjunkoski, I., and Baldea, M. (2015). A time scale-bridging approach for integrating production scheduling and process control. *Computers & Chemical Engineering*, 79:59–69.

Engell, S. and Harjunkoski, I. (2012). Optimal operation: Scheduling, advanced control and their integration. *Computers & Chemical Engineering*, 47:121–133.

Flores-Tlacuahuac, A. and Grossmann, I. E. (2006). Simultaneous cyclic scheduling and control of a multiproduct cstr. *Industrial & Engineering Chemistry Research*, 45(20):6698–6712.

Flores-Tlacuahuac, A. and Grossmann, I. E. (2010). Simultaneous scheduling and control of multiproduct continuous parallel lines. *Industrial & Engineering Chemistry Research*, 49(17):7909–7921.

Glover, F. (1975). Improved linear integer programming formulations of nonlinear integer problems. *Management Science*, 22(4):455–460.

Gurobi Optimization, LLC (2021). Gurobi optimizer reference manual. http://www.gurobi.com (accessed 09 February 2021).

Harjunkoski, I., Maravelias, C. T., Bongers, P., Castro, P. M., Engell, S., Grossmann, I. E., Hooker, J., Méndez, C., Sand, G., and Wassick, J. (2014). Scope for industrial applications of production scheduling models and solution methods. *Computers & Chemical Engineering*, 62:161–193.

Harjunkoski, I., Nyström, R., and Horch, A. (2009). Integration of scheduling and control—theory or practice? *Computers & Chemical Engineering*, 33(12):1909–1918.

Hart, W. E., Laird, C. D., Watson, J.-P., Woodruff, D. L., Hackebeil, G. A., Nicholson, B. L., and Siirola, J. D. (2017). *Pyomo–Optimization modeling in Python*, volume 67. Springer Science & Business Media, second edition.

Hart, W. E., Watson, J.-P., and Woodruff, D. L. (2011). Pyomo: Modeling and solving mathematical programs in python. *Mathematical Programming Computation*, 3(3):219–260.

Kelley, M. T., Pattison, R. C., Baldick, R., and Baldea, M. (2018a). An efficient milp framework for integrating nonlinear process dynamics and control in optimal production scheduling calculations. *Computers & Chemical Engineering*, 110:35–52.

Kelley, M. T., Pattison, R. C., Baldick, R., and Baldea, M. (2018b). An efficient milp framework for integrating nonlinear process dynamics and control in optimal production scheduling calculations. *Computers & Chemical Engineering*, 110:35–52.

Khajavirad, A. and Sahinidis, N. V. (2018). A hybrid lp/nlp paradigm for global optimization relaxations. *Mathematical Programming Computation*, 10(3):383–421.

Leenders, L., Bahl, B., Hennen, M., and Bardow, A. (2019). Coordinating scheduling of production and utility system using a stackelberg game. *Energy*, 175:1283–1295.

Lewis, J. M., Lakshmivarahan, S., and Dhall, S. (2006). Linear least squares estimation: method of normal equations. In *Dynamic Data Assimilation: A Least Squares Approach*, Encyclopedia of Mathematics and its Applications, pages 99–120. Cambridge University Press.

Merkert, L., Harjunkoski, I., Isaksson, A., Säynevirta, S., Saarela, A., and Sand, G. (2015). Scheduling and energy – industrial challenges and opportunities. *Computers & Chemical Engineering*, 72:183–198.

Mitra, S., Sun, L., and Grossmann, I. E. (2013). Optimal scheduling of industrial combined heat and power plants under time-sensitive electricity prices. *Energy*, 54:194–211.

Mitsos, A., Asprion, N., Floudas, C. A., Bortz, M., Baldea, M., Bonvin, D., Caspari, A., and Schäfer, P. (2018). Challenges in process optimization for new feedstocks and energy sources. *Computers & Chemical Engineering*, 113:209–221.

Neisen, V., Baader, F. J., and Abel, D. (2018). Supervisory model-based control using mixed integer optimization for stationary hybrid fuel cell systems. *IFAC-PapersOnLine*, 51(32):320–325.

Nicholson, B., Siirola, J. D., Watson, J.-P., Zavala, V. M., and Biegler, L. T. (2018). Pyomo.dae: A modeling and automatic discretization framework for optimization with differential and algebraic equations. *Mathematical Programming Computation*, 10(2):187–223.

Otashu, J. I. and Baldea, M. (2019). Demand response-oriented dynamic modeling and operational optimization of membrane-based chlor-alkali plants. *Computers & Chemical Engineering*, 121:396–408.

Pattison, R. C., Touretzky, C. R., Johansson, T., Harjunkoski, I., and Baldea, M. (2016). Optimal process operations in fast-changing electricity markets: Framework for scheduling with low-order dynamic models and an air separation application. *Industrial & Engineering Chemistry Research*, 55(16):4562–4584.

Petersen, D., Beal, L., Prestwich, D., Warnick, S., and Hedengren, J. (2017). Combined noncyclic scheduling and advanced control for continuous chemical processes. *Processes*, 5(4):83.

Risbeck, M. J., Maravelias, C. T., Rawlings, J. B., and Turney, R. D. (2017). A mixed-integer linear programming model for real-time cost optimization of building heating, ventilation, and air conditioning equipment. *Energy and Buildings*, 142:220–235.

Sass, S., Faulwasser, T., Hollermann, D. E., Kappatou, C. D., Sauer, D., Schütz, T., Shu, D. Y., Bardow, A., Gröll, L., Hagenmeyer, V., Müller, D., and Mitsos, A. (2020). Model compendium, data, and optimization benchmarks for sector-coupled energy systems. *Computers & Chemical Engineering*, 135:106760.

Sass, S. and Mitsos, A. (2019). Optimal operation of dynamic (energy) systems: When are quasi-steady models adequate? *Computers & Chemical Engineering*, 124:133–139.

Schäfer, P., Daun, T. M., and Mitsos, A. (2020). Do investments in flexibility enhance sustainability? a simulative study considering the german electricity sector. *AIChE Journal*, 66(11):e17010.

Seborg, D. E., Edgar, T. F., Mellichamp, D. A., and Doyle, F. J. (2010). *Process Dynamics and Control: 3rd Edition*. Wiley.

Tsay, C. and Baldea, M. (2019). 110th anniversary: Using data to bridge the time and length scales of process systems. *Industrial & Engineering Chemistry Research*, 58(36):16696–16708.

Voll, P., Klaffke, C., Hennen, M., and Bardow, A. (2013). Automated superstructure-based synthesis and optimization of distributed energy supply systems. *Energy*, 50:374–388.

Zhang, Q. and Grossmann, I. E. (2016). Planning and scheduling for industrial demand side management: Advances and challenges. In Martín, M. M., editor, *Alternative Energy Sources and Technologies*, Engineering, pages 383–414. Springer, Switzerland.