The Integration of Scheduling and Control: Top-Down vs. Bottom-Up

Adrian Caspari[a,1], Calvin Tsay[b,1], Adel Mhamdi[a], Michael Baldea[b,*], Alexander Mitsos[c,a,d,*]

[a] Process Systems Engineering (AVT.SVT), RWTH Aachen University, 52074 Aachen, Germany

[b] McKetta Dept. of Chemical Engineering, The University of Texas at Austin, Austin, TX 78712, USA

[c] JARA-CSD, 52056 Aachen, Germany

[d] Energy Systems Engineering (IEK-10), Forschungszentrum Jülich, 52425 Jülich, Germany

Abstract: The flexible operation of continuous processes often requires the integration of scheduling and control. This can be achieved by top-down or bottom-up approaches. We compare the two paradigms in-silico using an air separation unit as a benchmark process. To demonstrate the top-down paradigm, we identify data-driven models of the closed-loop process dynamics based on a mechanistic model and use them in scheduling calculations that are performed offline. The resulting target trajectories are passed to a linear model predictive control (LMPC) system and implemented in the process. To demonstrate the bottom-up paradigm, we define an economic nonlinear model predictive control (eNMPC) scheme, which performs dynamic optimization using the full model in closed-loop to directly obtain the control variable profiles to be implemented in the process. We provide implementations of the process model equations as both a gPROMS and a Modelica model to encourage future comparison of approaches for flexible operation, process control, and/or handling disturbances. The performance, advantages, and disadvantages of the two strategies are analyzed using demand-response scenarios with varying levels of fluctuations in electricity prices, as well as considering the cases of known, instantaneous, and completely unknown load changes. The similarities and differences of the two approaches as relevant to flexible operation of continuous processes are discussed. Integrated scheduling and control leverages existing infrastructure, can be immediately applied to real operation tasks. Both operation strategies achieve successful process operation with remarkable economic improvements (up to 8%) compared to constant operation. eNMPC requires more computational resources, and is—at the moment—not implementable in real-time due to maximum optimization times exceeding the controller sampling time. However, eNMPC achieves up to 2.5 times higher operating cost savings compared to the top-down approach, owing in part to the more accurate modeling of key process dynamics.

Keywords: Integrated Scheduling and Control, Economic Model Predictive Control, Air Separation Units, Demand Side Management

* Corresponding authors. *E-mail addresses:* mbaldea@che.utexas.edu (M. Baldea), amitsos@alum.mit.edu (A. Mitsos).
[1] These authors contributed equally to this work

1 Introduction

The flexible operation of continuous processes has potential economic and environmental benefits, e.g., from the implications of variable electricity prices that reflect, among others, the contribution of renewable energy sources to the power generation mix [1]. However, flexible operation is a difficult task, since decisions must be made over multiple time horizons. For example, electricity prices, which are highly relevant to process economics and production scheduling, fluctuate on a time scale of hours (weekly and seasonal variations are also present). On the other hand, process dynamics and control decisions evolve on time scales of seconds, minutes, and even hours. Thus, decisions must consider the time horizon(s) of both production scheduling and process dynamics and control. The flexible operation of continuous processes lies at the core of these interrelated operation tasks, i.e., at the nexus of scheduling and control [2, 3].

Methods for integrated decision-making across several time scales in chemical processes have been discussed in several recent review and perspective articles [2, 4, 5, 6, 7]. Approaches for flexible operation can broadly be categorized into two paradigms [2]: top-down and bottom-up. The former aims to include detailed process information (e.g., process dynamics and control) in calculations in the scheduling layer, while the latter seeks to consider the upper-level objective (e.g., process economics) in the process control layer.

The top-down paradigm takes into consideration the traditional hierarchical separation of operational decisions by relevant time scales [8]. The production schedules resulting from top-down calculations are presented in the form of setpoints/targets to be implemented on the process by a (typically multivariable, advanced) tracking controller, which calculates the profiles of the manipulated variables (MVs) and passes them either to the process directly or via subordinate base-layer (regulatory) controllers. These setpoint/target calculations have been addressed in the past using, e.g., real-time optimization (RTO) [9], and its extension using dynamic process models, dynamic real-time optimization (DRTO) [10]. While classical DRTO approaches, e.g., [10], consider the open-loop dynamics of a process, recent works [11, 12] propose embedding closed-loop process dynamics, or the dynamics of the process subject to lower-level controllers from the perspective of the DRTO layer, in setpoint calculations. This closed-loop DRTO (CL-DRTO) strategy requires an optimization problem formulation or model that is able to represent the closed-loop dynamics of the process, often including model predictive control (MPC) and/or regulatory control.

Jamaludin and Swartz [11] formulated the CL-DRTO problem as a bilevel problem, where the lower-level represents the embedded MPC optimization problem. Although the lower-level optimization problem

might be nonconvex in general, it is substituted by its first-order necessary optimality conditions using complementarity constraints to obtain a single-level problem. Later, Jamaludin and Swartz presented and analyzed multiple approaches to approximate closed-loop behavior to be used in a CL-DRTO formulation, assuming a subordinate linear model predictive control (LMPC) approach for setpoint tracking [12]. They showed that CL-DRTO outperforms DRTO and that the proposed formulations can successfully approximate the closed-loop behavior obtained from the rigorous approach. Depending on the formulation, they could substantially reduce computational requirements, while still retaining optimal closed-loop behavior. Li and Swartz [13, 14] extended the methods for distributed systems.

The same concept of embedding closed-loop process dynamics can be applied to production scheduling calculations, which typically involve longer time horizons than those considered in RTO. Simkoff and Baldea [15] embedded the optimality conditions of the tracking MPC problem (using complementarity constraints) in a dynamic optimization problem over a longer, scheduling time horizon. While [11, 12, 15] use an optimization problem formulation based on a detailed, first-principles model for flexible operation (via RTO or scheduling), several data-driven approaches have been proposed to embed information about process dynamics in production scheduling [5]. For example, Pattison et al. [16, 17, 18] identified data-driven dynamic surrogate models based on process closed-loop data to represent the dynamic behavior of the closed-loop system. The data-driven model representing the closed-loop behavior, termed a scale-bridging model (SBM), was then used in production scheduling calculations. The resulting vector of time-varying operating set-points was passed to a lower-level tracking controller to be implemented on the process.

In contrast to the top-down approaches, the bottom-up paradigm seeks to incorporate economic considerations into the process control system. This can be done either at the level of a subordinated distributed regulatory control system, cf., [19, 20, 21], or at the level of a supervisory controller, e.g., by economic nonlinear model predictive control (eNMPC) [22, 23, 24]. Here, eNMPC is used directly for feedback control and does not imply a hierarchical control structure with a subordinate tracking control scheme as in the above top-down paradigm [24]. Rather, eNMPC integrates the scheduling and control task by considering the scheduling targets at the controller level directly. In that sense, scheduling is an extension of the original eNMPC idea. Specifically, a dynamic optimization problem with an economic objective is solved with a sufficiently long time horizon, subject to an open-loop process model or the model including base-layer/regulatory controllers. Following this paradigm, a subordinated advanced tracking controller is not required; rather, the eNMPC is responsible for both maximizing economic performance and directly controlling the process. Recent works focused both on the theoretical properties, e.g.,

[25, 26], and on applications [27, 28, 29, 30] of eNMPC. Since eNMPC is computationally demanding, its application to large-scale processes requires numerically efficient approaches, such as fast-update methods [31, 32, 28, 29], hierarchical control structures [33], and/or model reduction techniques [34, 35, 36].

Both paradigms described above seek to integrate scheduling and control, but the question of whether a top-down or a bottom-up approach is more suitable for a specific process operation task has not been fully explored in the open literature [2]. Therefore, in this work, we compare two approaches representative of the top-down and bottom-up paradigms: i) scheduling with data-driven, closed-loop dynamic models (SBMs [16]) and subordinate tracking LMPC, and ii) economic nonlinear model predictive control (eNMPC). The two strategies, illustrated in Fig. 1, are applied in silico to a prototype air separation unit (ASU). Specifically, we employ the ASU model by Johansson [37] as a benchmark process for flexible process operation, and consider several representative demand-response scenarios.

Along the lines of [2], we propose the ASU model as a benchmark for the study of the integration of scheduling and control. ASUs are interesting case studies for the application of operating schemes and have been considered in numerous publications [38, 28, 27, 29, 30, 17]. They are electricity-intensive processes that accounted for 2.2% of the electricity consumption of the manufacturing industry in the US in 2014 [39]. In addition, ASUs are large-scale, multiple-input-multiple-output (MIMO) processes that include several standard process units, i.e., distillation columns, (multi-stream) heat exchangers, compressors, and turbines. They feature tight material and energy integration and have complex nonlinear dynamics that span multiple time scales. Furthermore, the process includes only physical effects that are well understood and can be represented using standard models.

To encourage comparisons of future methods using the proposed benchmark process, the model is implemented identically in both gPROMS [40] and Modelica [41]. Both models are made openly available online [42] at `https://data.mendeley.com/datasets/pfcc5gvzty` and can be used, e.g., to test methods for flexible operation, including scheduling techniques, (hierarchical) control systems, and methods for identifying and handling disturbances.

The remainder of the work is structured as follows: the two process operating strategies are first explained in Section 2. We then describe the benchmark ASU process and model implementations in Section 3 and show the comparison results in Section 4. Finally, Section 5 provides some concluding remarks.

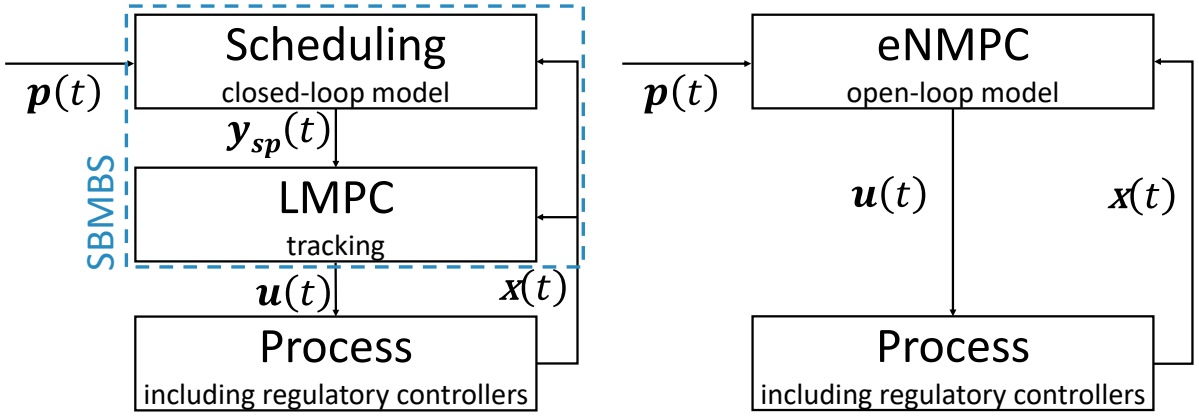


Fig. 1: Operating schemes for integrated scheduling and control. (a) top-down: SBM-based scheduling (SBMBS). A data-driven model (SBM) is used to predict the dynamic closed-loop behavior of the process under subordinate LMPC (and regulatory) control. The setpoints are passed to the LMPC which calculates the profiles of the manipulated variables based on a linear process model. The manipulated variable profiles are then implemented in the process directly or as setpoints for regulatory (e.g., PID) control. (b) bottom-up: eNMPC. The eNMPC solves a nonlinear dynamic optimization problem subject to a dynamic open-loop process model and other operational constraints. The resulting MV profiles are implemented in the process directly or as setpoints for regulatory control.

2 Methods

We describe the top-down approach considered in Section 2.1 and the bottom-up approach in Section 2.2.

2.1 Top-Down: SBMBS

The top-down (i.e., integrated scheduling and control) problem embeds dynamic information about the process and its control system into the production scheduling problem, in order to obtain dynamically feasible schedules [6]. The objective of the integrated problem is to determine the production setpoints to the process control system that maximize economic performance. A lower-level control system, e.g., LMPC, determines the optimal values of the manipulated variables to best guide the process to the

scheduled set points. The deterministic integrated scheduling and control problem can be stated as:

$$
\begin{aligned}
&\min_{\boldsymbol{y}_{sp},\boldsymbol{x},\boldsymbol{y},\boldsymbol{u},\boldsymbol{u^c}} \Phi(\boldsymbol{x}(t_{\text{sched}})) \\
s.t. \quad \boldsymbol{M}\dot{\boldsymbol{x}}(t) &= \boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{y}(t),\boldsymbol{u}(t),\boldsymbol{p}(t)) \ \forall t \in \mathcal{T}^s \\
\boldsymbol{0} &= \boldsymbol{g}(\boldsymbol{x}(t),\boldsymbol{y}(t),\boldsymbol{u}(t),\boldsymbol{p}(t)) \ \forall t \in \mathcal{T}^s \\
\boldsymbol{u}(t) &= \tilde{\boldsymbol{u}}_i, t \in [t_{i-1}, t_i) \\
\tilde{\boldsymbol{u}}_i &= \boldsymbol{\psi}(\tilde{\boldsymbol{x}}_{i-1}, \tilde{\boldsymbol{y}}_{sp,i-1}) \\
\boldsymbol{x}(t_0) &= \tilde{\boldsymbol{x}}_0 \\
\boldsymbol{0} &\geqslant \boldsymbol{c}(\boldsymbol{x}(t),\boldsymbol{y}(t),\boldsymbol{u}(t),\boldsymbol{p}(t)), t \in \mathcal{T}^s
\end{aligned}
\tag{1}
$$

where $\boldsymbol{x} : \mathcal{T}^s \to \mathbb{R}^{N_x}$ are the differential states, $\boldsymbol{y} : \mathcal{T}^s \to \mathbb{R}^{N_y}$ are the algebraic states, $\boldsymbol{u} : \mathcal{T}^s \to \mathbb{R}^{N_u}$ are the MVs, $\boldsymbol{p} : \mathcal{T}^s \to \mathbb{R}^{N_p}$ are the parameters, $\boldsymbol{f} : \mathcal{X} \to \mathbb{R}^{N_x}$, $\boldsymbol{g} : \mathcal{X} \to \mathbb{R}^{N_y}$ define the open-loop process model, given as a semi-explicit differential-algebraic equation system (DAE) with differential index 1, $\mathcal{X} = \mathbb{R}^{N_x} \times \mathbb{R}^{N_y} \times \mathbb{R}^{N_u} \times \mathbb{R}^{N_p}$, $\mathcal{T}^s = [t_0, t_f]$, and $\boldsymbol{c} : \mathcal{X} \to \mathbb{R}^{N_g}$ are path and terminal constraints. $\boldsymbol{M} \in \mathbb{R}^{N_x \times N_x}$ is the nonsingular mass matrix, $\mathbf{y}_{sp}(t)$ is the trajectory of set points and/or production targets, and Φ is an economic objective function. The problem is solved over a scheduling time horizon, $t_{sched} = t_f - t_0$, which is typically much longer than the dominant time constant of the process.

In effect, the optimization problem (1) determines the trajectory of $\mathbf{y}_{sp}(t)$ that minimizes Φ, while the evolution of the process states, $\boldsymbol{x}$ and $\boldsymbol{y}$, is predicted by the process model given values for the MVs and parameters, respectively, $\boldsymbol{u}$ and $\boldsymbol{p}$. The MVs $\boldsymbol{u}$ are discretized using constant-length time slots, and their values in time slot i are denoted as $\tilde{\boldsymbol{u}}_i, \forall\ t \in [t_{i-1}, t_i)$. $\tilde{\boldsymbol{x}}_i$ and $\tilde{\boldsymbol{y}}_{sp,i}$ similarly denote the values of $\boldsymbol{x}$ and $\boldsymbol{y}_{sp}$ in each time slot i. $\tilde{\boldsymbol{u}}_i$ are computed from the control law $\boldsymbol{\psi}(\circ)$, written in (1) as an explicit control law for simplicity. However, in this work we consider an (optimization-based) lower-level LMPC problem, which takes the following form:

$$
\begin{aligned}
&\min_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}} \frac{1}{2} \sum_{k=i}^{i+N-1} \left(\left|\tilde{\boldsymbol{y}}_k - \tilde{\boldsymbol{y}}_{sp,k}\right|_Q^2 + \left|\tilde{\boldsymbol{u}}_k - \tilde{\boldsymbol{u}}_{sp,k}\right|_R^2 \right) + \left|\tilde{\boldsymbol{y}}_N - \tilde{\boldsymbol{y}}_{sp,N}\right|_{P_f}^2 \\
s.t. \quad \tilde{\boldsymbol{x}}_{k+1} &= \boldsymbol{A}\tilde{\boldsymbol{x}}_k + \boldsymbol{B}\tilde{\boldsymbol{u}}_k \ \forall k \in \{i, ..., i+N-1\} \\
\tilde{\boldsymbol{y}}_k &= \boldsymbol{C}\tilde{\boldsymbol{x}}_k \ \forall k \in \{i, ..., i+N-1\} \\
\tilde{\boldsymbol{x}}_0 &= \boldsymbol{x}_{initial} \\
\boldsymbol{y}^{LB} &\leqslant \tilde{\boldsymbol{y}}_k \leqslant \boldsymbol{y}^{UB} \ \forall k \in \{i, ..., i+N-1\} \\
\boldsymbol{u}^{LB} &\leqslant \tilde{\boldsymbol{u}}_k \leqslant \boldsymbol{u}^{UB} \ \forall k \in \{i, ..., i+N-1\}
\end{aligned}
\tag{2}
$$

where $\boldsymbol{A} \in \mathbb{R}^{N_x \times N_x}$ is the state transition matrix, $\boldsymbol{B} \in \mathbb{R}^{N_x \times N_u}$ is the input matrix, and $\boldsymbol{C} \in \mathbb{R}^{N_y \times N_x}$ is the output matrix. The values in the matrices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$ are the parameters of the linear state-space model and can be identified from process data (usually through open-loop system identification experiments), or from simulations of an available dynamic model. The LMPC problem (2) determines the values of the MVs, $\boldsymbol{u}$, that minimize deviations of $\boldsymbol{y}$ from setpoints $\boldsymbol{y}_{sp}$, while also penalizing changes in $\boldsymbol{u}$. The process behavior is predicted by a simpler, linear model compared to $\boldsymbol{f}$ and $\boldsymbol{g}$ in (1). The LMPC problem is discretized in time, with k denoting the sample number ($k \in \{i, .., i+N-1\}$) at time $t = kT_s$. T_s is the sample time, N is the prediction horizon, and Q and R are tuning parameters penalizing deviations from the setpoint and control moves, respectively. The input variables are parameterized over $N-1$ time steps, giving the vector of control moves $\{\tilde{\boldsymbol{u}}_i, ..., \tilde{\boldsymbol{u}}_{i+N-1}\}$.

Optimization of (1) with (2) embedded involves passing values of $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{y}_{sp}$ from the upper-level problem (1) to the lower-level LMPC problem (2), solving for $\boldsymbol{u}$, and passing the optimal values of $\boldsymbol{u}$ back to (1). This information exchange repeats in a moving horizon fashion to enact closed-loop control. Therefore, solving the integrated problem over a long scheduling horizon, t_{sched}, requires significant computational effort, making this formulation intractable for many practical applications. Therefore, in this work, we accelerate solution of the integrated scheduling and control problem, using a reduced-order SBM to approximate the closed-loop response of the process to changes in $\boldsymbol{y}_{sp}(t)$ [43]. Following this approach, the dynamic model of the process and its control system is replaced with an input-output model relating the process output variables to changes in $\boldsymbol{y}_{sp}$. The SBM-based scheduling problem (SBMBS) can be written:

$$
\begin{aligned}
&\min_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{y}_{sp}} \Phi(\boldsymbol{x}(t_{sched})) \\
s.t. \quad &\boldsymbol{M}\dot{\boldsymbol{x}}(t) = \boldsymbol{f}_{SBM}(\boldsymbol{x}(t), \boldsymbol{y}(t), \boldsymbol{y}_{sp}(t), \boldsymbol{p}(t)) \\
&\boldsymbol{0} = \boldsymbol{g}_{SBM}(\boldsymbol{x}(t), \boldsymbol{y}(t), \boldsymbol{y}_{sp}(t), \boldsymbol{p}(t)) \\
&\boldsymbol{x}(t_0) = \boldsymbol{x_0} \\
&\boldsymbol{0} \geqslant \boldsymbol{c}(\boldsymbol{x}(t), \boldsymbol{y}(t), \boldsymbol{y}_{sp}(t), \boldsymbol{p}(t)), t \in \mathcal{T}^s
\end{aligned}
\tag{3}
$$

where $\boldsymbol{f}_{SBM} : \mathbb{R}^{N_x} \times \mathbb{R}^{N_{sp}} \rightarrow \mathbb{R}^{N_x}$ and $\boldsymbol{g}_{SBM} : \mathcal{X} \rightarrow \mathbb{R}^{N_y}$ denote the SBM. Similar to the original bi-level optimization problem (1), the SBM-based optimization problem (3) determines the trajectory of $\boldsymbol{y}_{sp}$ that minimizes Φ. However, the evolution of the process is predicted by a closed-loop process model, and there is no embedded, lower-level optimization problem $\tilde{\boldsymbol{u}}_i = \boldsymbol{\psi}(\circ)$. Rather, the SBM, directly predicts the closed-loop process behavior given $\boldsymbol{y}_{sp}$.

The SBM can be derived using model reduction principles from a detailed dynamic model [44, 43].

Alternatively, the SBM can be constructed via a data-driven approach, comprising system identification using a set of dedicated experiments, historical operating data with routine setpoint changes (as in [45]), or data generated by simulation of a detailed dynamic model [16]. The variables modeled in the SBM need only include the subset of process variables that are scheduling relevant, i.e., they are involved in Φ and/or represent important constraints in $\boldsymbol{c}(\circ)$, such as quality or safety-critical constraints. While the selection of scheduling-relevant variables can be carried out using empirical process insights, Tsay and Baldea [46] showed that a latent-variable approach could be used to reduce the dimensionality of dynamical system $\boldsymbol{f}_{SBM}$ included in SBMBS.

2.2 Bottom-Up Approach: eNMPC

In contrast to the top-down approach, where knowledge of the control system is embedded in scheduling calculations, an economic objective function—which might be identical to the scheduling objective function in (1)—can be employed at the control level directly. In the eNMPC approach, the dynamic optimization problem solved at the control level takes the process operating cost (or profit) explicitly into account. We consider the formulation of an optimization problem subject to a semi-explicit index-1 differential-algebraic equation system (DAE) and additional constraints:

$$\begin{aligned} &\min_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{u}} \Phi(\boldsymbol{x}(t_f)) \\ \text{s.t.} \quad &\boldsymbol{M}\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{y}(t),\boldsymbol{u}(t),\boldsymbol{p}(t)), \forall t \in \mathcal{T}^{\text{eNMPC}} \\ &\boldsymbol{0} = \boldsymbol{g}(\boldsymbol{x}(t),\boldsymbol{y}(t),\boldsymbol{u}(t),\boldsymbol{p}(t)), \forall t \in \mathcal{T}^{\text{eNMPC}} \\ &\boldsymbol{x}(t_0) = \boldsymbol{x_0} \\ &\boldsymbol{0} \geqslant \boldsymbol{c}(\boldsymbol{x}(t),\boldsymbol{y}(t),\boldsymbol{u}(t),\boldsymbol{p}(t)), t \in \mathcal{T}^{\text{eNMPC}} \end{aligned} \tag{4}$$

where $\mathcal{T} = [t_0, t_f']$, and $\boldsymbol{c} : \mathbb{R}^{N_x} \times \mathbb{R}^{N_y} \times \mathbb{R}^{N_u} \times \mathbb{R}^{N_p} \rightarrow \mathbb{R}^{N_g}$ are path and terminal constraints. $\boldsymbol{f}$ and $\boldsymbol{g}$ define the open-loop process model as in (1). t_0 and t_f' are the initial and final time, respectively. $\Phi : \mathbb{R}^{N_x} \rightarrow \mathbb{R}$ is the economic objective function. The ENMPC problem (4) determines the trajectory of the MVs $\boldsymbol{u}(t)$ that minimizes Φ, while the evolution of the process states, $\boldsymbol{x}$ and $\boldsymbol{y}$, is predicted by the open-loop process model in $\boldsymbol{f}$ and $\boldsymbol{g}$.

In contrast to the top-down approach, the DAE system describing the process is an open-loop process model, i.e., it does not include a lower level tracking MPC. However, it may include subordinated regulatory controllers (typically PID). The optimization problem (4) is solved on a moving horizon with a sampling time t_s and a control and prediction horizon $t_c = t_f' - t_0$. The resulting optimal control variable

profiles are implemented on the process directly or as setpoints to the subordinate regulatory controllers. We note here that the time interval $t_c = t'_f - t_0$ is typically shorter than $t_{sched} = t_f - t_0$ considered in Problem (1).

The constraint sets of the dynamic optimization problems (1), (2), (3), and (4) include DAE systems Several software packages, e.g., Dymola [47] and gPROMS [40], automatically reduce the DAE to a system of ordinary differential equations (ODEs) by performing index reduction. Nevertheless, designating constraints as DAEs is more general and convenient, and the methods presented in this work are not restricted to ODEs.

3 Air Separation Process

The ASU shown in Fig. 2 and corresponding process model described by Johansson [37] is considered as a benchmark process to compare flexible operation approaches. This section describes the process model of the benchmark ASU and a set of practically motivated demand-response scenarios.

3.1 Process and Model Description

The process uses a single cryogenic distillation column to produce a stream of high-purity nitrogen product. The inlet stream is first compressed in the main compressor (MC) to 6.8 bar and is cooled against warming gaseous products in a multi-stream primary heat exchanger (PHX). For modeling purposes, the PHX is divided into two sections (PHX1 and PHX2). At the end of PHX1, a portion $\xi_{turbine}$ of the feed stream is removed and passed through a turbine to generate some electricity. The rest of the feed stream passes through PHX2, where it is liquefied. The two portions of the feed stream are re-combined and fed to the bottom of the high-pressure distillation column (HPC). The bottoms product of the column is expanded to atmospheric pressure and passes through the reboiler side of an integrated reboiler/condenser (IRC) unit. The distillate of the column consists of high-purity nitrogen, and a portion ξ_{top} is removed as a product stream and sent to the PHX. The remainder of the distillate is sent to the condenser side of the IRC and returned to the column as reflux. The product stream is removed at the midpoint of PHX1, expanded across a second turbine, and passed through the entire PHX again. The waste stream from the reboiler also passes through the PHX to provide cooling for the feed air.

To enable load shifting, the process flowsheet includes a nitrogen liquefier, a storage tank, and an evaporator. Stored liquid nitrogen from the product tank can be evaporated and used to meet instantaneous gas nitrogen demand when production rate is decreased. On the other hand, excess product

nitrogen (beyond instantaneous demand) can be liquefied and stored in the storage tank, allowing liquid nitrogen inventory to be replenished, typically during periods of low electricity prices.

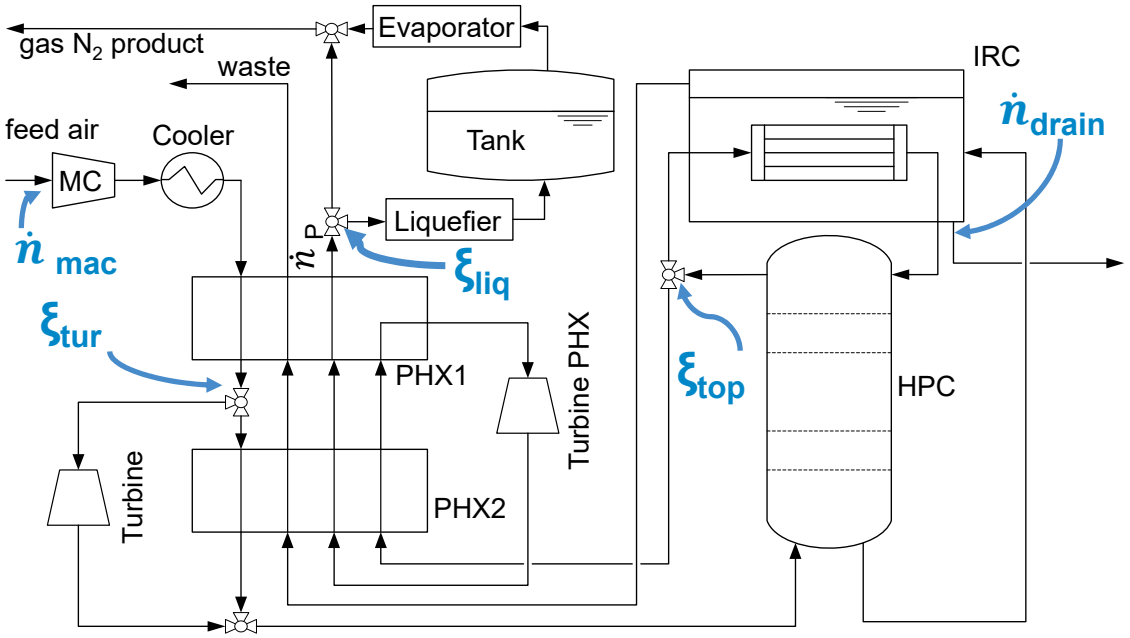


Fig. 2: Air separation process flowsheet. MVs: $\dot{n}_{\text{mac}}$ feed air flowrate, ξ_{tur} split factor to turbine, ξ_{top} split factor to PHX2, ξ_{liq} split factor to liquefier, $\dot{n}_{\text{mac}}$ feed air flowrate.

To encourage the comparison of methods in addition to those considered herein, we provide implementations of the process model in both gPROMS [40] and Modelica [41], and the model files are available at [42]. The main parts of the process model are summarized below. The interested reader is referred to [37] for further details.

Physical Properties: Air is modeled as a ternary mixture of nitrogen, oxygen, and argon. The vapor phase is modeled as an ideal gas, while the Margules equation [48] is used to compute thermodynamic properties for the liquid phase. Liquid Densities are computed using the Rackett equation [49]. Standard correlations are used for enthalpy and entropy calculations, and the Antoine equation [49] is used to compute saturation pressures.

Turbines/compressor: The dynamics of the turbines and the compressor are assumed to be fast, and thus these units are modeled using steady-state equations. We further assume that the compression and expansion can be modeled as polytropic processes. The turbines and compressor are modeled with a constant adiabatic efficiency of 0.8 using polytropic equations. The turbines/compressor are assumed to operate within the surge or choke lines of the compressor map. Previous work [50] suggests that this

simplification of the compressor model is justified when the ASU is operated within a relatively narrow operation range. Nevertheless, the models can easily be adapted to include a compressor map and additional constraints for further investigations.

Distillation column: The distillation column is modeled as stage-by-stage column with 30 equilibrium stages. Index-1 reduction is performed as in [51]. The column is assumed to be well-insulated (no heat loss), with well-mixed stages. The condenser operates at 6.4 bar, and we assume a 0.2 bar total pressure drop across the column with a linear pressure profile over the column stages.

Multi-stream heat exchangers: The PHX is modeled using a one-dimensional spatial domain. Dynamic mass and energy balance equations are used for the wall, while quasi-stationary balance equations are used for the fluid. Most of the heat exchange occurs in PHX1, which is discretized into 50 finite elements. The smaller region, PHX2 is treated as a single element.

Liquefier/evaporator: It is assumed that the dynamics of the liquefier are much faster than those of the ASU, and the liquefier is modeled using steady-state material and energy balances. The liquefier is represented as an ideal refrigeration cycle with overall efficiency of 0.8. The evaporator operates at ambient conditions and is assumed to not require further energy input.

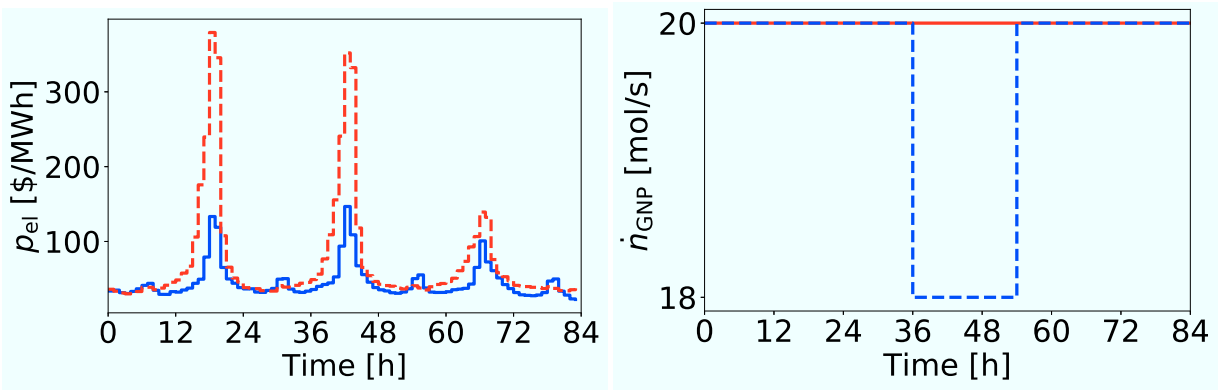
In total, the ASU process model is an index-1 DAE system with 415 differential states and ∼2400 algebraic states (the gPROMS implementation contains a few additional algebraic states as tracking variables). The model does not include any subordinate PID controller, and we employ the LMPC described by [52]. The MV setpoints computed by the LMPC are passed directly to the process.

The LMPC has four controlled variables (CVs): $\dot{n}_{product}$, $I_{product}$, ΔT_{RC}, and n_R. The first CV, $\dot{n}_{product}$ represents the amount of product produced by the process, and the corresponding value is time-varying and determined by the scheduling layer. The process is designed for a nominal production rate of $\dot{n}_{product} = 20$ mol/s, but can modulate its production as long as operational constraints are met. The remaining CVs are involved in important constraints that must be satisfied. $I_{product}$ is the level of impurity (oxygen and argon) in the product nitrogen stream, which must not exceed a value of 1500ppm at any time. ΔT_{RC} is the temperature difference between the reboiler and condenser side of the IRC, which must not fall below a value of 2K at any time. Finally, n_R is the holdup in the reboiler, which cannot be depleted or exceed the physical capacity of the reboiler. These operational constraints are summarized in Table 2. The split fraction to the liquefier ξ_{liq} is set such that the product demand is exactly met at all times, i.e., $\dot{n}_{product}(1-\xi_{liq}) = \dot{n}_{demand}$ when $\dot{n}_{product} > \dot{n}_{demand}$, and $\xi_{liq} = 0$ otherwise.

4 Closed-Loop Case Studies

In this section, specifics regarding the closed-loop simulations employing the top-down and bottom-up approaches are first provided. Afterwards, we present and discuss the results found for optimal flexible operation in the various scenarios.

4.1 Operation Scenarios and Settings



(a) (b)

Fig. 3: (a) Electricity price profiles. Blue, solid line: moderate electricity price profile. Red, dashed line: extreme electricity price profile. (b) Product demand profile. Red, solid line: constant load. Blue, dashed line: temporary load change.

Table 1: MVs and their bounds.

MV name	lower bound	upper bound
$\dot{n}_{MAC}$ [mol/s]	30	50
ξ_{tur} [-]	0	0.1
ξ_{liq} [-]	0	1
ξ_{HPC} [-]	0.51	0.54
F_{drain} [mol/s]	0	2

Table 2: Summary of operational constraints.

variable name	lower bound	upper bound	constraint type
$I_{product}$ [ppm]	0	1500	path
ΔT_{RC} [K]	2	5	path
N_{R} [mol]	5000	40000	path
N_{tank} [mol]	864000	3456000	path
F_{tank} [mol/s]	0	-	path
N_{R} [mol]	20000 = initial	-	terminal
N_{tank} [mol]	1728000 = initial	-	terminal

Table 3: Overview operational scenarios.

scenario	electricity price fluctuations	load change
S1	moderate	no
S2	moderate	yes
S3	extreme	no
S4	extreme	yes

Four benchmark demand-response scenarios that exhibit representative, practical challenges for flexible operation are presented here. These scenarios include two different electricity price profiles, with the first exhibiting moderate price fluctuations and the second extreme fluctuations. Both are shown in Fig. 3a. The electricity price profiles were retrieved from the California Independent System Operator (CAISO), corresponding to the day-ahead electricity prices from October 16-19, 2017 (moderate fluctuations), and August 1-4, 2017 (extreme fluctuations). The electricity price profiles are taken from historical data, and our ratings (moderate vs. extreme) are subjective and relative to current electricity pricing situations. Although we consider a scheduling horizon of three days, prices are provided for four days for the moving horizon approaches.

Accounting for moderate and extreme fluctuations in electricity price, as well as with and without the occurrence of the unplanned maintenance event, the four scenarios used to compare SBMBS and eNMPC in this work are summarized in Table 3.

In addition to the dynamics introduced by the electricity market, we consider the case of a change in the product demand. As the product demand has to be exactly satisfied at all times, the product is withdrawn from the storage tank when the actual product rate of the process is below the demand.

Excess production is directed to the storage tank. The product demand profiles are illustrated with and without the load change disturbance in Fig. 3b. Specifically, this comprises a schedule-level disturbance, where the customer demand decreases by 10% (from the nominal 20mol/s to 18mol/s) for a period of time between 36 and 54 hours. This disturbance is assumed to be caused by an unplanned maintenance event downstream of the ASU. In this scenario, the occurrence of the change is unknown before it takes place, but once the change begins, the new demand and duration are known (i.e., the length of time required for the maintenance procedure is known). Other alternatives could be to consider both the start and end times, as well as the magnitude of the demand change, to be known (planned maintenance), or to all be unknown (random failure). The results for a random failure situation, where the occurrence of the temporary demand change is completely unknown, are given in the Supplementary Material.

The objective of both approaches is to minimize the operating cost of the ASU, given by:

$$\Phi = \int_{t_0}^{T} p_{\text{elec}}(t) \left(P_{\text{comp}}(t) + P_{\text{liq}}(t) - P_{\text{tur}}(t)\right) \mathrm{d}t \tag{5}$$

where p_{elec} is the electricity price per MWh, P_{comp} the compressor power demand, P_{liq} the liquefier power demand, P_{tur} the power supplied by the two turbines, and the final time T is t_f for the SBMBS and t'_f for the eNMPC.

The MVs and their bounds are summarized in Table 1. Different initial points for the MVs are used for the two methods: the initial point for the SBM schedule corresponds to the steady-state found at the nominal MPC setpoint values, as the controller tends to return the process to these conditions. On the other hand, the initial point for the first eNMPC iteration corresponds to the optimal time-invariant production. The process is initialized by simulating the mathematical model with the relevant MVs until steady-state is reached. The path and terminal constraints given in Table 2 are included in the optimization problems.

The nominal values in Table 1 correspond to the optimal time-invariant operation with respect to the process power demand and subject to the constraints summarized in Table 2. This optimal time-invariant operation provides a base case for economic performance. This corresponds to conventional optimal operation without responding to (predicted) changes in electricity prices All solution times reported were obtained on a 64-bit Windows 10 desktop computer with an Intel Core i7-8700 CPU at 3.20 GHz and 16GB RAM.

4.1.1 SBMBS settings

SBMBS is performed using the closed-loop process and LMPC, with linear state-space models developed for the considered ASU by Dias et al. [52]. The LMPC is implemented in the Matlab Model Predictive Control Toolbox [53] with a sampling time of six mininutes and a prediction horizon of 10 hours. The LMPC problem is solved almost instantaneously (<0.1 s) and could potentially also be implemented with a shorter sampling time. Closed-loop simulations are performed by simulating the full-order gPROMS dynamic model and passing the outputs to the controller implementation in Matlab (i.e., full state feedback). The load change is handled in SBMBS following the approach of Pattison et al. [18], wherein the production scheduling horizon is resolved for the remainder of the three-day horizon, and a high-gain observer updates the states of the SBM.

A data-driven approach [16] is used to identify an SBM for the process and its LMPC: first, the optimal schedule found using a steady-state model is simulated to create a data set for closed-loop system identification. The identified Hammerstein-Wiener models provided by Pattison et al. [16] (for a simpler controller rather than LMPC) were then updated by using the previous parameters as an initial guess in the Matlab System Identification Toolbox [54]. The structure of the models remains unchanged. The performance of the updated models on training and test data is given in Table 4. The identified SBM predicts the process response to changes in the production rate setpoint, whose trajectory is then optimized in SBMBS calculations. Scheduling calculations were performed using the built-in dynamic optimization capabilities of gPROMS, with an hourly discretization of the production setpoint that matches the hourly electricity price intervals. The production rate setpoint/target was constrained between 15 and 25 mol/s during each hourly interval. The default tolerances for DASOLV and the SQP optimizer in gPROMS are used.

Table 4: Identified Model Details.

variable	# P-L input segments	linear system order	output model type	output poly. order	# P-L output segments	NMSE (training)	NMSE (validation)
$\dot{n}_{product}$	5	3	polynomial	2		0.93	0.90
$I_{product}$	4	4	PW linear		6	0.82	0.65
$\dot{n}_{MAC}$	3	2	polynomial	2		0.89	0.83
N_R	3	4	polynomial	1		0.79	0.81
ΔT_{RC}	9	4	polynomial	2		0.82	0.80

4.1.2 eNMPC settings

The eNMPC utilizes the full-order, mechanistic process model and is implemented with a sampling time of 15 min and a control horizon of 12 hours. We use equal prediction and control horizons and again assume full state feedback. The MVs are discretized using a piecewise-constant parameterization with uniform intervals of 15 min. In contrast to the optimization problems of the SBMBS, the eNMPC directly utilizes all manipulated variables and constraints in Table 2. A warm-start strategy is used, wherein each eNMPC optimization problem is initialized with the MV profiles resulting from the solution of the previous eNMPC problem. The time delay resulting from the solution of the dynamic optimization problem is neglected in this work, providing insight into the optimal performance and MV trajectories with eNMPC.

The eNMPC handles the load change when it occurs, since the beginning of the load change coincides with one of the online dynamic optimization problems. The eNMPC problems (4) are solved to local NLP convergence with direct single-shooting [55, 56] using the DyOS (Dynamic Optimization Software) framework [57]. The dynamic optimization problem is solved sequentially, using NIXE (NIXE Is eXtrapolated Euler) [58] as DAE integrator and SNOPT (Sparse Nonlinear OPTimizer) [59] as NLP solver. The model is implemented in Modelica [41] using Dymola [47] and coupled to DyOS as a Functional Mockup Unit (FMU) [60], generated with Dymola. FMU only supports ODEs, and Dymola performs symbolic reformulation and numerical reduction of the DAE system to provide an FMU. We employ settings of 10^{-6} as DAE integrator tolerance and 10^{-4} as NLP feasibility and optimality tolerances.

Remark 1 gPROMS and Modelica use different numerical solvers (e.g., DASOLV and NIXE, respectively) for time integration. However, we find that the two implementations give near-identical results. For the solutions to the scenarios reported in this work, the average relative difference between the predictions of the output variables between the two models is 5.6×10^{-6}. The maximum relative difference is 4.2×10^{-5}. Therefore, while the differences between schedules arise primarily from the different strategies used, there may be small differences caused by the different numerical methods. For consistency, all results presented in this work are given as the values generated by simulating the gPROMS implementation for the given MV profiles.

4.2 Numerical Results

Solution times: The solution of the SBMBS problems required 2229s and 2258s of CPU time for scenarios S1 and S3, respectively. While this solution time is significant, the scheduling problem is solved on a slower frequency (e.g., daily or every three-day scheduling period) compared to the online eNMPC

Table 5: Summary of economic performance of SBMBS and eNMPC for different control scenarios. Base represents the constant production rate case.

	Base	SBMBS	Δ	eNMPC	Δ
S1	\$1025.09	\$1010.18	-1.5%	\$991.17	-3.3%
S2	\$997.76	\$988.12	-1.0%	\$966.15	-3.2%
S3	\$1700.86	\$1585.86	-6.8%	\$1561.71	-8.2%
S4	\$1645.21	\$1556.36	-5.4%	\$1527.62	-7.2%

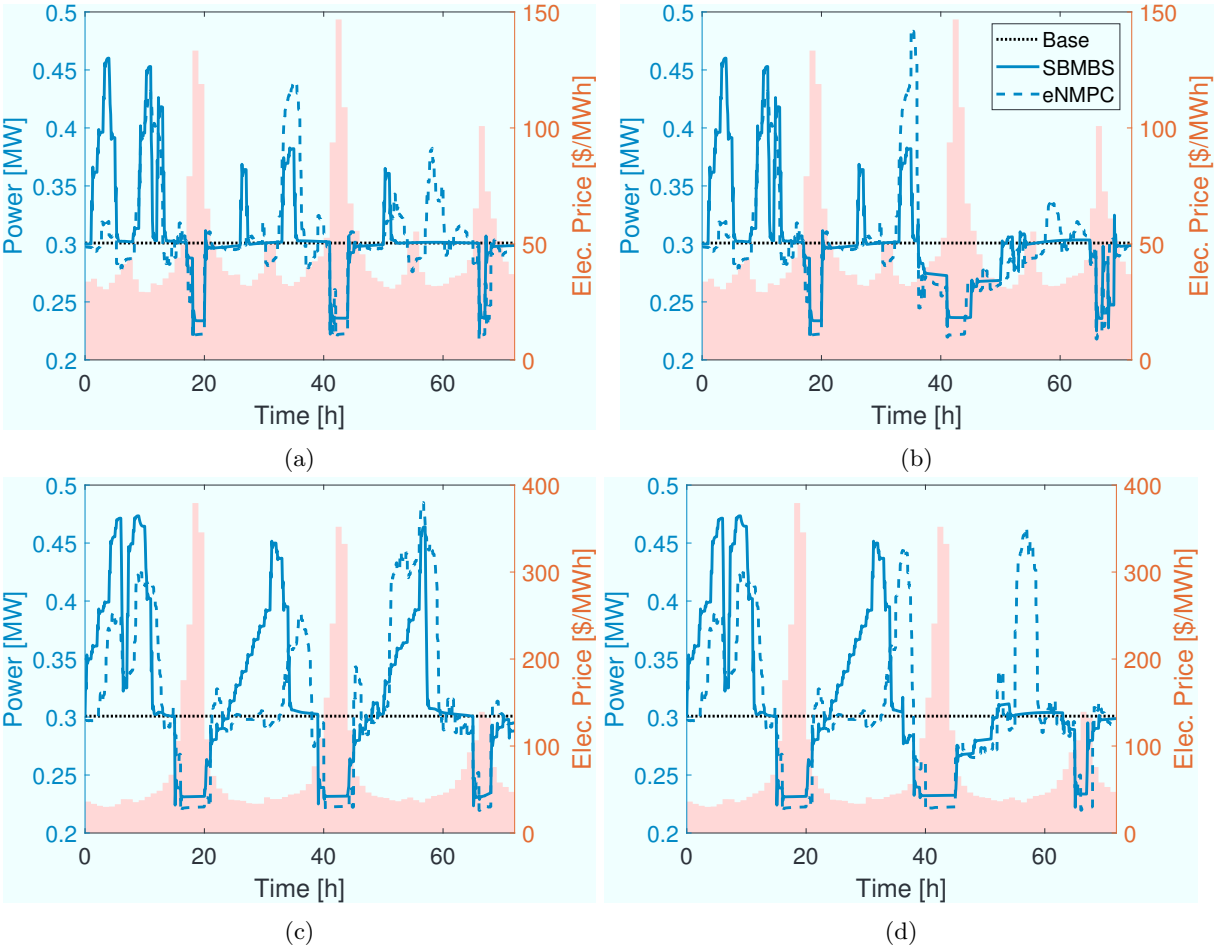
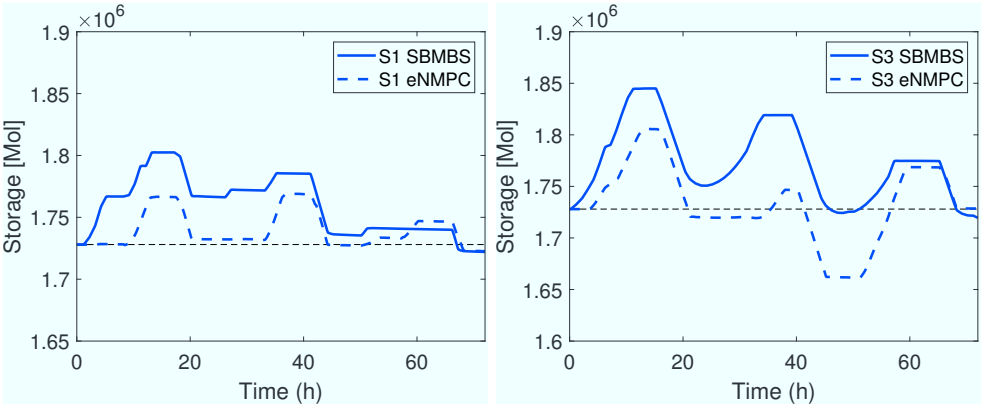


Fig. 4: Power consumption for optimal operation with eNMPC and the SBMBS. SBMBS profiles blue, solid. eNMPC profiles blue, dashed. Electricity prices are shaded. Constant production rate operation black, dotted. (a) Scenario S1. (b) Scenario S2. (c) Scenario S3. (d) Scenario S4.

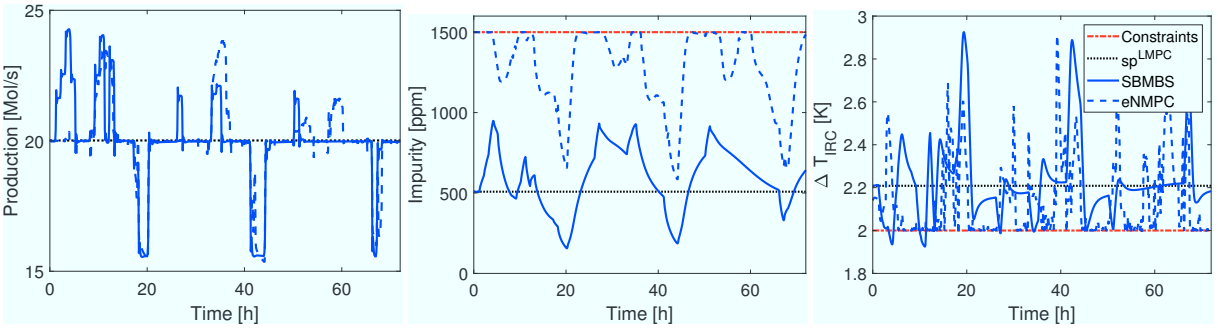
Table 6: Summary of CPU times for eNMPC. The mean and maximum values, and the standard deviations are calculated based on all eNMPC optimizations for each operation scenario.

scenario	mean CPU time [s]	maximum CPU time [s]	standard deviation [s]
S1	285	6269	482
S2	374	5444	332
S3	406	8059	530
S4	358	1198	178



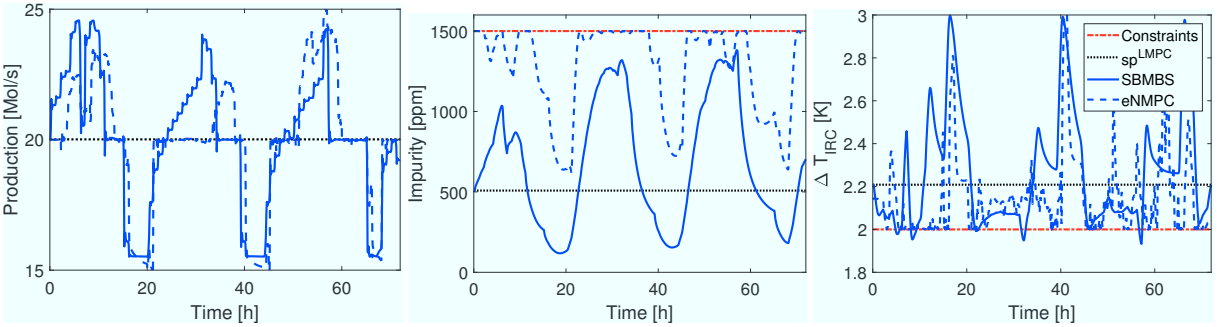
(a) (b)

Fig. 5: Storage level during optimal schedules for scenarios S1 and S4. Constant production rate operation black, dotted. SBMBS operation blue, solid. eNMPC operation blue, dashed. (a) Scenario S1. (b) Scenario S4.



(a) (b) (c)

Fig. 6: Trajectories of selected state variables $\mathbf{y}$ for SBMBS and eNMPC operation and scenario S1 (constant demand, moderate electricity price fluctuation). SBMBS profiles blue, solid. eNMPC profiles blue, dashed. Bounds red, dash-dotted. LMPC setpoints black, dotted. (a) Production rate $\dot{n}_{\mathrm{P}}$. (b) Product stream impurity. (c) Temperature difference in IRC.



(a) (b) (c)

Fig. 7: Trajectories of selected state variables $\mathbf{y}$ for SBMBS and eNMPC operation and scenario S3 (changing demand, extreme electricity price fluctuation). SBMBS profiles blue, solid. eNMPC profiles blue, dashed. Bounds red, dash-dotted. LMPC setpoints black, dotted. (a) Production rate $\dot{n}_{\mathrm{P}}$. (b) Product stream impurity. (c) Temperature difference in IRC.

Table 7: Performance of eNMPC with fast-update method.

scenario	mean CPU time [s]	max CPU time [s]	standard deviation [s]	cost	savings
S1	164	250	32	\$997.83	-2.5%
S3	172	493	43	\$1574.39	-7.2%

implementation. The solution of the SBMBS re-scheduling problem at hour 36 using SBMs required 401s and 410s for scenarios S2 and S4, respectively. The eNMPC problem is solved at every sample time, and statistics about the computation times are given in Table 6. In particular, while SBMBS is solved in <1h for a 72-hour horizon, the eNMPC solutions are solved in ∼5min for a 15-minute sample time. The SBM scheduling dynamic optimization problem involves 72 decision variables (one setpoint discretized by 72 one-hour intervals), while each eNMPC problem involves 192 decision variables (4 manipulated variables discretized by 48 fifteen-minute intervals).

Importantly, while time delays from solving the eNMPC problem were neglected, Table 6 shows that the maximum CPU times found in all four scenarios exceed the 15-minute sample time of the eNMPC system. Therefore, we also investigate a sub-optimal, fast-update method [28] on the scenarios without load change (S1 and S3). In this approach, the number of SQP iterations is restricted to two at each sampling time. This method significantly reduces both the mean and maximum CPU times (Table 7), while still leading to a feasible solution and economic improvements. Specifically, the maximum CPU time is reduced to 30–55% of the sampling time, while the mean CPU time is reduced to <20% of the sampling time. Note that the CPU time equal to ∼20% of the sampling time may still represent a non-negligible delay in controller actuation. Advances to optimization solvers and/or computing hardware could expedite these computations, or enable more SQP iterations (and improved economic performance) while still guaranteeing real-time capability. We note that other fast-update methods have been proposed, cf., [31]. Moreover, recent works [36, 34, 30] have shown that reduced-order modeling approaches may further accelerate eNMPC calculations while still producing control variable profiles similar to those obtained using a full-order model.

Economic evaluation: The objective function values of the implemented solutions are given in Table 5. Both approaches achieve high economic improvements compared to the constant production rate base case. The improvements are higher for the extreme electricity price profile and lower if the load change takes place, compared to the respective scenario without load change. The eNMPC achieves 0.2 up to 2.4 (scenario S2) times higher improvements compared to the improvements of the SBMBS, with the improvements found from the two operation strategies being closer to each other for the more extreme electricity price profiles. Note that even the worst-case 1% savings obtained by SBMBS in scenario S2

is significant in a commodity market such as industrial gases. The eNMPC improvements are higher relative to the improvements of the SBMBS for the moderate electricity price profiles (S1 and S2) than for the extreme profile (S3 and S4). The eNMPC achieves higher economic improvements since it can drive the process faster and more aggressively. Specifically, while the LMPC actually has a shorter sample time, it relies on setpoints that are generated from a higher-level production scheduling problem. These setpoints change at hourly intervals, while the eNMPC directly incorporates the process economics into its 15-minute sample time, which is longer than the LMPC sample time, but shorter than a scheduling time slot. Furthermore, with the exception of production rate, the setpoints for LMPC are left at their steady-state values, and the LMPC explicitly drives the process away from its bounds. While this is representative of practical operation, it translates to more conservative behavior.

On the other hand, the eNMPC uses the exact model of the controlled system and can reliably maintain the CVs at their bounds. Furthermore the eNMPC does not penalize or restrict MV moves, as is typical of LMPC, which would make the MV profiles smoother and less aggressive, in turn slowing the closed-loop behavior. A similar smoothing effect could be achieved using a coarser MV discretization. With the faster and more aggressive operation, the eNMPC can exploit shorter and more moderate electricity price peeks better than the SBMBS. On the other hand, a more aggressive process operation could affect the life time of the process units, e.g., compressor and turbines, and their actuators. While these effects could be incorporated into the production scheduling problem, e.g., [61], we did not consider shortened life spans of actuators or other units as a consequence of flexible operation in this work.

While the eNMPC achieves significant economic benefits, it suffers from the aforementioned computational issues. The optimal operation costs for eNMPC with a fast-update method are given in Table 7. The savings are slightly reduced by constraining the number of SQP iterations (compared to the eNMPC results in Table 5), but the reduced computational times make the method more practically relevant. Higher computational savings for the eNMPC can be achieved by using reduced dynamic models [30, 36]. Moreover, the economic improvements remain higher than those obtained from the SBMBS for the respective operation scenarios.

Power demand profiles: The operation cost follows directly from the power consumption profiles, which are provided for the three-day horizon in Figure 4. Both methods can handle demand-side management well, in response to the given electricity prices (demand response): the instantaneous power consumption in both schedules is approximately in antiphase with electricity price. Specifically, both eNMPC and the SBMBS realize an intuitive economic process operation: the power demand is high when the electricity price is low and vice-versa. This behavior is more apparent for the extreme electricity price

profile than for the moderate one, explaining the higher improvements for the extreme price profile. The eNMPC leads to power demand profiles with larger amplitude fluctuations than the SBMBS, supporting that the eNMPC operates the process more aggressively to obtain higher economic benefits. The eNMPC can exploit in particular the moderate electricity price fluctuations better than the SBMBS; it can make use of small electricity price fluctuations, whereas the SBMBS keeps the production constant during these times. While the production rate (and power consumption in turn) is not explicitly constrained in the eNMPC problem, it still saturates at a minimum value, which corresponds to the lower bound of $\dot{n}_{\mathrm{MAC}}$, or the flow rate through the main air compressor. The production rate also reaches an upper value, which implicitly results from the CV constraints discussed below.

The eNMPC can operate the process faster; the electricity demand is rapidly increased just before the electricity price increase in order to increase production to load the storage tank and reduce the power demand and production when the electricity price is high. On the other hand, SBMBS handles process transitions more slowly; with only a single setpoint profile $\boldsymbol{y}_{\mathrm{sp}}$ to be manipulated for optimization at the scheduling layer, i.e., problem 3, the closed-loop process requires more time for increasing the electricity demand without violating operational constraints (e.g., Figure 4d from 20 h to 40 h). Since the CVs involved in operating constraints are not considered in the upper-level scheduling problem, SBMBS can only avoid constraint violations by the LMPC driving the CVs to their steady-state setpoints. The CVs are not moved away from their setpoints and/or bounds in anticipation of future behavior, and tradeoffs between the CVs are not considered during scheduling. Finally, slower process operation from SBMBS may result from using a LMPC as the tracking controller. Previous work [62] has suggested that, while more computationally intensive, nonlinear MPC can result in faster closed-loop performance compared to LMPC.

Storage tank holdup profiles: The extent of the flexible operation can be seen from the storage tank profiles, which are shown in Figure 5 for scenarios S1 and S3 (no load change). The profiles for the remaining scenarios are provided in the Supplementary Material. The tank is filled when the electricity price is low, and product is withdrawn from the tank otherwise. The tank level profiles from SBMBS operation and eNMPC operation appear qualitatively very similar, which is remarkable, since the eNMPC uses only a 12 h control horizon. Although eNMPC achieves higher economic improvements (Tables 5 and 7), it makes less use of the storage tank compared to SBMBS. Apparently, the eNMPC operates the process closer to the production boundaries: since the storage tank is used less extensively, the production has to be satisfied by the actual process production rate more often. On the other hand, the tank profiles show that SBMBS is able to consider electricity-price trends over a longer time horizon (72 h in this case).

The SBMBS can hold excess storage for extended periods of time, in anticipation of higher electricity prices later in the scheduling horizon. Similar to the electricity demand profiles, the storage tank holdup profiles show the aggressive process operation by eNMPC; product is withdrawn from the tank leading the tank level to decrease below its base level until the production rate is rapidly increased and the tank is refilled (Figure 5b from 40 h to 60 h). However, the endpoint constraint of returning N_R at $t = 72h$ is not exactly met by the eNMPC, since the eNMPC does not consider the full 72-h scheduling horizon and is therefore more strategically myopic.

Profiles of selected state variables: The profiles of the four controlled variables for both eNMPC and SBMBS are shown again for scenarios S1 and S3 (scenarios without load change) in Figures 6 and 7. As expected, the output variables in the SBMBS schedule are guided to their setpoints by the LMPC, while the same variables are guided to their respective bounds by the eNMPC. The LMPC problem seeks to bring the controlled variables, other than production rate, to their base-case setpoint values. These setpoints are not considered in the production scheduling layer, resulting in a smaller scheduling problem 3, but with fewer setpoint profiles $\boldsymbol{y}_{\text{sp}}$ to be manipulated for optimization. In particular, the setpoint for product impurity appears conservative, as the impurity never reaches the bound of 1500 ppm in the SBMBS (Figures 6b, 7b). Adjusting this setpoint may improve process performance (in terms of energy consumption), but it is difficult to determine a priori a value that balances process robustness with operational flexibility. In fact, the impurity levels found in the SBMBS are much closer to the bound of 1500 ppm for the extreme electricity price profile (Figure 7b), even though the same setpoint is used. This suggests that the high degree of process fluctuations (driven by larger fluctuations in electricity price) pushes the SBMBS to operate more similar to the eNMPC, which is confirmed by the similarity in economic performance. While decreasing the weight on impurity in the LMPC may also allow to SBMBS to push the process closer to its bounds for the case of moderate price fluctuations, this would likely result in violation of the impurity constraint given extreme electricity prices. In a broader sense, this observation confirms that the performance of production scheduling is highly dependent on the controller tuning and setpoints, in addition to the schedule itself.

This behavior reveals a primary difference between the approaches: the eNMPC seeks to push the system to its limits to maximize economic performance, while the SBMBS relies on a control system that seeks to maintain output variables at their LMPC setpoint values. Therefore, eNMPC demonstrates improved economic performance by using all four MVs to optimize the economic objective function. Operation at the process bounds is enabled by using the accurate mechanistic process model for eNMPC. Thus, the eNMPC can exploit in particular the moderate electricity price fluctuation with higher improvements

than the SBMBS. On the other hand, SBMBS uses the production rate setpoint as the only setpoint $\boldsymbol{y}_{sp}$ to be manipulated in order to optimize the economic objective function (at the scheduling level), yet can achieve relatively similar economic performance depending on controller tuning. The remaining MVs are set by LMPC, preserving the infrastructure of the conventional control system and the corresponding dynamic stability. Scheduling with SBMBS finds remarkably similar performance to eNMPC in the extreme electricity price scenarios (S3 and S4) even without knowledge of the full process model: SBMBS involves only a low-order data-driven SBM and a linear model for control. Note that electricity price fluctuations similar to the extreme scenario may become more commonplace with increased penetration of time-varying electricity generation from renewable sources (e.g., wind and solar energy).

The production rates (Figures 6a, 7a) for the SBMBS and eNMPC solutions appear similar, although the faster and more aggressive process operation by the eNMPC is again revealed. Intuitively, the production rate profiles closely mirror process power consumption (Figure 4): eNMPC is able to ramp up production more quickly, while the SBMBS operates more conservatively (e.g., at t =20 h) to avoid the violation of operation constraints.

The SBMBS temporarily violates the 2K lower bound on the temperature difference ΔT_{IRC} in the integrated reboiler-condenser (Figures 6c, 7c). For this reason, a back-off approach (e.g., [63]) should be used to account for inaccuracy in the identified SBMs. A different controller tuning may also help in this regard, similar to the above discussion of the impurity constraint. Since the integrated scheduling and control problem is solved over the full scheduling horizon (t_{sched} = 72 h), dynamic optimization with the full-order process model is not tractable in a practical amount of time [52]. In contrast, the eNMPC operation uses the exact process model and does not violate 2K lower bound on the temperature difference ΔT_{IRC} at any point in time. In practical application, the eNMPC may also require back-offs to satisfy path constraints in the presence of disturbances, modeling errors, delay, etc. [64]. This would result in more a conservative operation compared to the eNMPC using an exact process model, as reported here.

Additional profiles: For additional insight into the optimal solutions presented in this work, the unit-wise energy consumption of all solutions (i.e., attribution to turbines, compressor, and liquefier) are given in the Supplementary Material. In these profiles we note that the most significant contributors to the electricity demand of the ASU are the MC and the liquefier, whereas the energy supply from the turbine seems to be negligible. The liquefier is only used temporarily during extreme electricity price peaks, since liquefying surplus product for later usage consumes electricty itself. Therefore, the savings from load shifting must exceed the costs introduced by liquefying excess product. The profiles of the manipulated variables $\boldsymbol{u}(t)$ for all solutions are also provided in the Supplementary Material. The MVs

$\boldsymbol{u}$ reach their bounds often under eNMPC, with potential implications in schedule robustness and/or equipment degradation. On the other hand, the LMPC penalizes large changes in $\boldsymbol{u}$, and $\xi_{turbine}$ and ξ_{HPC} tend to remain close to their nominal steady-state values when the SBMBS schedule is implemented with the LMPC.

In the Supplementary Material, we also provide results from closed-loop simulations with completely unknown load-changes. Here, the change in product demand must be detected through feedback of the storage tank holdup. The results show that both operation strategies are able to handle this scenario well: they achieve even higher economic improvements with respect to the constant production benchmark. The eNMPC again outperforms the LMPC for the case of moderate electricity price fluctuations, while the savings are very similar for the case of extreme electricity price fluctuations.

5 Conclusions

The flexible operation of continuous chemical processes enables economic improvements in dynamic modern markets with fast-changing feedstocks and diversified energy sources. Many top-down and bottom-up approaches for flexible operation have been proposed in the literature, and in this work two approaches representative of these two paradigms are compared. We implement a model of a small air separation unit process as a benchmark process for comparing flexible operation methods and provide a set of prototypical scenarios. An LMPC system is developed for the process, along with a reduced-order scale-bridging model that can be solved to produce optimal schedules while considering the closed-loop dynamics of the process and its existing control system. Further, an eNMPC system is developed for the process and is implemented to simultaneously control the process and maximize its economic performance.

The results on the benchmark process show that both methods enable economic improvements by flexible operation, while each has its own advantages and disadvantages. SBMBS finds lower economic improvements compared to the eNMPC approach, but takes advantage of the existing control system and does not require any new infrastructure. Furthermore, the existing control system maintains the dynamic stability of the system, and the scheduling problem can be solved over a longer time horizon owing to the use of a reduced-order model (the SBM). On the other hand, the eNMPC has access to the full open-loop process model and thus accurately predicts the process behavior, while the SBMs are approximations. Therefore, the eNMPC system operates the process more aggressively and finds larger economic improvements from flexible operation (especially with lower fluctuations in electricity prices), but the nonlinear economic control problem requires new infrastructure for process operations and is more computationally challenging. The computation times found in this study suggest that fast-update

methods and/or model reduction may be required for eNMPC to be real-time applicable in many practical applications.

While this work demonstrates the high potential of eNMPC in flexible operation, future work should investigate further the impact of fast-update methods and/or model reduction that may be required in practice. With respect to industrial application it is also necessary to analyze how the flexible process operation affects the life time of the process equipment as a future work. These methods are particularly relevant when applying eNMPC to chemical processes of larger scale, e.g., multi-product ASUs. On the other hand, future work for SBMBS could investigate alternative data-driven model structures, re-scheduling techniques, and optimization formulations (e.g., backoff constraints, rate-of-change constraints, transfer of degrees of freedom between control and scheduling layers). Motivated by the observation that SBMBS performance is highly dependent on the process control system, future work could study the combination of an upper layer closed-loop scheduling with alternative controllers, such as a subordinate nonlinear MPC that may accelerate process operation. In general, alternative control strategies for both eNMPC and SBMBS can be tested using the benchmark process model, including hierarchical structures, nonlinear lower-level tracking controllers, and/or decentralized control strategies. It would further be interesting to compare top-down and bottom-up approaches for the case of an imperfect model, or a case involving disturbances at the process level, where online state estimation would be required to update the optimal schedule. Similarly, in the absence of full state feedback, a state estimator would be required to retrieve the initial states for the controller model. Finally, the approaches for dynamic optimization could be studied, including alternative DAE integration/shooting methods, warm-starting methods, and optimization solvers. The models provided in gPROMS and Modelica are easily linked with Matlab, Python, and DyOS, which can access a wide variety of optimization solvers and numerical algorithms.

6 Acknowledgment

A.C., A.M., and A.M. gratefully acknowledge the financial support of the Kopernikus project SynErgie by the Federal Ministry of Education and Research (BMBF) and the project supervision by the project management organization Projektträger Jülich (PtJ). C.T. and M.B. gratefully acknowledge funding from the US National Science Foundation (NSF) through CAREER Award 1454433 and Award CBET-1512379. C.T. thanks the German Academic Exchange Service (DAAD) for support through a Short-Term Research Grant, and is grateful for his association to the International Research Training Group (DFG) IRTG-2379 Modern Inverse Problems.

References

[1] A. Mitsos, N. Asprion, C. A. Floudas, M. Bortz, M. Baldea, D. Bonvin, A. Caspari, and P. Schäfer, "Challenges in process optimization for new feedstocks and energy sources," *Computers & Chemical Engineering*, vol. 113, pp. 209–221, may 2018.

[2] M. Baldea and I. Harjunkoski, "Integrated production scheduling and process control: A systematic review," *Computers & Chemical Engineering*, vol. 71, pp. 377–390, 2014.

[3] E. N. Pistikopoulos and N. A. Diangelakis, "Towards the integration of process design, control and scheduling: Are we getting closer?," *Computers & Chemical Engineering*, vol. 91, pp. 85–92, 2016.

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

[5] C. Tsay and M. Baldea, "*110th Anniversary:* Using data to bridge the time and length scales of process systems," *Industrial & Engineering Chemistry Research*, vol. 58, no. 36, pp. 16696–16708, 2019.

[6] L. S. Dias and M. G. Ierapetritou, "From process control to supply chain management: An overview of integrated decision making strategies," *Computers & Chemical Engineering*, vol. 106, pp. 826–835, 2017.

[7] P. M. Castro, I. E. Grossmann, and Q. Zhang, "Expanding scope and computational challenges in process scheduling," *Computers & Chemical Engineering*, vol. 114, pp. 14–42, 2018.

[8] R. Scattolini, "Architectures for distributed and hierarchical model predictive control – a review," *Journal of Process Control*, vol. 19, no. 5, pp. 723–731, 2009.

[9] M. L. Darby and M. Nikolaou, "MPC: Current practice and challenges," *Control Engineering Practice*, vol. 20, no. 4, pp. 328–342, 2012.

[10] L. Würth, R. Hannemann, and W. Marquardt, "A two-layer architecture for economically optimal process control and operation," *Journal of Process Control*, vol. 21, no. 3, pp. 311–321, 2011.

[11] M. Z. Jamaludin and C. L. Swartz, "Closed-loop formulation for nonlinear dynamic real-time optimization," *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 406–411, 2016.

[12] M. Z. Jamaludin and C. L. Swartz, "Approximation of closed-loop prediction for dynamic real-time optimization calculations," *Computers & Chemical Engineering*, vol. 103, pp. 23–38, 2017.

[13] H. Li and C. L. Swartz, "Approximation techniques for dynamic real-time optimization (DRTO) of distributed MPC systems," *Computers & Chemical Engineering*, vol. 118, pp. 195–209, oct 2018.

[14] H. Li and C. L. Swartz, "Dynamic real-time optimization of distributed MPC systems using rigorous closed-loop prediction," *Computers & Chemical Engineering*, 2018.

[15] J. M. Simkoff and M. Baldea, "Production scheduling and linear MPC: Complete integration via complementarity conditions," *Computers & Chemical Engineering*, vol. 125, pp. 287–305, 2019.

[16] R. C. Pattison, C. R. Touretzky, T. Johansson, I. Harjunkoski, and M. Baldea, "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*, vol. 55, no. 16, pp. 4562–4584, 2016.

[17] R. Pattison, C. R. Touretzky, T. Johansson, M. Baldea, and I. Harjunkoski, "Moving horizon scheduling of an air separation unit under fast-changing energy prices," *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 681–686, 2016.

[18] R. C. Pattison, C. R. Touretzky, I. Harjunkoski, and M. Baldea, "Moving horizon closed-loop production scheduling using dynamic process models," *AIChE Journal*, vol. 63, no. 2, pp. 639–651, 2017.

[19] S. Skogestad, "Self-optimizing control: the missing link between steady-state optimization and control," *Computers & Chemical Engineering*, vol. 24, no. 2-7, pp. 569–575, 2000.

[20] S. Skogestad, "Plantwide control: the search for the self-optimizing control structure," *Journal of Process Control*, vol. 10, no. 5, pp. 487–507, 2000.

[21] S. Skogestad, "Control structure design for complete chemical plants," *Computers & Chemical Engineering*, vol. 28, no. 1-2, pp. 219–234, 2004.

[22] S. Engell, "Feedback control for optimal process operation," *Journal of Process Control*, vol. 17, no. 3, pp. 203–219, 2007.

[23] R. Amrit, J. B. Rawlings, and L. T. Biegler, "Optimizing process economics online using model predictive control," *Computers & Chemical Engineering*, vol. 58, pp. 334–343, 2013.

[24] M. Ellis, H. Durand, and P. D. Christofides, "A tutorial review of economic model predictive control methods," *Journal of Process Control*, vol. 24, no. 8, pp. 1156–1178, 2014.

[25] D. Angeli, R. Amrit, and J. B. Rawlings, "On average performance and stability of economic model predictive control," *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1615–1626, 2012.

[26] M. A. Müller and L. Grüne, "Economic model predictive control without terminal constraints for optimal periodic behavior," *Automatica*, vol. 70, pp. 128–139, 2016.

[27] A. Caspari, C. Offermanns, P. Schäfer, A. Mhamdi, and A. Mitsos, "A flexible air separation process: 2. Optimal operation using economic model predictive control," *AIChE Journal*, 2019.

[28] A. Caspari, J. M. Faust, P. Schäfer, A. Mhamdi, and A. Mitsos, "Economic nonlinear model predictive control for flexible operation of air separation units," *IFAC-PapersOnLine*, vol. 51, no. 20, pp. 295–300, 2018.

[29] A. Caspari, Y. M. Perez, C. Offermanns, P. Schäfer, A.-M. Ecker, A. Peschel, F. Schliebitz, G. Zapp, A. Mhamdi, and A. Mitsos, "Economic nonlinear model predictive control of multi-product air separation processe," *Computer-Adided Chemical Engieering*, vol. 46, 2019.

[30] P. Schäfer, A. Caspari, A. Mhamdi, and A. Mitsos, "Economic nonlinear model predictive control using hybrid mechanistic data-driven models for optimal operation in real-time electricity markets: In-silico application to air separation processes," *Journal of Process Control*, 2019.

[31] I. J. Wolf and W. Marquardt, "Fast NMPC schemes for regulatory and economic NMPC – a review," *Journal of Process Control*, vol. 44, pp. 162–183, 2016.

[32] M. Diehl, H. G. Bock, and J. P. Schlöder, "A real-time iteration scheme for nonlinear optimization in optimal feedback control," *SIAM Journal on Control and Optimization*, vol. 43, no. 5, pp. 1714–1736, 2005.

[33] C. R. Touretzky and M. Baldea, "Integrating scheduling and control for economic MPC of buildings with energy storage," *Journal of Process Control*, vol. 24, no. 8, pp. 1292–1300, 2014.

[34] P. Schäfer, A. Caspari, K. Kleinhans, A. Mhamdi, and A. Mitsos, "Reduced dynamic modeling approach for rectification columns based on compartmentalization and artificial neural networks," *AIChE Journal*, vol. 65, no. 5, p. e16568, 2019.

[35] Y. Cao, C. L. E. Swartz, J. Flores-Cerrillo, and J. Ma, "Dynamic modeling and collocation-based model reduction of cryogenic air separation units," *AIChE Journal*, vol. 62, no. 5, pp. 1602–1615, 2016.

[36] A. Caspari, C. Offermanns, A.-M. Ecker, M. Pottmann, G. Zapp, A. Mhamdi, and A. Mitsos, "A wave propagation approach for reduced dynamic modeling of distillation columns: Optimization and control," 2019 under review.

[37] T. Johansson, "Integrated scheduling and control of an air separation unit subject to time-varying electricity prices," Master's thesis, KTH Royal Institute of Technology, Stockholm, Sweden, 2015.

[38] Q. Zhang, I. E. Grossmann, C. F. Heuberger, A. Sundaramoorthy, and J. M. Pinto, "Air separation with cryogenic energy storage: Optimal scheduling considering electric energy and reserve markets," *AIChE Journal*, vol. 61, no. 5, pp. 1547–1558, 2015.

[39] U.S. Energy Information Administration, "Manufacturing energy consumption survey," *https://www.eia.gov/consumption/manufacturing/data/2014/pdf/table11_1.pdf*, accessed 2019.

[40] Process Systems Enterprise, "general PROcess Modeling System (gPROMS)." www.psenterprise.com/products/gproms, 1997-2019.

[41] Modelica, "https://www.modelica.org/," accessed 2020.

[42] C. Tsay, A. Caspari, R. Pattison, T. Johansson, A. Mitsos, and M. Baldea, "A benchmark air separation unit for process control and flexible operation," *Mendeley Data v1*, 2020. doi:10.17632/pfcc5gvzty.1.

[43] J. Du, J. Park, I. Harjunkoski, and M. Baldea, "A time scale-bridging approach for integrating production scheduling and process control," *Computers & Chemical Engineering*, vol. 79, pp. 59–69, 2015.

[44] M. Baldea, J. Du, J. Park, and I. Harjunkoski, "Integrated production scheduling and model predictive control of continuous processes," *AIChE Journal*, vol. 61, no. 12, pp. 4179–4190, 2015.

[45] C. Tsay, A. Kumar, J. Flores-Cerrillo, and M. Baldea, "Optimal demand response scheduling of an industrial air separation unit using data-driven dynamic models," *Computers & Chemical Engineering*, vol. 126, pp. 22–34, 2019.

[46] C. Tsay and M. Baldea, "Integrating production scheduling and process control using latent variable dynamic models," *Control Engineering Practice*, vol. 94, p. 104201, 2020.

[47] Dassault Systemes, "https://www.3ds.com/de/produkte-und-services/catia/produkte/dymola/," accessed 2020.

[48] A. Harmens, "Vapour-liquid equilibrium N_2-Ar-O_2 for lower argon concentrations," *Cryogenics*, vol. 10, no. 5, pp. 406–409, 1970.

[49] D. W. Green and R. H. Perry, "Perry's chemical engineers' handbook," *Choice Reviews Online*, vol. 45, no. 08, pp. 45–4393–45–4393, 2008.

[50] Y. Cao, C. L. Swartz, M. Baldea, and S. Blouin, "Optimization-based assessment of design limitations to air separation plant agility in demand response scenarios," *Journal of Process Control*, vol. 33, pp. 37–48, 2015.

[51] R. Huang, V. M. Zavala, and L. T. Biegler, "Advanced step nonlinear model predictive control for air separation units," *J. Process Control*, vol. 19, no. 4, pp. 678–685, 2009.

[52] L. S. Dias, R. C. Pattison, C. Tsay, M. Baldea, and M. G. Ierapetritou, "A simulation-based optimization framework for integrating scheduling and model predictive control, and its application to air separation units," *Computers & Chemical Engineering*, vol. 113, pp. 139–151, 2018.

[53] Mathworks, "MATLAB Model Predictive Control Toolbox," *https://www.mathworks.com/products/mpc.html*, accessed 2019.

[54] Mathworks, "MATLAB System Identification Toolbox," *https://www.mathworks.com/help/ident/*, accessed 2019.

[55] R. G. Brusch and R. H. Schapelle, "Solution of highly constrained optimal control problems using nonlinear programing," *AIAA Journal*, vol. 11, no. 2, pp. 135–136, 1973.

[56] R. W. H. Sargent and G. R. Sullivan, "The development of an efficient optimal control package," in *Optimization Techniques*, pp. 158–168, Springer-Verlag, 1978.

[57] A. Caspari, J. M. M. Faust, F. Jung, C. Kappatou, S. Sass, Y. Vaupel, R. Hannesmann-Tamás, A. Mhamdi, and A. Mitsos, "DyOS - a framework for optimization of large-scale differential algebraic equation systems," *Computer-Aided Chemical Engineering*, vol. 46, 2019.

[58] R. Hannemann, W. Marquardt, U. Naumann, and B. Gendler, "Discrete first- and second-order adjoints and automatic differentiation for the sensitivity analysis of dynamic models," *Procedia Comput. Sci.*, vol. 1, no. 1, pp. 297–305, 2010.

[59] P. E. Gill, W. Murray, and M. A. Saunders, "SNOPT: An SQP algorithm for large-scale constrained optimization," *SIAM Rev.*, vol. 47, no. 1, pp. 99–131, 2005.

[60] Functional Mock-up Interface for Model Exchange and Co-Simulation, "https://fmi-standard.org/," accessed 2019.

[61] J. Wiebe, I. Cecílio, and R. Misener, "Data-driven optimization of processes with degrading equipment," *Industrial & Engineering Chemistry Research*, vol. 57, no. 50, pp. 17177–17191, 2018.

[62] P. Schäfer, L. F. Bering, A. Caspari, A. Mhamdi, and A. Mitsos, "Nonlinear dynamic optimization for improved load-shifting agility of cryogenic air separation plants," in *13th International Symposium on Process Systems Engineering (PSE 2018)*, pp. 547–552, Elsevier, 2018.

[63] Y. I. Valdez-Navarro and L. A. Ricardez-Sandoval, "A novel back-off algorithm for integration of scheduling and control of batch processes under uncertainty," *Industrial & Engineering Chemistry Research*, 2019.

[64] E. M. B. Aske, S. Strand, and S. Skogestad, "Coordinator MPC for maximizing plant throughput," *Computers & Chemical Engineering*, vol. 32, no. 1-2, pp. 195–204, 2008.