Integration of Scheduling and Control for Chemical Batch Plants under Stochastic Uncertainty: A Back-off Approach

by

Héctor Uriel Rodríguez Vera

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Chemical Engineering

Waterloo, Ontario, Canada, 2021

Author’s Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Due to competitiveness in the chemical industrial sector, methodologies must be developed for an optimal programming of activities of a plant and a safe operation of its equipment. The formulation of a problem that involves the scheduling layer and control layer of the hierarchical manufacturing process results in a very complex model as each layer aims to satisfy different objectives and has specific constraints. Historically, in order to reduce the complexity of the model, each layer was considered as an individual problem, with assumptions that neglected phenomena that was observed in real processes. Thus, the solutions would seem economically attractive at first glance, but very difficult to implement in the practice due to a high level of information incoherence between the inter-phase variables. Consequently, methodologies that integrate the scheduling layer and the control layer need to be developed while considering aspects that may emerge during operation such as model uncertainty or plant-model mismatch.

In this work, two back-off methodologies are presented to address Mixed Integer Dynamic Optimization (MIDO) formulations that arise when modeling the scheduling and control of flow-shop batch plants under stochastic parametric uncertainty. The core idea of the methodologies is to generate scheduling decisions, find control decisions and determine unit operation times that offer dynamic feasibility in the presence of stochastic parametric uncertainty. The MIDO problem is decomposed into a set of problems, with the aim to reduce the required computational time necessary to solve a full fletched MIDO formulation. The set of problems is then solved iteratively. The first methodology (Algorithm A) decomposes the MIDO problem into a scheduling problem, a dynamic optimization problem, a set of dynamic feasibility problems (dynamic feasibility test) and a unit time operation minimization problem. Since Algorithm A is limited by the sequential calculation of control decisions and unit

operation times and that the scheduling does not accurately reflect the *back-off* dynamics of the system, a second algorithm (Algorithm B) was developed to address these issues. This algorithm decomposes the MIDO problem into a parametric sensitivity analysis, a scheduling problem, a dynamic optimization problem and a set of dynamic feasibility problems (dynamic feasibility test). The parametric sensitivity analysis is performed to create correlations that will allow the scheduling problem to consider the *back-off* dynamics of the system. Back-off terms are introduced in the model constraints of the dynamic optimization problem to represent the variability of the system caused by the uncertainty. Stochastic uncertainty is modeled using statistical distribution functions and are embedded in the set of dynamic feasibility problems to test the dynamic feasibility of the optimal control actions under random realizations in the uncertain parameters. The variability in the observed variables caused by the uncertain parameters while performing the dynamic feasibility test are used to calculate the back-off terms. To appreciate and evaluate the effect of the variations in the unit operations times caused by the back-off effect, in the scheduling problem, a continuous-time formulation has been considered and implemented.

A case study featuring a flow-shop batch plant consisting of two dynamic reaction processes and two steady state separation processes is used to illustrate the benefits and limitations of the proposed back-off methodologies. An scenario consisting of a one unit available per process was use to compare Algorithm A with the methodology developed by Yael-Alvarez & Ricardez-Sandoval[1]. The results show that considering varying unit operation times in the back-off methodology increments the computational effort, but the economics of the are improved up to a 42%. Algorithm B was evaluated using scenarios to measure the effects of varying the value of variability considered in the back-off terms and the effects of having multiple available units with different processing capacities. The results show that the amount of variability considered in a back-off term may improve the profits up to a 22% per scheduled job compared to a nominal case (no back-off terms). Due to the stochastic nature of the uncertainty propagation models, only solutions that offer dynamic feasibility under uncertainty are assured.

In general, unit operation times chosen from optimization are better suited to accommodate stochastic parametric uncertainty while the control actions enforce process operational and product quality constraints at reasonable economic costs. Hence, the two methods proposed in this work have the potential of addressing optimal scheduling and control problems under stochastic realizations in flow-shop batch plants.

The first method (Algorithm A) was presented at the 11th International Federation of Automatic Control (IFAC) Symposium on Advanced Control of Chemical Processes (ADCHEM) 2021[2]. The second method (Algorithm B) has been submitted for publication to Industrial & Engineering Chemistry Research (I&EC)[3].

Acknowledgements

Firstly, I would like to deeply thank Prof. Luis A. Ricardez Sandoval for being a very supportive and attentive supervisor, who patiently reviewed and advised my work and performance, guiding me throughout the realization of my Master's degree. I appreciated his guidance through difficult times, and his constant dedication to her students.

I would like to thank my Thesis committee, Dr. Peter Douglas and Dr. Jeff Gostick, for their invaluable feedback and suggestions.

I would like to thank Prof. Jorge Ortega Pérez, Prof. Fabienne Espitalier and Prof. Gloria Elizabeth Moreno Morales for being professors that since my Bachelor's studies have encouraged me to continue learning and to keep looking for opportunities to do so. They are all very supportive and knowledgeable inside and outside the classroom.

I would like to thank my mother, Ana, for always being a loving and caring mother, who has always supported me in every decision through all my life, in all her capacity and who has always encouraged me "to leave my little town and see that there is more in this world". Thank you to my Grandmother, María, and Sister, Annelly, which, even across the distance, gave me their full support and company.

Finally, I would like to thank the Consejo Nacional de Ciencia y Tecnología (CONACyT) of Mexico for providing the financial support necessary to perform my Master studies through a federal scholarship from the Mexican government.

Table of contents

List of Figures

List of Tables

List of Acronyms

DAE	Differential Algebraic Equation
PDF	Probabilistic Distribution Function
MC	Monte Carlo
MIDO	Mixed Integer Dynamic Optimization
MINLP	Mixed Integer Non-Linear Programming
MILP	Mixed Integer Linear Programming
MIQP	Mixed Integer Quadratic Programming
NLP	Non-Linear Programming
LP	Linear Programming
STN	State-Task Network
UOT	Unit Operation Time
EWO	Enterprise-Wide Optimization
IT	Information Technology
PSE	Power Series Expansions
NLMPC	Non-Linear Predictive Model Control

List of Symbols

Indexes

a	Product index
e	Event index
i	Back-off algorithm, master loop index
j	Unit index
k	Task/Process index
m	Back-off algorithm, inner loop index
n	Stochastic scenario index
o	Parametric sensitivity analysis scenario index
p	Mechanistic unit model index
q	Inequality constraint index
r	Equality constraint index
s	State index
t	Time index
w	Parameter index

Sets

N_{Pr}	Set of the processes (tasks) of the chemical batch plant
N_E	Set of available equipment
N_R	Set of recipes
N_p	Set of mechanistic unit models
N_q	Set of inequality constraints
N_r	Set of equality constraints
N_x	Set of system states
N_u	Set of manipulated variables
N_ψ	Set of system parameters
C	Set of cost information
C_i	Subset of costs at the i^{th} iteration
Ψ	Set of model parameters
ψ_{Nom}	Subset of nominal parameters
ψ_{Unc}	Subset of uncertain/random parameters
T	Set of unit operation times
τ_{Dyn}	Subset of unit operation times for dynamic processes
τ_{Fix}	Subset of unit operation times for static processes
STN_i	State-Task Network information set at the i^{th} iteration
δ_i^{In}	Subset for State-Task relation at the i^{th} iteration
δ_i^{Out}	Subset for Task-State relation at the i^{th} iteration
Λ_i	Subset of slopes at the i^{th} iteration (linear correlations)
Γ_i	Subset of intercepts at the i^{th} iteration (linear correlations)

P	Subset of states prices
R	Subset of market demand
Z	Set of solutions of optimization problems
β_{Sch_i}	Subset of scheduling decisions at the i^{th} iteration
κ	Parametric sensitivity analysis set (material load)

Variables

$\zeta_{i,k,j,q,t,n}$	Observed system variable
$\tau_{i,k,j}$	Unit operation time
$\psi_{Unc_{i,w}}$	Random realization of a stochastic uncertain parameter
$B_{i,k,j,e}$	Unit processing material variable
$b_{i,k,j,q,t}$	Back-off term
$D_{a_{i,j}}$	Substance quantity
$d_{i,s,e}$	Satisfied market demand
$f_{p,k,j}$	Mechanistic model equation
$h_{q,k,j}$	Inequality constraint
$\tilde{J}_{i,k,j,e}$	Unit operation cost
$k_{r,k,j}$	Equality constraint
$mid_{k,j}$	Average unit processing material value
N_{MC}	Total number of stochastic realizations
$Q_{i,j,e}$	Material quantity inside a unit
r_s	Total state demand
$ST_{i,s,e}$	State quantity variable
$T^F_{i,k,j,e}$	Finalization time variable
$T^S_{i,k,j,e}$	Initialization time variable
$u_{Dyn_{i,k,j}}$	Control trajectory profiles
$W_{i,k,e}$	Unit assignment decision variable
$x_{i,j,k,t}$	System state variable
$\dot{x}_{i,p,k,j,t}$	System state rate of change
$Y_{i,j,e}$	Binary scheduling decision variable

Parameters

η	PDF parameter
λ	Back-off strength
τ_{max_k}	Maximal unit operation time value
τ_{min_k}	Minimal unit operation time value
$B_{Max_{k,j}}$	Maximal unit processing material value
$B_{Min_{k,j}}$	Minimal unit processing material value
$c_{i,k,j}$	Unit operation cost
E	Event points

e_f	Last event point
H	Time horizon
O_s	State capacity
p_s	State price
t_s	Initial global time
t_f	Final global time
L	Batch size
Tol_{BO}	Tolerance criteria for Eq. 7
Tol_{SS}	Tolerance criteria for Eq. 6
u_{max_k}	Maximal control decision value
u_{min_k}	Minimal control decision value
$\rho_{i,k,s}^{In}$	State-Task relation value
$\rho_{i,k,s}^{Out}$	Task-State relation value
$\delta_{i,k,j,s}^{In}$	State-Task correlation value
$\delta_{i,k,j,s}^{Out}$	Task-State correlation value
$\Gamma_{i,k,j}^{\tau}$	Dynamic time intercept value (linear correlations)
$\Gamma_{i,k,j}^{C}$	Dynamic cost intercept value (linear correlations)
$\Lambda_{i,k,j}^{\tau}$	Dynamic time slope value (linear correlations)
$\Lambda_{i,k,j}^{C}$	Dynamic cost slope value (linear correlations)

Chapter I. Introduction

Traditionally, chemical processes are modelled and optimized as non-interactive problems for each of the different decision-making layers that conform the hierarchical manufacturing process[4–6]. More than often, the assumption of perfect (ideal) conditions of operation is made to allow for a simplification of the mathematical complexity of the model for each layer. While this approach may yield economically attractive solutions for each one of decision layers, the solutions often face multiple complications when an attempt of implementation into real systems is considered. This is mainly due to information mismatch occurring in the intermediary states between the layers, which may eventually result in suboptimal or infeasible solutions that cannot be implemented.

Due to the elevated pressure for high performance and the continuous search for optimal process operation, significant efforts have been made to develop robust strategies that can accommodate the most typical conditions found in chemical manufacturing systems. Enterprise-wide optimization and smart manufacturing advocate for a higher integration of the information by the implementation of Information Technologies (IT), expecting that such procedure will lead into better decision-making processes[7–9]. Numerous methodologies for the integration of different sets of decision-making layers have already been reported in the literature[4–6,10] with a demonstrated improvement in the performance of the studied systems with a varying level of success[11–14]. Integration approaches are performed with the aim of reducing process infeasibility and sub-optimal solutions, while taking advantage of the natural interconnection between the decision-making layers[11,15].

Multi-unit, multi-product chemical batch plants are of great importance for the chemical industry. A batch operation allows for a highly controlled operation and production of highly valued chemical, while remaining quite adaptable to changing market trends. Thus, a multi-unit, multi-product chemical batch plant is a plant consisting of multiple units that can perform different tasks, with units that can be easily rearranged for the manufacturing of various products. The pharmaceutical industry is a prime example of this type of production plants, as some drugs are seasonal, where production must be designed to account for maximum efficiency, as materials and products tend to be quite expensive and require significant energy.

Scheduling is key in manufacturing systems as it aims to allocate resources to tasks that need to be performed to meet process goals and market demands. On the other hand, process control aims to manipulate the available process variables in real time to achieve satisfaction of quality requirements while maintaining the operation stable and within their feasibility limits.

A simultaneous approach of scheduling and control is often set to solve complex, large-scale MIDO problems that operate in multiple time horizons and that are required to ensure stability and dynamic feasibility on the various processes conforming the system. Since MIDO problems are often difficult to solve explicitly, discretization and decomposition techniques are often employed[16]. After a discretization of the differential-algebraic process model equations, a MIDO problem becomes a large-scale Mixed-Integer Non-Linear Programming (MINLP) problem that may become computationally intractable due to model inflation and increase in the complexity of the model. Thus, it is usual to find methodologies in the literature that further decompose MINLP problems involving scheduling and control decisions into sets of Mixed-Integer Linear Programming (MILP) problems and Non-Linear Programming (NLP) problems. The solutions of the subset of problems are then reconciliated by an iterative procedure. A few approaches for integration of scheduling and control that follow this general framework have been proposed in the literature[17–21]. Studies addressing the integration of planning, scheduling and control[22,23] and the integration of design and control[24–32] are also available.

Uncertainty may be inherent to the model (mathematical representation), the process (dynamics and online measurements), external variables (environment) or discrete events (like equipment availability)[33]. Typically, the assumption of the presence of uncertainty in the methodologies for the integration of scheduling and control is omitted or quite often simplified using deterministic assumptions to significantly reduce the computational burden. Nonetheless, scheduling and control studies implementing stochastic parametric uncertainty have been conducted using a two-stage stochastic programming approach[20] and earlier versions of the back-off methodology developed within our group[1]. In those previous works, the optimization of unit processing times and its implications on the scheduling have not been explored. This condition is of great interest as it might lead to an improvement on the efficiency of a system that is subject to stochastic realizations in the uncertain parameters.

In the back-off approach, the key idea is to move away from a highly attractive economic solution which in practice, may be infeasible or sub-optimal due to the presence of uncertainty, into a competitive economic solution that remains dynamically feasible under stochastic realizations in the uncertain parameters. Back-off methodologies have already been previously implemented for control structure selection[34,35]; integration of design, scheduling and control[18]; integration of control and design[31,36]; and integration of scheduling and control for batch systems[1]. In the latter, the performance of the methodology was compared, for multiple scenarios, with the resolution of the MINLP formulation.

I.1. Research Objectives

This work presents two novel decomposition methodologies for the integration of scheduling and control under stochastic parametric uncertainty based on the back-off approach for multi-unit, multi-product chemical batch plants. This study is performed to show that varying unit operation times improve scheduling decisions, control profiles and the overall plant economics. The back-off approach was selected because back-off terms can capture the variability of a system when it is subjected to stochastic parametric uncertainty. Also, back-off terms can be introduced into the formulation of the constraints as parameters to help guide the system into finding a feasible solution that is economically attractive. Uncertainty is assumed to be stochastic because it allows for a better representation of the phenomena occurring in actual chemical batch processes. Uncertainty can be represented by implementing Probabilistic Density Functions (PDF), or a combination of PDFs, that best fits the observed behavior.

The specific objectives of the current thesis are as follows:

- Propose a new back-off methodology for the integration of scheduling and control that can address the simultaneous scheduling and control of multi-unit, multi-product chemical batch plants under the presence of stochastic parametric uncertainty. The MIDO formulation of the integrated problem will be decomposed into a set of NLP and MILP problems that will be solved sequentially and iteratively until a dynamically feasible solution that can accommodate a user-defined level of process variability is found. Back-off terms are used to represent the deviation of a system caused by uncertain realizations in the process parameters.
- Study the impact of varying the unit operation times while implementing back-off terms and measure the corresponding economic benefits.
- Implement a continuous-time formulation for the scheduling formulation that may offer improvements over a discretized time grid. Furthermore, it opens the possibility of implementing correlations that can translate the effects of the back-off terms on the dynamics of the model into the scheduling problem.

- ✐ Inspect the effect of multiple units with different processing capacities under the effect of multiple uncertain parameters, which are propagated through the processing units considered in the chemical batch plant.

By performing this thesis work, it is expected to demonstrate the importance of the simultaneous optimization of unit operations times and unit control profiles in the presence of stochastic parametric uncertainty, while the scheduling decisions account for the optimal regimes of operation.

I.2. Thesis Structure

The thesis is structured as follows:

- Chapter II: In this section a discussion and review of the general state of the art of the field of the integration of scheduling and control that takes into consideration uncertainty is performed. Necessary background concepts are explained in this chapter to facilitate comprehension of the key topics covered in this research.
- Chapter III: This chapter presents the general problem statement of the integrated scheduling and control problem that was considered for the development of the back-off methodologies presented in this thesis. Then, the chapter continues with the presentation of the back-off Algorithm A, the case study used to evaluate its performance and the results. Algorithm A was presented in the 11th International Federation of Automatic Control (IFAC) Symposium on Advanced Control of Chemical Processes (ADCHEM) 2021[2].
- Chapter IV: Algorithm B is an improved version of Algorithm A and is presented in this section. The same case study used for algorithm A is used in this section; in addition, a larger case study involving more processing units are also considered here. This work is currently under review in Industrial & Engineering Chemistry Research[3].
- Chapter V: This chapter presents the conclusions and suggestions for future lines of research.

Chapter II. Literature Review

This chapter presents the current state-of-the-art involving the integration of scheduling and control and the methods used to consider uncertainty into the model. These aspects are key since the difficulty in finding an optimal solution of an integrated scheduling and control problem resides in the inherent interactions between scheduling and control and the complexity associated with the process model. Also, the difficulty is related to the mathematical representation and the numerical approach used to search for an optimal solution. Each of the different scheduling and control approaches reported in the literature comes with assumptions that lead into advantages and restrictions that differentiates them from each other and that need to be weighted before their implementation[4].

This literature review has been performed to clarify the scope of the present work and to better outline their contributions to this emerging area in process systems engineering. This chapter begins with a discussion of the main aspects and challenges faced in process integration for chemical manufacturing systems. Generalities on Scheduling and Dynamic Optimization, the layers of interest for this work, are then discussed. Methodologies available for the integration of scheduling and control is reviewed next. Uncertainty is discussed, and the back-off approach is introduced as method to deal with model uncertainty. The chapter concludes with a summary.

II.1. Enterprise-Wide Optimization and Smart Manufacture

Due to the evolving nature of the markets, continuous efforts are made for the development of robust strategies that can lead to optimal operation and management of chemical industrial plants. Enterprise-Wide Optimization (EWO) suggests that it is possible to achieve an optimal decision-making process and exchange of information through the implementation of IT technologies[8]. EWO seeks the full integration of the hierarchical manufacturing decision layers without compromising the mathematical modeling of each layer. Subsequently, EWO leads to Smart Manufacturing, which envisions a disruptive reorganization and integration of the complete model by allowing an intelligent *generate-plan-apply* process that fully involves all the stakeholders[7]. **Figure 1** exhibits the decision-making hierarchy of the manufacturing layers. The hierarchy describes the typical flow of information and the time frequency at which decisions are made at each stage[4,10,37,38].

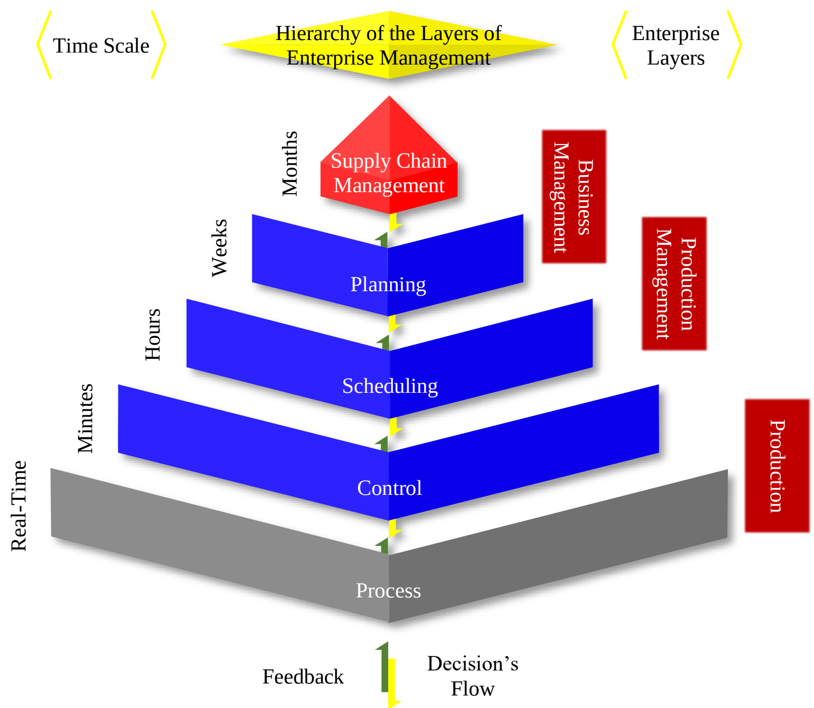

Figure 1. Hierarchy of the Layers of Enterprise Management.

As described by Shobrys and White[37], the decision layers include the following operational aspects:

- Supply Chain Management. This layer corresponds to the Business Management/Logistics of an enterprise. It includes strategic decisions that have long lasting effects (in the order of months) and that affect wide geographical areas. This layer is often related with arrangements in the acquisition of material and logistics. The decisions made at this layer are strategic. Reacting to sudden changes is very difficult at this stage, as it would imply a change of strategy that may result in monetary losses due to changes in planned activities that may require to *pause* the general operations until a new strategy is idealized (e.g., start-up/shut-down units).
- Planning (Production Management). It involves tactical decisions as their time span is shorter (in the order of weeks) and the range of effect is more limited. The decisions taken in this layer are usually limited to one facility while following the operational strategies defined in the previous layer. The decisions made at this layer may adapt to seasonal market trends and forecasts, but it is still quite difficult to react to sudden drastic market changes. This layer must account for the decisions made at the Supply Chain Layer.
- Scheduling (Production Management). This layer establishes the distribution of materials and times for activities necessary to perform a specific process, task or realize a product within the facility. A scheduling plan is usually made to account for the activities of a day of operation; thus, the duration of a decision is usually measured in hours. This layer offers more flexibility of reaction to market trends and external perturbations, e.g., rush orders or changes in product demands. Though a schedule accounts for a short amount of time, it has as objective to meet the production goals decided at the Planning layer.
- Control (Production). This layer makes a real time observation of a process and makes use of interventive actions in the process. The actions made at this layer seek to minimize the effects of disturbances in the process, ensuring a smooth plant operation while achieving the economic goals and objectives set in previous layers. Control actions in chemical plants are usually implemented in the order of minutes or seconds after a change of conditions has been detected.
- Process (Production). The activity or task taking place at the present (real time).

The traditional approach for the modeling and optimization of the decision-making layers in the chemical industry is to consider each as separate problems that must be solved sequentially[4–8]. While this approach may yield economically attractive solutions, the solutions obtained from this approach often face multiple complications when they are implemented into real systems. This is mainly due to information mismatch occurring in the intermediary states between the layers, which eventually results in suboptimal or infeasible solutions that cannot be implemented online. To address this issue, numerous efforts have been made to develop frameworks for the integration of different layers[4,5,10,12,39].

Integrative approaches have been gradually developed and adopted as the technological progress has made it possible[37]. The core concept behind the integrated approach is the formulation of a single mathematical model that fully describes the decisions of multiple layers. Integration approaches are performed with the aim of reducing process infeasibility and sub-optimal solutions, while taking advantage of the natural interconnection between the decision-making layers[11,15]. In an integrated problem, each hierarchical layer still operates in a different time horizon and has their own set of objectives, making it a very complex problem. In addition, the size of an industrial problem adds a high level of dimensionality to the model plus the fact that enterprises vary in complexity depending on its business category. Furthermore, the difficulty of integrating layers aggravates as general framework to do so still does not exist. Nonetheless, numerous methodologies for the integration of different sets of decision-making layers have already been reported in the literature[4–6,10] with a demonstrated improvement in the performance of the studied systems with a varying level of success[11–14].

Another important activity that is also taken into consideration in the area of process integration is Process Design, which consists in the mathematical conceptualization of the phenomena (physical and/or chemical) that describes the transformation of raw materials (ingredients) into products. This aspect involves the design of the equipment where the transformation processes will be performed, defining the mode of operation and recipes of production. Process Design is not often presented in the hierarchy of the layers of enterprise management as it is an activity of project management. Nevertheless, process design decisions also impacts the design of control schemes, which refers to the actions that are taken during operation to ensure that production objectives are met. The typical approach has been to first design the operation units and then adapting control schemes to the process. As this is not optimal, since process designs may impose constraints on the dynamic performance in closed-loop, integrated design and control schemes are also considered[10,22,23,28,40,41]. The present thesis focuses on the integration of the scheduling and the control layers since it deals with the improvement of the operation of chemical plants that have already been designed and are operational. The following sections will therefore discuss some generalities about these two layers and their integration.

II.2. Scheduling

This section serves to provide a general background on scheduling, which is one of the layers of interest in this thesis work. As stated in the previous section, scheduling is a decision-making process, which is key for the industry on the planification of daily plant activities. In general, scheduling aims to determine the optimal allocation of finite resources available to the tasks (processes) that need to be performed to meet the demands of the market and satisfy the economic objectives of the plant.

II.2.1. Time Representation

Scheduling formulations can be mainly classified into two types of categories accordingly to the way time is mathematically represented in the model[39,42–44]:

- Discrete-Time Formulations: This approach divides the time horizon into a finite number of uniform intervals, thus defining a time grid for the variables and constraints of the model. An event may only take place at the boundaries of these time periods. Processing times are assumed to be constant and integer multiples of the Δt used for the intervals. This representation offers ease of modelling and interoperability. Furthermore, a high degree of flexibility can be achieved with a finer time discretization. However, a small discretization may incur in larger number of variables causing intractability for large-scale applications or minute discretization periods.
- Continuous-Time Formulations: This approach introduces a time grid where events may take place at arbitrary time points in the given time horizon. The concept of variable event times, which can be defined globally or for each unit, is also introduced. Variables are required to determine the timings of events. Since the duration of events, start time and finish time are not defined *a priori*, the mathematical model tends to be small, compared to discrete-time formulations, especially because no inactive time intervals must be modeled. However, the scheduling process is overall more challenging to model and complex structures might be needed.

Figure 2 shows the graphical representations of a scheduling plan according to their time representation.

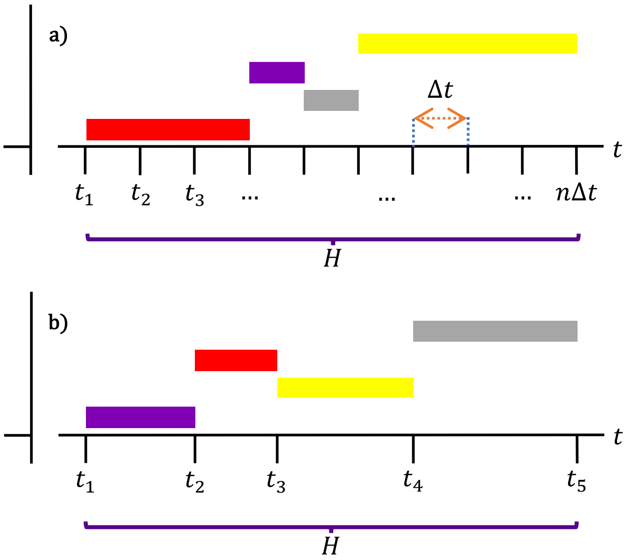

Figure 2. Scheduling Representations.
(a) Discrete-Time Representation, **(b)** Continuous-Time Representtation.

II.2.2. Process Representation

According to the complexity of the process, processes can be described as follows[43,44]:

- Sequential Processes: Multiple products are manufactured following the very same processing sequence. Processing stages may be defined. Since batches are the representation of production, it is not necessary to consider mass balances explicitly.

- Network Processes: This type of networks is used to represented complex production sequences where the flow of material merges or splits from batch to batch. Material balances must be considered explicitly.

Furthermore, there are two types of network representations[39,42–44]:

- State-Task Network (STN): Introduced by Kondili et al.[45], the STN of a chemical process is diagram constituted by two types of nodes: State Nodes (circles), which represent raw materials, intermediary materials or final products; and Task Nodes (rectangles), representing a task/subprocess. Fractions of material incoming and outgoing from a task is given near the corresponding linking line (except if it is 1). No unit is preassigned on the model and batches are of variable size.

- Resource-Task Network (RTN): Introduced by Pantelides et al.[46] as an extension of the STN, has as main feature the unified representation of processing equipment, storage, material transfers and utilities.

Figure 3 shows the graphical representations of a general sequential process, a general network process, a general STN and a general RTN, as previously described. The advantages and disadvantages of each representation are thoroughly discussed elsewhere[47]. The two back-off methodologies presented in this work make use of STN representations since it is able to capture the main features of the case studies used to evaluate the back-off algorithms.

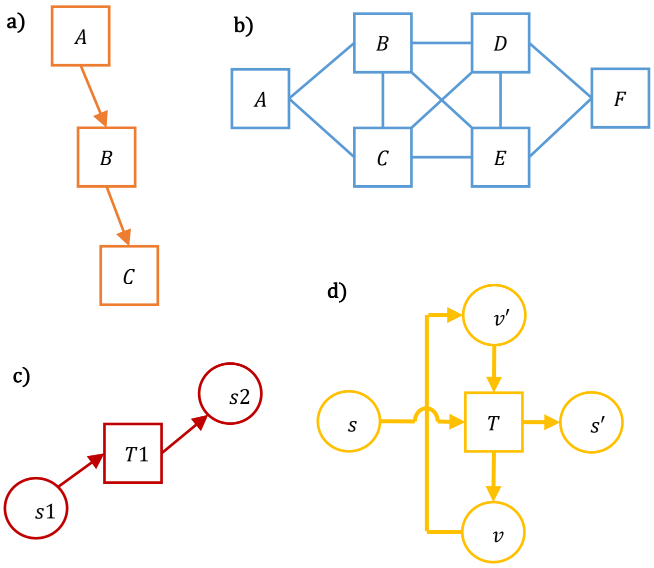

Figure 3. Process Representations.
(a) Sequential Process, **(b)** Network Process, **(c)** State-Task Network, **(d)** Resource-Task Network.

II.3. Dynamic Optimization

In the two back-off decomposition algorithms presented in this work, manipulated variables are treated as decision variables when the dynamic model of the system (Problem 1, section III.1) is optimized. This section serves to introduce Dynamic Optimization, which refers to the optimization and control of a process in open loop, i.e., no feedback strategy is considered.

Dynamic optimization encompasses a wide area of applications such as design of distributed systems, online and offline control applications, trajectory optimization, batch plant optimization and parameter estimation. A dynamic optimization problem may be that which is described by an implicit set of differential-algebraic equations (i.e., conservation laws) with initial conditions or boundary conditions and algebraic equations (i.e., constitutive equations and equilibrium conditions). The decision variables (degrees of freedom) are control variables (manipulated variables) and time-independent variables (i.e., parameters, initial conditions, etc.).

The process of finding a solution for dynamic optimization problems may rely on concepts from optimal control theory for the simpler problems and numerical methods for more complex problems. The numerical methods can be divided on 2 approaches: (1) First optimize then discretize (indirect approaches) and (2) First discretize then optimize (Sequential and Simultaneous approaches). The direct transcription approach is utilized for the development of the methodologies described in this thesis. Direct transcription approaches are fully simultaneous approaches, meaning that there is a full discretization of all the system variables (states, control profiles and state equations). The typical discretization methodology is orthogonal collocation on finite elements because of its usefulness in problems whose solution has steep gradients, bacuese it can be applied to time-dependent problems and because it is numerically stable on stiff problems (A-stable). The resulting large nonlinear problem may be addressed with the usage of NLP solvers[48,49].

II.4. Integration of Scheduling and Control

According to Engell & Harjunkoski[12] scheduling and control share a certain number of similarities that can be exploited for their integration. For instance, both are “real-time” decision-making layers, as they operate in the lower time scales (*vs.* the other layers) and need to quickly adapt to sudden changes in the plant. Also, both layers must deal with uncertainties, unit operation times, product yields and material consumption. Engell & Harjunkoski[12] also note some advantages while performing the integration of scheduling and control:

- Reduction in maintenance requirements and improvement in equipment life-time expectancy.
- Smarter exploitation of the degrees of freedom.
- Information more readily and timely available in scheduling, thus, scheduling decisions are less prone to errors or infeasibilities.
- Reduction in profit losses related to set-ups, changeovers and transitions.

A complex large-scale MIDO problem often arises when a simultaneous approach for scheduling and control is considered. A MIDO problem operates in multiple time horizons, and it is required that stability and dynamic feasibility is ensured for every system or unit included in the operation. Discretization and decomposition methods are often used on MIDO problem formulations due to their inherent difficulty to be solved explicitly[16]. Once the differential-algebraic and control equations of a MIDO problem have been discretized, it becomes a large-scale MINLP problem. A MINLP problem may be computationally intractable due to the elevated number of equations. Consequently, it is usual to further decompose MINLP problems into sets of MILP problems and NLP problems. The individual solutions of these problems are then reconciliated by an iterative procedure. Numerous methodologies to deal with the integration of scheduling and control can be found on the literature. In a chronological order the contributions to the field will be revised next.

As the importance of integrate solutions attracted more interest, at the end of the last century Bhatia & Biegler[50] showed the importance to include dynamic models that describe batch processes when plant activities are being scheduled and the economic benefits of doing so. Then, Mishra et al.[51] compared the standard recipe approach with the overall optimization approach and demonstrated the challenges associated with a MINLP formulation for batch processes and the advantages when these problems are solved using commercial solvers. Nyström et al.[52] then introduced an Iterative Decomposition Algorithm for MIDO problems with the possibility of online implementation for non-large non-complex processes, which later was generalized for multiple units[53]. Flores-Tlacuahuac & Grossmann[54] proposed a methodology for MIDO formulations where the dynamic modeling is directly embedded in the scheduling formulation. The procedure relies in discretization by using orthogonal collocation points. The methodology was implemented for the optimization of the operation of a CSTR. Later, Flores-Tlacuahuac & Grossmann[55] expanded the methodology for multiple parallel production lines and by making the assumption that the system dynamics are represented in terms of lumped parameters systems. Their method was successfully implemented for the operation of tubular reactors. Chu & You[21] developed a decomposition methodology using a General Benders Decomposition (GBD) algorithm framework for a batch plant. Touretzky & Baldea[56,57] then introduced the MPC to the area by developing an Economic MPC (E-MPC) for the control of building Heating, Ventilation and Air Conditioning (HVAC). The results show and excellent control performance and fast computational times, which provide potential for online implementations. Zhuge & Ierapetritou[58] introduced a multi-parametric MPC (mp-MPC) for online implementation in batch processes. The results suggest that a trade-off needs to be done for online implementations: a simplification of the model results into a suboptimal solution. Du et al.[59] presented a robust Scale-Bridging Model (SBM) framework for continuous processes. The results show that this framework has comparable performance to approaches using a perfect process model. Zhuge & Ierapetritou[60] developed a decomposition approach for MIDO problems where the scheduling problem and the dynamic optimization problem can be solved independently from the other. With the implementation of this methodology to a CSTR it was proven that simpler systems do not need to update to satisfy varying market demands. Beal et al.[17] proposed a nonlinear discrete-time formulation aided by with sample pseudo-binary variable functions for a CSTR. The implementation returned favorable results. Dias & Ierapetritou[10] implemented a methodology for an air separation unit and obtained results with an improved performance on the unit's operation, but uncertainty was not considered during the unit operation optimization. Kelley et al.[61] introduced a MILP framework based on deriving data-driven surrogate models of closed-loop process dynamics (SBMs) and linearizations of Hammerstein–Wiener and Finite Step Response (FSR) model forms in the scheduling formulation. The results showed low computational times and similar (or improved) economic results when compared to other methodologies reported in literature. Caspari et al.[62] compared the performance of top-down and bottom-up approaches. While a discussion in depth of the advantages and disadvantages of each approach is made, top-down approaches offer lower computational times at the expense of reduced economic profits compared to bottom-down approaches. Simkoff & Baldea[63] developed a MPC framework where the Karush-Kuhn-Tucker (KKT) conditions of the controlled are embedded in a NLP scheduling formulation. Their results show that implementing this methodology decreases the information mismatch between the scheduling plan and the actual performance of the process. More recently, Andréz-Martínez and Ricardez-Sandoval[15] presented a switched system formulation for multi-product continuous processes. The methodology avoids the use of integer variables by employing an NLP formulation. Their results show that it offers to similar solutions than those obtained with a

MIDO formulation at lower computational costs and without the need of solving integer decisions thus making it a more efficient framework compared to studies reported in the literature. The modeling of the scheduling problem may rely on discrete-time formulations[10,50,52,53], continuous-time formulations[17,20,21,51,56–59] or cyclic scheduling[15,19,54,55,60,64].

It is important to account for the uncertainty and disturbances when performing the optimization of a model because there is the potential of rendering optimal solutions sub-optimal, unfeasible or even making a plant to operate under undesirable (e.g., unstable) conditions. Very few studies have considered the presence disturbances[19,64,65] or uncertainty[1,13,14,18,20,65] in their formulations. Zhuge & Ierapetritou[19] presented a closed-loop implementation for chemical processes subject to disturbances. In their methodology, the integrated MIDO problem is discretized using the implicit Runge-Kutta method. The implementation of the methodology into their case studies effectively rejected most of the disturbances. Chu & You[20] considered parametric uncertainty in the kinetics and a scheduling with varying processing times in their two-stage stochastic programming and GBD approach, though interaction between units affected by uncertainties were not accounted for. Zhuge & Ierapetritou[64] proposed a fast MPC framework for online implementation in batch processes subject to disturbances. The results from implementation of the methodology suggest that it is faster than a previous work[58] and that it is better suited for economic objectives. The methodology uses cyclic scheduling. The back-off approach[1,13,14,18], which will be discussed in more detail in the following sections, has been implemented with iterative decomposition algorithms to account for uncertainty. Simkoff & Baldea[42] developed a data driven NLP methodology applied to the operation of a Chlor-Alkali plant in which uncertainty is accounted for with the capacity of offering solutions in seconds or minutes. Nonetheless, machine learning algorithms rely on historical data of operation which might not always be readily available or insufficient. More recently, Santander & Baldea[65] considered bounded disturbances and uncertainties when developing a two-stage stochastic programming scheduling formulation with processing time relationships from historical data which offers robust solutions.

The effects that parametric uncertainty might have on the total time of operation of the processing units are not directly explored on most of the previous studies. This is also very important, since the operation time is usually set to be a fixed parameter when, due to uncertainty, it might need to be varied to account for the variations in these parameters. This thesis focuses on the effects that stochastic parametric uncertainty may pose when performing the optimization of operation times for each scheduled unit operation, and how this affects scheduling and control decisions for chemical batch plants. Given that uncertainty is a central topic in this thesis, this concept is reviewed next.

II.5. Uncertainty

According to Pistikopoulos[33], uncertainty may be classified into 4 categories: (I) Model inherent uncertainty (mathematical representation), (II) Process inherent uncertainty (dynamics and online measurements), (III) Uncertainty inherent to external variables (environment) or (IV) Uncertainty inherent to discrete events (like equipment availability). The quantification of uncertainty may be deterministic (finite number of possible scenarios known *a priori*) or stochastic (uncertainty is a random variable) and uncertainty can be dealt by reacting to its appearance or by preventing its effects on the system[4,39]. Furthermore, uncertainty may be modeled by a bounded description (lack of sufficient information to accurately describe the uncertainty), a probabilistic description (described by probability density functions) or a fuzzy description (use of fuzzy sets)[39].

Although the importance of accounting for uncertainty was recognized in the second half of the past century[29], as its presence has important effect on the operation of processes, it still remains a challenge[4,37,44]. Uncertainty becomes even more important when there is an exchange of information between different layers, as each layer has different sources of uncertainty, causing a more nuanced information mismatch as it propagates from one layer to another, leading to highly suboptimal or infeasible solutions. The importance of uncertainty is exacerbated when constraint violations rise safety and environmental concerns[4]. In addition to this, the propagation of uncertainty, performed through numerous simulations to quantify its effects on the operation of a given system, quickly becomes a computationally expensive stage in many methodologies[66]. To address the problematic that uncertainty impose in real system's operation, two main types of approaches are often considered, i.e., reactive and preventive approaches[67].

Reactive approaches first look to obtain the solution of the integrated problem at nominal conditions (no uncertainty). Then, as uncertainty occurs, the solutions are updated. These approaches are computationally expensive and are unable to ensure feasibility on all the time horizons of operation. Decomposition and parametric programing are usual techniques employed in integration of scheduling and control[39,67].

Preventive approaches integrate uncertainty when the solution of the integrated problem is being calculated. These approaches aim to generate solutions that remain feasible for most of the range of uncertainty considered. Stochastic programming, fuzzy programming, robust optimization and sensitivity analysis are techniques used in these approaches[39,67].

Another methodology that has been used to deal with uncertainty is the back-off approach. The key idea behind the back-off approach is to perform a systematic iterative movement from a nominal point of operation, which is highly economically attractive, but dynamically infeasible in the presence of uncertainty, to another point of operation that remains feasible. This is possible through the implementation of back-off terms, which are the representation of the deviation of the system under the presence of uncertainty. An important feature of this approach is that it can deal with stochastic uncertainty and the level of robustness of the implementation can be specified, avoiding the need to explore for the worst-case scenario[1,18]. The framework of the back-off method was first proposed by Perkins et al.[68] for simultaneous design and control, and it was later expanded by[69] Kookos et al.[31,35,36,44]. More recently, our research group has developed other implementations of the back-off method. Mehta & Ricardez-Sandoval[70] and Rafiei & Ricardez-Sandoval[24,26] developed a framework using Power Series Expansions (PSE) for the integration of design and control under uncertainty. Palma-Flores et al.[27] proposed a NLMPC implementation for the integration of design and control. Koller & Ricardez-Sandoval[18] successfully implemented the back-off method for the integration of design, scheduling and control under stochastic uncertainty of a multipurpose CSTR. Valdez-Navarro & Ricardez-Sandoval[1] implemented the back-off method for the integration of scheduling and control under stochastic uncertainty of a multiproduct reactor-filter-distillation system. Palma-Flores & Ricardez-Sandoval proposed a methodology based on PSE to simplify NLMPC implementations for the integration of design and control[71].

The work presented in this thesis expands upon these previous works by implementing the back-off approach for the integration of scheduling and control under stochastic parametric uncertainty, described by probabilistic density functions, for multi-unit, multi-process batch plants. The goal is to offer a methodology that finds optimal operation times and control profiles, for all the scheduled unit operations, that offer a feasible and economically attractive operation of chemical batch plants.

II.6. Summary

Numerous methodologies have been developed for the integration of scheduling and control with the objective of reducing the computational burden that represents the search of optimal solutions for complex operations in chemical batch plants. Although it has been shown that it is possible to find an optimal solution for simple case studies, most of the methods involve the use of decomposition techniques for the MIDO formulation of the scheduling and control problem. It is still quite difficult to solve a full MIDO problem formulation of a complex chemical plant in relatively low computational times. Moreover, scheduling formulations that account for varying unit operation times can be found on the literature. Nevertheless, none of the procedures seem to implement a feedback procedure where the parameters used to calculate the operation times are updated through the implementation. In addition, the application of data driven models has already been explored and reported in the literature for optimization procedures of chemical plants. Recent publications account for the presence of disturbances and uncertainties in the model, while returning solutions that do show robustness and an improved operation regimen. Despite these efforts, these methodologies rely on extensive sets of historical data of the operation of the chemical plant.

Based on the above, there is the need to develop a methodology for the integration of scheduling and control that can offer a more accurate scheduling plan of the process and operating policies that remain dynamically feasible in the presence of stochastic parametric uncertainty.

Chapter III. Back-Off Decomposition Algorithm A

In this chapter, the general problem formulation for the integration of scheduling and control is presented, followed by the first Back-Off Decomposition Algorithm developed in this study (Algorithm A).

The general formulation of an integration problem for scheduling and control under stochastic parametric uncertainty is introduced in section III.1 to show the mathematical complexities involved when solving integrated scheduling and control problems. Algorithm A was developed to address the issues that arise from solving the full MIDO problem introduced in section III.1 while still providing a feasible attractive solution in acceptable computational times. A case study is used to analyze the performance of this algorithm, and the results are presented in this chapter. The information regarding Algorithm A has been presented at the 11th IFAC on ADCHEM 2021[2].

III.1. Problem Statement

Batch processes are very important to the industry since they are employed in a wide range of industries, ranging from the pharmaceutical industry to the food industry, where flexibility to sudden market changes is important and where quick product reformulations must be performed. Solving an integrated problem of large-scale multi-unit, multi-product batch plants is very difficult due to the high quantity of variables and equations.

The mathematical conceptual basis for the general formulation of the optimization problem, which is also considered for the back-off frameworks presented in this thesis, is established in this section. The considerations that will be stated next, define the scope of the integration of scheduling and control problem.

Given:

- A flow-shop multi-unit multi-product batch plant that is composed by N_{Pr} set of tasks, with N_E set of equipment that rely on N_R set of recipes.
- A chemical process (task) described by mechanistic dynamic functions f for the N_p states of the system and expressions h that encompasses the set of N_q environmental, safety, product quality and/or operational constraints.
- The plant is described by a set of parameters Ψ, composed by deterministic model parameters defined a priori (ψ_{Nom}) and uncertain stochastic parameters (ψ_{Unc}). The latter can be characterized by probability density functions, assumed to be known a priori.
- A set C that considers the cost information of all raw materials, waste, by-products, and the price information of the products.
- The plant is assumed to operate under a finite timespan (H) specified a priori, from an initial time t_s to a final time t_f, with a finite number of event points (E) that can be estimated using the procedure described in the literature[42].
- A set of unit operation times (T), comprised of a subset for tasks that are assumed to operate at steady-state (τ_{Fix}) and tasks that are driven by their transient operation (τ_{Dyn}).

- ✐ An economic function (Z_{MIDO}) that considers associated costs determined by operational costs related to regime and unit assignment, unit operational times, material costs, product profits and penalties incurred during operation (e.g., by products/waste generation).

An optimization problem that leads to an improvement in the process economics of the batch plant described above can be formulated. This problem is expected to search for an optimal scheduling plan (S_C), optimal control profiles (u) and optimal unit operation times (τ_{Dyn}), which under a set of uncertain (stochastic) parameters (ψ_{Unc}), are ought to hold dynamic feasibility of the flow-shop batch plant. This problem can be mathematically formulated as follows:

$$\min_{u_{k,j}(t),\tau_{Dyn_{k,j}},s_{C_{e,k,j}}} Z_{MIDO}\left(x(t)_{k,j},u(t)_{k,j},\psi,\tau_{Dyn_{k,j}},s_{C_{e,k,j}},c\right) \tag{1}$$
$$s.t.$$
$$f_p\left(x(t)_{k,j},\dot{x}(t)_{k,j},u(t)_{k,j},\psi,t,\tau_{Dyn_{k,j}},s_{C_{e,k,j}}\right)=0,\forall\, t,e\in E,p\in N_p,k\in N_{Pr},j\in N_E$$
$$h_q\left(x(t)_{k,j},\dot{x}(t)_{k,j},u(t)_{k,j},\psi,t,\tau_{Dyn_{k,j}},s_{C_{e,k,j}}\right)\le 0,\forall\, t,e\in E,,q\in N_q,k\in N_{Pr},j\in N_E$$
$$u_{min_k}\le u_{k,j}(t)\le u_{max_k},\ \forall\, t,\ j\in N_{E_k},\ k\in N_{Pr}$$
$$\tau_{min_k}\le \tau_{Dyn_{k,j}}\le \tau_{max_k},\forall\, e,\ j\in N_{E_k},\ k\in N_{Pr}$$
$$\tau_{Dyn_{k,j}}\in\tau_{Dyn},\forall\, j\in N_{E_k},k\in N_{Pr}$$
$$x\in X\subseteq\mathbb{R}^{1\times N_x\times N_E\times N_{Pr}},\ u\in U\subseteq\mathbb{R}^{1\times N_u\times N_E\times N_{Pr}},\ c\in C\subseteq\mathbb{R}^{1\times N_C}$$
$$\psi_{Nom},\psi_{Unc}\in\Psi\subseteq\mathbb{R}^{1\times N_\psi},\ \tau_{Fix},\ \tau_{Dyn}\in T\subseteq\mathbb{R}^{1\times N_\tau}$$
$$s_{C_{e,k,j}}\in\{0,1\},\forall\, e\in E,\ j\in N_{E_k},\ k\in N_{Pr}$$
$$t\in\left[t_s,t_f\right],\ H=t_f-t_s$$

where f_p represents the p^{th} differential-algebraic equation (DAEs) of the system's model. h_q represents the q^{th} model constraint. $x(t)_{k,j}$ represents a state variable of the system for task k taking place in unit j and $\dot{x}(t)_{k,j}$ represents its time derivative. $u_{k,j}(t)$ represents the time-dependent control decisions necessary for unit j to operate while carrying out task k. This variable considers a pre-specified upper (u_{max_k}) and lower (u_{min_k}) saturation limits. $\tau_{Dyn_{k,j}}$ represents the unit operation time of unit j necessary to carry out task k and is delimited by a maximum (τ_{max_k}) and a minimum (τ_{min_k}) allowed time defined *a priori*. $s_{C_{e,k,j}}$ represents the set of integer and continuous scheduling decisions that specify the production schedule for the batch plant at event e for each unit j and task k. An event point e represents a time instance allocation of a task being realized and/or the utilization of a unit[43]. Note that sharing an event point does not mean collusion, as the usage of a unit or the realization of a task may happen at different points in the time domain.

Problem 1 can be casted as an infinite-dimensional stochastic MIDO problem, which is often difficult to solve explicitly, particularly for systems involving multiple processing units. To circumvent this issue and find optimal solutions in reasonable computational times, decomposition techniques leading to MILP/NLP problem formulations have been suggested. In those decomposition methods, each problem defines a simplified specific sub scenario of interest from the main problem.

In this work, two back-off decomposition methodologies are presented to obtain a feasible, approximated solution of Problem 1.

A key novelty of the methods presented in this work is the consideration of unit operation times as decision variables while optimal control profiles are optimized for the tasks to be implemented during plant operation. Each methodology will be explained in their corresponding chapter. The key idea of the back-off approach is to make use of back-off terms, which aim to represent the variability of the system under the effect of stochastic (random) realizations in the uncertain parameters (ψ_{Unc}). Back-off terms are implemented into the constraints of the dynamic system and then utilized to drive the system to a new dynamically feasible and attractive economic solution that can accommodate stochastic parametric uncertainty. In addition, back-off terms are relatively easy to implement. Other methodologies have been previously implemented to deal with similar formulations to Problem 1, which either solve the problem using a direct approach, decompose the MIDO problem and then solve the resulting subset of problems or make assumptions in order to simplify Problem 1. Those previous methodologies had been limited by the reduction of the possible values for the uncertainty, not considering a continuous-time representation in the scheduling problem or by not correctly portraying how the uncertainty that affects one unit carries to the next processing units. Also, the new back-off algorithms were developed because previous back-off approaches considered fixed unit operation times in the scheduling problem and in the dynamic optimization formulation (i.e., unit operation times were considered as model parameters along all the methodology[1]).

The methodology of the first attempt at giving a solution to Problem 1, i.e., Algorithm A, is presented next.

III.2. Methodology

The key novelty of Algorithm A consists in the optimization of unit operation times while implementing back-off terms into the model constraints. The back-off terms represent the variability observed in the controlled variables due to parametric uncertainty. The calculated scheduling plan and control decisions will be capable of accommodating stochastic parametric uncertainty, which in combination with optimal unit operation times, results in dynamically feasible and economically attractive solution to Problem 1 (section III.1).

In the back-off decomposition approach A, the MIDO Problem 1 is decomposed as follows: (1) A scheduling problem, performed to obtain scheduling decisions for the plant that determines material transfer between units. (2) A dynamic optimization problem, which determines the optimal control regimes of the process in the presence of the back-off terms imposed on the process constraints. (3) Monte Carlo simulations, used to generate sufficient statistical data to update the calculation of the back-off terms. (4) A set of optimization problems aimed to obtain the minimal time values at which each unit can operate, dictated by the back-off terms and control profiles determined by step (2) and (3). The set of the four problems is solved in an iterative fashion until a feasible scheduling and control solution that can accommodate the uncertainty in the system is identified. An illustration of the algorithm is shown in **Figure 4**. Each of the steps in this algorithm are presented next.

III.2.1. Initialization

To initiate the algorithm, the sets T and Ψ must be defined; also, initial values are needed for the specification of the State-Task Network (STN) for the scheduling problem (i.e., E, H, ρ_0, P, C_0). Moreover, it is necessary to define the probabilistic distribution function (PDF) and their corresponding parameters (η) that will describe each parameter conforming ψ_{Unc} (i.e., for an element w: $\psi_{Unc_w} = PDF_w(\eta_w)$. This data could be obtained from process heuristics, a sensitivity analysis or from historical data. Tolerance criteria (Tol_{SS} & Tol_{BO}) is also required for initialization. The iteration index for the algorithm is also set, i.e., $i = 0$.

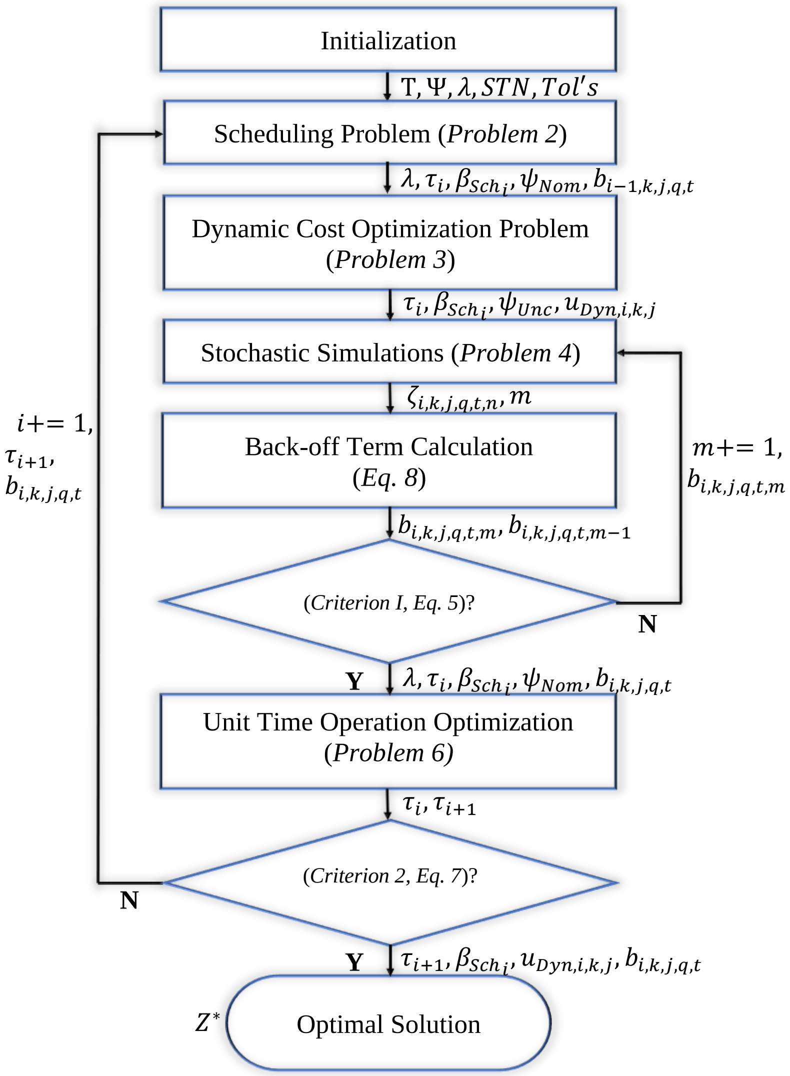

Figure 4. Back-off Decomposition Algorithm A Diagram.

III.2.2. Scheduling Problem

A continuous-time formulation and the State-Task- Network (STN) are used as basis for the formulation of the scheduling problem. The solution from this problem will determine the sequence of unit operation and the quantity of material processed per unit. The scheduling formulation is represented by the following MILP problem:

$$\max_{W_{i,k,e}, Y_{i,j,e}, B_{i,k,j,e}} Z_{Sch}\left(c_{i,k,j}, p_s, \rho_{i,k,s}^{in}, \rho_{i,k,s}^{out}, d_{i,s,e}, E_i, H\right) \quad (2)$$

$s.t.$

$$\sum_{k \in N_{Pr_j}} W_{i,k,e} \leq Y_{i,j,e}, \forall j \in N_E, e \in E \quad (2a)$$

$$W_{i,k,e} B_{Min_{k,j}} \leq B_{i,k,j,e} \leq W_{i,k,e} B_{Max_{k,j}}, \forall k \in N_{Pr}, j \in N_{E_k}, e \in E \quad (2b)$$

$$ST_{i,s,e} = ST_{i,s,e-1} - d_{i,s,e} + \sum_{k \in N_{Pr_s}} \sum_{j \in N_{E_k}} \left(\rho_{i,k,s}^{Out} B_{i,k,j,e-1} - \rho_{i,k,s}^{In} B_{i,k,j,e}\right), \quad \forall s \in S, e \in E \quad (2c)$$

$$ST_{i,s,e} \leq O_s, \forall s \in S, e \in E \quad (2d)$$

$$Q_{i,j,e} = Q_{i,j,e-1} + \sum_{k \in N_{Pr_j}} B_{i,k,j,e} - \sum_{k \in N_{Pr_j}} \sum_{s \in S_k^{out}} \rho_{i,k,s}^{out} B_{i,k,j,e-1}, \quad (2e)$$

$$Q_{i,j,e_f} = 0, \forall j \in N_E, e \in E$$

$$\sum_{e \in E} d_{i,s,e} \geq r_s \ \ \forall s \in S \quad (2f)$$

$$T_{i,k,j,e}^F = T_{i,k,j,e}^S + \tau_{i,k} W_{i,k,e}, \forall k \in N_{Pr_j}, j \in N_E, \ e \in E \quad (2g)$$

$$T_{i,k,j,e}^F \geq T_{i,k,j,e}^S, \forall k \in N_{Pr_j}, j \in N_E, e \in E \quad (2h)$$

$$T_{i,k,j,e+1}^S \geq T_{i,k,j,e}^F, \forall k \in N_{Pr_j}, j \in N_E, e \in E \quad (2i)$$

$$T_{i,k,j,e+1}^F \geq T_{i,k,j,e}^F, \forall k \in N_{Pr_j}, j \in N_E, e \in E \quad (2j)$$

$$T_{i,k,j,e+1}^S \geq T_{i,k,j,e}^S, \forall k \in N_{Pr_j}, j \in N_E, e \in E \quad (2k)$$

$$T_{i,k,j,e}^F \leq H, \forall k \in N_{Pr_j}, j \in N_E, e \in E \quad (2l)$$

$$T_{i,k,j,e}^S \leq H, \forall k \in N_{Pr_j}, j \in N_E, e \in E \quad (2m)$$

$$T_{i,k,j,e_f}^F = H, \forall k \in N_{Pr_j}, j \in N_E, e \in E \quad (2n)$$

$$T_{i,k,j,e+1}^S \geq T_{i,l,j,e}^F - H(1 - W_{i,l,e}) \quad \forall j \in N_E, k \in N_{Pr_j}, l \in N_{Pr_j}, k \neq l, e \in E, e \neq e_f \quad (2o)$$

$where$

$c_{i,k,j} \in C_i, \forall k \in N_{Pr_j}, j \in N_E$

$p_s \in P, r_s \in R, \forall s \in S$

$\rho_{i,k,s} \in \rho_i, \forall k \in N_{Pr}, s \in S$

In Problem 2 the objective is to maximize the profit from producing a variety of products by implementing their recipes. Note that other objective functions can be implemented. $c_{i,k,j}$ represents the cost of task k in unit j at the i^{th} iteration, P_s is the sale price of state s, E is the number of event points. Eq. (2a) represents the allocation constraints where $W_{i,k,e}$ is set to zero unless task k at event e is taking place in unit j at the i^{th} iteration, also, this constraint depends on the value of $Y_{i,j,e}$, which represents the assignment of unit j at event e at the i^{th} iteration (with a value of 1). Eq. (2b) represents the capacity constraint, where $B_{i,k,j,e}$ is the material holdup of unit j while processing task k at event e.. Eq. (2c) represents the state material balances. $ST_{i,s,e}$ is the total quantity of state s at event e whereas $\rho_{i,k,s}^{In}$ and $\rho_{i,k,s}^{Out}$ represent the proportion of state s that is consumed or produced by task k, respectively. Eq. (2d) ensures that the quantity produced by each state is not greater than its capacity O_s. Eq. (2e) is the material balance inside each unit j with the condition that at the end of the last event e_f, must be equal to zero as all units must be emptied (no remaining material inside the processing units is allowed). Eq. (2f) represents the market demand constraints, where r_s is the total demand of state s and $d_{i,s,e}$ is the demand satisfied of such state at event e. r_s is the market demand that has to be satisfied according to the decisions made at the planning level. Eq. (2g) to Eq. (2n) are time logic constraints, i.e., they represent the time at which a task k starts to take place at unit j and in which event e takes place; they also enforce that no unit can be operated beyond the time horizon H. Eq. (2o) ensures the consecutiveness in the occurrence of two tasks sharing the same unit. Note that in problem (2), the parameters in the sets C_i and ρ_i change at each iteration i in accordance with the information gathered from the previous iteration(s).

III.2.3. Dynamic Cost Optimization Problem

The set of scheduling decisions β_{Sch_i} ($W_{i,k,e}$ and $Y_{i,j,e}$) can be obtained from the solution of Problem (2) ($Z_{Sch_i}^*$). This represents the key inputs required for the formulation of the dynamic cost optimization problem. In particular, the nonlinear dynamic optimization Problem 3 aims to find the control actions $u_{i,k,j,t}$ for each unit j at time t in the i^{th} iteration that maximize the profit while the order of operation is specified by the scheduling decisions β_{Sch_i} found from Problem 2. The formulation of this problem is as follows:

$$\max_{u_{i,k,j,t}} Z_{Dyn}\left(x(t)_{i,k,j}, \dot{x}(t)_{i,k,j}, u(t)_{i,k,j}, \Psi, t, \tau_{i,k,j}, C_i, \beta_{Sch_{i,k,j}}\right) \tag{3}$$

$s.t.$

$$f_p\left(x(t)_{i,k,j}, \dot{x}(t)_{i,k,j}, u(t)_{i,k,j}, \psi_{Nom}, t, \tau_{i,k,j}, \beta_{Sch_{i,k,j}}\right) = 0, \quad \forall\, t, p \in N_p, k \in N_{Pr}, j \in N_E \tag{3a}$$

$$h_q\left(x(t)_{i,k,j}, \dot{x}(t)_{i,k,j}, u(t)_{i,k,j}, \psi_{Nom}, t, \tau_{i,k,j}, \beta_{Sch_{i,k,j}}\right) \pm \lambda b_{i-1,k,j,q,t} \leq 0, \quad \forall\, t, q \in N_q, k \in N_{Pr}, j \in N_E \tag{3b}$$

where

$\psi_{Nom} \in \Psi$, $\tau_{i,k,j} \in \tau_i$, $\tau_i \in \mathrm{T}$

In Problem 3, the back-off terms $\lambda b_{i-1,k,j,q,t}$ represent the deviation of system from the nominal point under realizations in uncertain model parameters at the inequality constraint q of unit j performing task k at time t at the i^{th} iteration. More details about the back-off terms are provided in section III.2.5. Eq. (3a) represents the mechanistic model of the processes performed by the units of the plant. Eq. (3b) represents the constraints of the model. Consideration of the back-off terms in Eq. (3b) forces the system to find control decisions that will account for this back-off while ensuring dynamic feasibility. Note that back-off terms at the initial iteration ($i = 0$) have a value equal to zero. As shown in Eq. (3b), the back-off terms are preceded by a multiplier-factor λ, which can be thought as the level confidence given to each of the back-off terms. To obtain a coherent solution between the scheduling and control layer, both problems should consider objective functions with similar purposes (i.e., increase the overall profit).

III.2.4. Stochastic Simulations

By solving Problem 3 the optimal control profiles ($u_{Dyn_{i,k,j}}$) are obtained under the effects of back-off terms identified from the previous iteration. Hence, such control actions are not guaranteed to remain feasible under uncertainty.

In this step, the resilience of $u_{Dyn_{i,k,j}}$ to drive the system to feasible solutions is evaluated under a set of n stochastic realizations in the uncertain parameters ($\psi_{Unc_{i,n}}$), i.e., the value of ψ_{Unc} changes at each realization n. The realization in ψ_{Unc} are taken from their corresponding probability function (PDF) for each uncertain parameter and specified in the initialization step. This step is needed to generate statistical data that will allow for the recalculation of back-off terms at the current iteration. The problem under consideration is as follows:

$$max\ \zeta_{i,k,j,q,t,n} = h_q\left(x(t)_{i,k,j}, \dot{x}(t)_{i,k,j}, u_{Dyn_{i,k,j}}, \psi_{Unc_{i,n}}, t, \tau_{i,k,j}\right) \quad (4)$$

$$s.t.$$

$$f_p\left(x(t)_{i,k,j,n,m}, \dot{x}(t)_{i,k,j,n,m}, u_{Dyn_{i,k,j}}, \psi_{Unc_{i,n}}, t, \tau_{i,k,j}\right) = 0, \quad (4a)$$

$$\forall\ t, p \in N_p, k \in N_{Pr}, j \in N_E$$

where

$$\psi_{Unc_{i,n}} \in \Psi,\ \ \tau_{i,k,j} \in \tau_i,\ \ \tau_i \in \mathrm{T}$$

The considerations required for Problem 4 are as follows: i) only the mechanistic models (f_{pn}) are enforced; ii) control variables ($u_{Dyn_{i,k,j}}$) remain fixated; iii) only the feasibility of the system is assessed under different realizations in the uncertain parameters; iv) each uncertain parameter in $\psi_{Unc_{i,n}}$ is described by a PDF, thus, the values used while solving Problem 4 are selected from Monte Carlo (MC) sampling techniques. Problem 4 is solved a total of N_{MC} times, which is the total of stochastic realizations considered. N_{MC} is unknown a priori as Problem 4 must be solved in batches of n realizations of $\psi_{Unc_{i,n}}$ until the user-defined Criterion 1 (Eq. 5) is met.

Criterion 1 (Eq. 5) quantifies the deviations in the back-off term between the m^{th} and $m-1^{th}$ data populations. Note that each population m is composed of $n * m$ realizations in the uncertain parameters. This criterion makes this step repeat until the errors between the back-off terms in the actual data set (i.e., population m) and the previous data collected (i.e., population $m-1$) is below a user defined tolerance. Note that population m includes the information of previous populations and that N_{MC} is equal to the number of data points in the last m^{th} population. Also, note that Criterion (5) can only be enforced when $m \geq 2$. The specific procedure to estimate the back-off terms shown in Criterion 1 (Eq. 5) is presented in the next section (III.2.5).

$$\left|1 - b_{i,k,j,q,t,m-1}/b_{i,k,j,q,t,m}\right| \leq Tol_{SS}\ \forall\ \mathrm{t}, k \in N_{Pr}, \mathrm{j} \in N_E, \mathrm{q} \in N_q \tag{5}$$

III.2.5. Back-Off Term Calculation

A back-off term ($b_{i,k,j,q,t}$) is the representation of the deviation in the q^{th} constraint function ($\zeta_{i,k,j,q,t,n}$) at a time t for a unit j performing task k at the i^{th} iteration for each stochastic simulation n in the uncertain parameters. The introduction of $b_{i,k,j,q,t}$ into constraints h_q is to search for a solution capable of accommodating uncertainty, backed off from the optimal solution at nominal model parameters[18]. Note that there may be a back-off term for each inequality q at each time point t, i.e.:

$$b_{i,k,j,q,t} = \sqrt{\frac{1}{N_{MC}}\sum_{n=1}^{N_{MC}}\left[\zeta_{i,k,j,q,t,n} - \frac{1}{N_{MC}}\sum_{n=1}^{N_{MC}}\zeta_{i,k,j,q,t,n}\right]^2} \quad \forall\ t, k \in N_{Pr}, j \in N_{E_{i,q}}, q \in N_{q_i} \tag{6}$$

Eq. (6) represents the calculation of the normal standard deviation for discrete random variables, which is used for the calculation of the back-off terms in this work. Eq. (6) may vary correspondingly to the statistical distribution of $\zeta_{i,k,j,q,t,n}$. Note that it is Eq. (6) that is used to calculate the back-off terms at each m population in the procedure described in section III.2.4. Also, note that the iteration index i has been updated in Eq. (6) to indicate that these back-off terms have been updated and may be used in subsequent calculations. In this work, it is assumed that the data used for the back-off terms calculation, $\zeta_{i,k,j,q,t,n}$, follows a statistical distribution that can be approximated to a normal distribution.

III.2.6. Unit Operation Time Optimization Problem

Since the back-off terms represent the degree of variability that the constraints need to accommodate using an optimal control profile, in this step, such variability is used to determine the optimal operation times for each unit j. Problem 7 aims to determine the minimum time required for each scheduled unit operation to achieve their corresponding production goals under parameter uncertainty, which is expressed through the back-off terms.

$$\min_{\tau_{Dyn_{i+1,k,j}}} Z_{UOT}\left(x(t)_{i,k,j}, \dot{x}(t)_{i,k,j}, u(t)_{i,k,j}, \Psi, t, \tau_{Dyn_{i+1,k,j}}, C_i\right) \quad (7)$$

$s.t.$

$$f_p\left(x(t)_{i,k,j}, \dot{x}(t)_{i,k,j}, u(t)_{i,k,j}, \psi_{Nom}, t, \tau_{Dyn_{i+1,k,j}}\right) = 0,\ \forall\ t, p \in N_{p_i} \quad (7a)$$

$$h_q\left(x(t)_{i,k,j}, \dot{x}(t)_{i,k,j}, u(t)_{i,k,j}, \psi_{Nom}, t, \tau_{Dyn_{i+1,k,j}}\right) \pm \lambda b_{i+1,k,j,q,t} \leq 0, \quad (7b)$$
$$\forall\ t,\ \ j \in N_{E_i},\ \ q \in N_{q_i}$$

$$\tau_{Min_j} \leq \tau_{Dyn_{i+1,k,j}} \leq \tau_{Max_j},\ \tau_{Dyn_{i+1,k,j}} \in \tau_{i+1},\ \forall\ \ k \in N_{Pr}, j \in N_{E_i}$$

$where$

$$\psi_{Nom} \in \Psi,\ \tau_{i+1} \in \mathrm{T}$$

From Problem 7, τ_{i+1} is the set of all the unit operation times under model parameter uncertainty. Once the optimization of the unit operation time has been performed, the variation between the back of terms used in the i^{th} iteration ($b_{i,k,j,q,t}$), and those calculated in the current iteration ($b_{i+1,k,j,q,t}$) is used as Criterion 2 (Eq. 8) which terminates the algorithm if such deviation is lesser than a user-defined tolerance (Tol_{BO}); otherwise, the algorithm proceeds with the next iteration, as shown in **Figure 4**. Criterion 2 is defined as follows:

$$\left|1 - b_{i-1,k,j,q,t}/b_{i,k,j,q,t}\right| \leq Tol_{BO}, \forall\ \mathrm{t}, k \in N_{Pr}, \mathrm{j} \in N_E, \mathrm{q} \in N_q \quad (8)$$

If Criterion 2 (Eq. 8) is met, a solution Z^* that is composed by the control decisions u_{Dyn}, unit operation times τ_{Dyn} and process scheduling decisions β_{Sch}, that accommodates the variability (λ) in the uncertain model parameters (ψ_{Unc}), has been found.

III.3. Case Study

The performance of the methodology presented in the previous section has been evaluated using a case study adapted from the literature that involves a chemical batch plant[20]. As shown in **Figure 5**, the chemical batch plant is composed by 4 batch processes: a first set of chemical reactions (RI), a filtration process (FI), a second set of chemical reactions (RII) and a separation process (SI), hence, $N_{Pr} = \{RI, RII, FI, SI\}$ and they follow a sequential path. The plant is allowed to operate in a time horizon (H) set to 13 h.

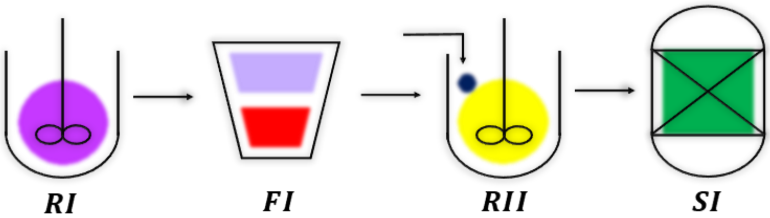

Figure 5. Case Study – Process Scheme.

The process model consists of substance A transforming into an intermediate product B in process RI, taking place in a jacketed non-isothermal batch reactor where the temperature is controlled by valves regulating the flow of the cold or hot auxiliary services passing through the jacket. The resulting mixture, which contains substance B, is filtered in process FI to obtain pure substance B. The intermediate species B is then stored inside the semi-batch reactor where process RII will take place. RII initiates when substance D is fed in a controlled fashion to regulate the production of species E and F. The mix containing species B and E (where E is the desired product) is separated from the mix of species D and F in the separation process SI. Processes FI and SI are stationary and assumed to achieve a perfect separation of substances within a fixed operation time defined *a priori*. $F1$ filters perfectly substance B, which is fed pure to process RII. SI is capable of perfectly separating E from the mixture exiting RII. On the other hand, RI and RII are time dependent dynamic processes; thus, the operation times for these units (τ_{Dyn}^{RI} & τ_{Dyn}^{RII}) need to be obtained from optimization. The present model assumes that there are storage units between each batch process that can hold the outgoing materials until the following unit that will process such material is available.

The Set of Equations (9), which describe the dynamic process RI are reported in **Table 1**.

Table 1. Process Dynamics of set of reactions ***RI***

Chemical Reaction(s)	(9)
$$A \overset{r_1}{\rightarrow} B \overset{r_2}{\rightarrow} C$$	
Mass Balance(s)	
$$\frac{dC_A}{dt} = -r_1, \frac{dC_B}{dt} = r_1 - r_2$$	
Energy Balance(s): Reactor	
$$\frac{dT_R}{dt} = -\frac{\Delta H_1 r_1 + \Delta H_2 r_2}{\rho_R c_R} + \frac{UA_j(T_J - T_R)}{V_R \rho_R c_R}$$	
Energy Balance(s): Reactor's Jacket	
$$\frac{dT_J}{dt} = \frac{F_{Hot}(T_{Hot} - T_J)}{V_J} + \frac{F_{Cold}(T_{Cold} - T_J)}{V_J} + \frac{UA_j(T_R - T_J)}{V_R \rho_R c_R}$$	
Rate(s) of Reaction	
$$r_1 = k_1 e^{-E_1/T_R} C_A, r_2 = k_2 e^{-E_2/T_R} C_B$$	
Innitial Condition(s)	
$$C_A(0) = C_{A_0}, C_B(0) = C_{B_0}, T_R(0) = T_{R_0}, T_J(0) = T_{J_0}$$	
Constraint(s)	
$$C_B(t_f) \geq C_{B_{Fix}}, T_R(t_f) \leq T_{R_{Fix}}$$	
Manipulated Variable(s)	
$$u(t) = [F_{Hot} \;\; F_{Cold}]$$	

Similarly, the Set of Equations (10), corresponding to process RII, is reported in **Table 2.**

Table 2. Process Dynamics of set of reactions ***RII***

Chemical Reaction(s)	(10)
$$B + D \overset{r_3}{\rightarrow} E, 2D \overset{r_4}{\rightarrow} F$$	
Mass Balance(s)	
$$\frac{dC_B}{dt} = -r_3 - \frac{F_{Feed}}{V_R} C_B$$ $$\frac{dC_D}{dt} = -r_3 - 2r_4 + \frac{F_{Feed}}{V_R}\left(C_{D_{Feed}} - C_D\right)$$ $$\frac{dC_E}{dt} = r_3 - \frac{F_{Feed}}{V_R} C_E$$ $$\frac{dC_F}{dt} = r_3 - \frac{F_{Feed}}{V_R} C_F$$	
Volume Balance(s)	
$$\frac{dV_R}{dt} = F_{Feed}$$	
Rate(s) of Reaction	
$$r_3 = k_3 C_B C_D, r_4 = k_4 C_D^2$$	
Innitial Condition(s)	
$$V_R(0) = V_{R_0}, C_B(0) = C_{B_{RI}}, C_D(0) = C_{D_0}, C_E(0) = C_{E_0}, C_F(0) = C_{F_0}$$	
Constraint(s)	

$$V_R(t_f)C_E(t_f) \geq V_{Fix}C_{E_{Fix}}, V_R(t_f) \leq V_{R_{Fix}}, V_R(t_f)C_F(t_f) \leq V_{Fix}C_{F_{Fix}}$$

Manipulated Variable(s)

$$u(t) = [F_{Feed}]$$

On **Table 3** are listed the chemical reaction parameters and the process specifications and parameters for the reactors used for the processes RI and RII. For this case study it is assumed that k_{AB} and k_{BC} are the uncertain parameters in the system that follow a normal distribution with a standard deviation equivalent to 5% of their expected (nominal) values (reported in Table 3). Concentration values for species A, B, C, D, E and F are: A) For process RI, at the start of operation only species A is present with a concentration of 1 $kmol/m^3$ and the process stops when a concentration of 0.7 $kmol/m^3$ for species B is achieved. B) For process RII, at the start of operation only species B is present with a concentration of 1 $kmol/m^3$; species D is injected with a concentration of 10 $kmol/m^3$; species F is monitored to never surpass the 0.15 $kmol/m^3$ and the process operation finalizes when E reaches a concentration above 0.5 $kmol/m^3$.

Table 3. Case Study – Chemical reaction parameters and reactor specifications.

Variable Name	Value	Description
$E_{A\rightarrow B}$ $[K]$	4500	Activation Energy
$E_{B\rightarrow C}$ $[K]$	8250	Activation Energy
ΔH_{AB}^{RX} $[KJ/mol]$	-1E4	Enthalpy of Reaction in RI
ΔH_{BC}^{RX} $[KJ/mol]$	-1E6	Enthalpy of Reaction in RI
k_{AB} $[h^{-1}]$	1E7	Kinetic Velocity in RI
k_{BC} $[h^{-1}]$	1E10	Kinetic Velocity in RI
k_{BDE} $[m^3/Kmol/h]$	2	Kinetic Velocity in RII
k_{2DF} $[m^3/Kmol/h]$	1	Kinetic Velocity in RII
$T_{RI,Initial}^{Reactor}$ $[K]$	294	Reactor's Temperature
$T_{RI,Initial}^{Jacket}$ $[K]$	294	Jacket's Temperature
$T_{RI,Final}^{Reactor}$ $[K]$	300	Reactor's Temperature
$T_{Cold}^{Aux.Serv.}[K]$	294	Aux. Serv. Temperature
$T_{Hot}^{Aux.Serv.}[K]$	350	Aux. Serv. Temperature
$\dot{v}_{Hot}^{Aux.Serv.}[m^3/h]$	15	Aux. Serv. Flow
$\dot{v}_{Cold}^{Aux.Serv.}[m^3/h]$	15	Aux. Serv. Flow
$\dot{v}_{RII,D}^{Feed.}[m^3/h]$	5	Species D Feed Flow
$V_{RII,Max}^{Reactor}[m^3]$	10	Reactor's Volume
$\rho_{RI}^{Reactor}[kg/m^3]$	1E3	Reactor's Density

$\rho_{RI}^{Jacket}[kg/m^3]$	1E3	Jacket's Density
$Cp_{RI}^{Reactor}[kJ/kg/K]$	2.5	Reactor's Heat Capacity
$Cp_{RI}^{Jacket}[kJ/kg/K]$	2.5	Jacket's Heat Capacity
$UA_{RI}[kJ/m^2/K/h]$	8E4	Heat Transfer Coeff.

Table 4 presents the costs and prices of the material incomes and outcomes and the auxiliary services. Note that for the scheduling problem, the associated costs and prices are defined from data obtained from sensitivity analysis of the operation of the plant.

Table 4. Case Study – Prices and Costs of species, mixtures, and auxiliary services.

Variable Name	Value	Description
$p_{F_{Cool}}\ [m.u/m^3]$	0.5	Aux. Serv. Price
$p_{F_{Hot}}\ [m.u/m^3]$	0.5	Aux. Serv. Price
$p_{F_{Feed}}[m.u./kmol]$	50	Raw Feed Price
$p_A[m.u./kmol]$	100	Raw Material Price
$p_{Waste}[m.u./kmol]$	1000	Waste Cost
$p_{Pure}[m.u./kmol]$	10000	Product Price
$p_{Impure}[m.u./kmol]$	1000	Sub-product Cost

The profit function for Problem 3, i.e., the dynamic optimization problem, is as follows:

$$Z_{Dyn} = -c_{RI} - c_{A_{RI}} - c_{FI} - c_{RII} + c_{SI_1} - c_{SI_2} \tag{9}$$

$$c_{RI} = \sum_j^{N_{E_{RI}}} \left(p_{F_{Hot}} \int_0^{\tau_{Dyn_{i,j}}^{RI}} F_{Hot_j}(t)dt + p_{F_{Cool}} \int_0^{\tau_{Dyn_{i,j}}^{RI}} F_{Cool_j}(t)dt \right) \tag{10}$$

$$c_{A_{RI}} = p_A \sum_j^{N_{E_{RI}}} n_{A_{i,j}}(t)|_{t=0} \tag{11}$$

$$c_{FI} = p_{Waste} \sum_j^{N_{E_{FI}}} \left(n_{A_{i,j}}(t) + n_{C_{i,j}}(t) \right)|_{t=\tau_{Dyn_{i,j}}^{RI}} \tag{12}$$

$$c_{RII} = p_{F_{Feed}} \sum_j^{N_{E_{RII}}} \int_0^{\tau_{Dyn_{i,j}}^{RII}} V_{R_{i,j}}(t) C_{F_{i,j}}(t)dt \tag{13}$$

$$c_{SI_1} = p_{Pure} \sum_j^{N_{E_{RII}}} n_{E_{i,j}}(t)|_{t=\tau_{Dyn_{i,j}}^{RII}} \tag{14}$$

$$c_{SI_2} = p_{Impure} \sum_j^{N_{E_{RII}}} \left(n_{D_{i,j}}(t) + n_{F_{i,j}}(t) \right)|_{t=\tau_{Dyn_{i,j}}^{RII}} \tag{15}$$

where Eq. (9) involves the cost of auxiliary services for RI (Eq. 10), the cost for raw species A at the beginning of process RI (Eq. 11), the cost for the waste generated in FI (Eq. 12), the cost for species D fed into process RII (Eq. 13), the revenues for selling pure E (Eq. 14) and the cost for having unreacted D and generated F in the product (Eq. 15).

The profit function used in the scheduling Problem 2 is:

$$Z_{Sch} = c_{ST} - c_{Op} \tag{16}$$
$$c_{Op} = \sum_{e}^{E} \sum_{k}^{N_{Pr}} \sum_{j}^{N_{E_k}} W_{k,e} \tau_j c_{k,j} \tag{17}$$
$$c_{ST} = \sum_{e}^{E} \sum_{s}^{S} d_{s,e}\, p_s \tag{18}$$

Eq. (16) involves the costs associated with the operation ($c_{k,j}$) of unit j to realize task k at event e (Eq. 17) and the cost or revenue caused by the consumption or generation of the state s (Eq. 18), i.e., $p_s < 0$ for costs and $p_s > 0$ for revenues). Additional details can be found elsewhere[20].

The objective of this case study is to identify a scheduling and control strategy that can maximize the batch plant profits in the presence of stochastic uncertainty in the model parameters.

III.4. Results

The case study was implemented using Pyomo optimization suite within Python 3.7. The Interior Point algorithm IPOPT™ with ma57 HSL linear solver was used to solve Problems 3, Problem 4 and Problem 7. CPLEX was used to solve the MILP Problem 2. The model was solved in a PC with an Intel® Core™ i7-8700 CPU @ 3.2 GHz and 16 GB of RAM. The Differential-Algebraic Equations (DAE)[72] module from Pyomo was utilized for the discretization of the differential equations. Orthogonal collocation on finite elements was chosen as the discretization method. For the present case study, 40 finite elements and 3 collocation points for the problems solved with ipopt were considered. These parameter values returned acceptable solutions in an average CPU time of ~1 s per simulation.

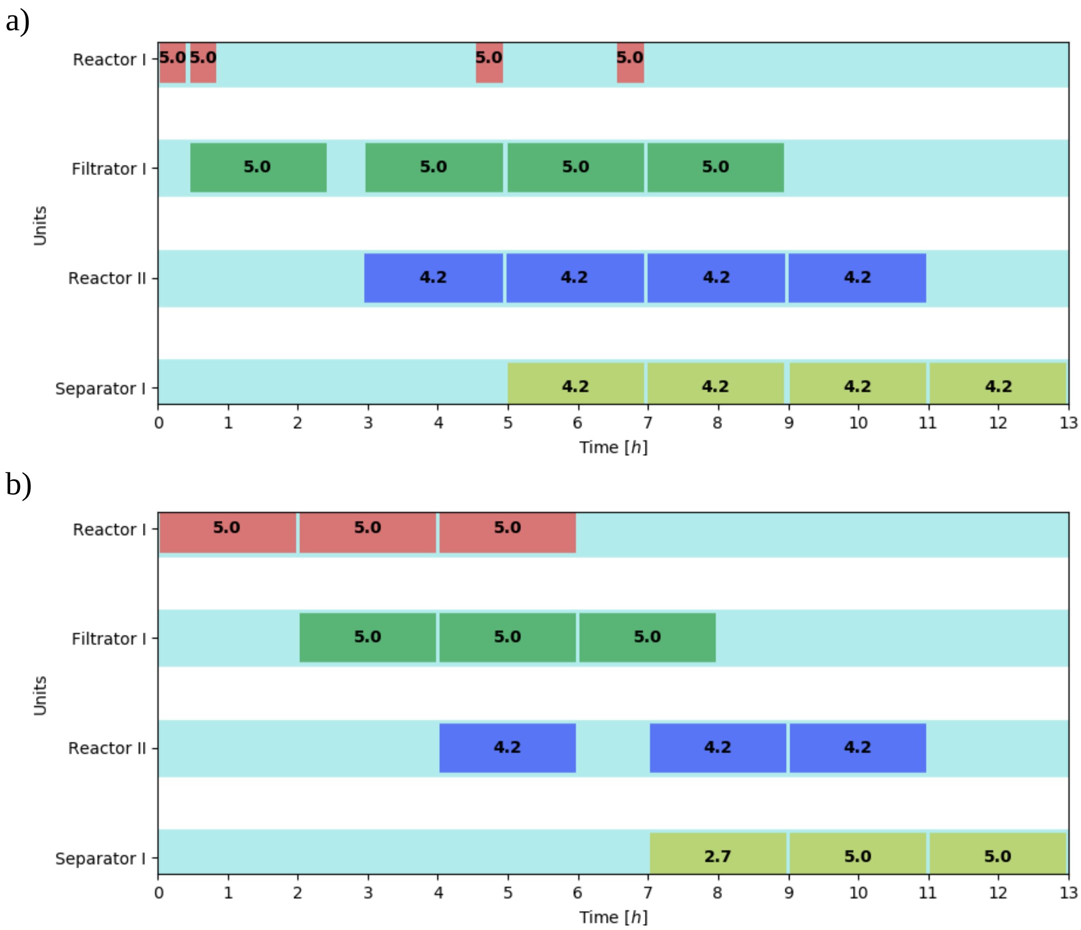

Figure 6. Algorithm A – Scheduling Plans:
(a) Algorithm A ($\lambda = 2$), **(b)** Valdez-Navarro's algorithm ($\lambda = 2$). Each unit allocation is portrayed with their corresponding material processing quantity in m^3.

For comparison purposes, the present case study was solved using the Algorithm A and that proposed by Valdez-Navarro and Ricardez-Sandoval[1] where unit operations times remain fixed during the calculations. In this scenario the effects of varying the unit operation times were studied. Uncertainty was considered only to be present on the kinetic velocities of the first set of reactions (RI). Note that Valdez-Navarro and Ricardez-Sandoval[1] showed the advantages and limitations of performing an integrated approach against a non-iterative sequential approach for scheduling and control.

As shown in **Figure 6**, the proposed algorithm can accommodate another batch sequence by finding new unit operations times that increase plant production (material processing batch size is depicted inside each unit assignment in the figure). While all the unit operation times were set to 2h for Valdez-Navarro's algorithm (**Figure 6b**), it can be observed that the values τ_{Dyn}^{RI} and τ_{Dyn}^{RII} for the present approach (**Figure 6a**) are 0.436h and 2.017h.

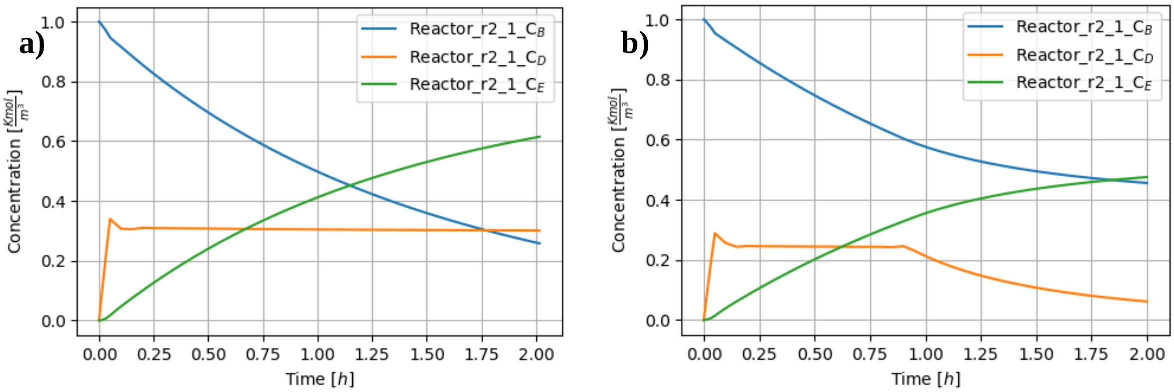

Figure 7. Algorithm A – Concentration profiles of species B, D and E for RII. **(a)** Algorithm A ($\lambda = 2$), **(b)** Valdez-Navarro's algorithm ($\lambda = 2$). In each graph's legend, r2 represents task RII. Only the operation of one unit is shown, as the rest are identical.

The operation regimes found by Algorithm A are more economically expensive, as more control actions are required to maintain the dynamic feasibility (the control profiles are not shown for brevity). Nevertheless, as it can be seen in **Figure 6a**, due to the addition of another job and the increased production in task RII (as shown in **Figure 7a**, compared to **Figure 7b**, where the concentration of E is higher at the end of operation) the profits grow a 42% with respect to the case where no unit time operation is performed (a total of 90,386 m.u. of profit for Algorithm A compared to 63,369 m.u. of profit obtained from Valdez-Navarro's algorithm). The increase in production of species E shown in **Figure 7a**, is due to the back-off effect caused by the methodology in the model constraints to mitigate the effects of the uncertainty in the kinetic velocities k_{AB} and k_{BC}.

Regarding the CPU times, each iteration of the present algorithm takes on average 8 h, with 3 iterations required to solve the present case study. On the other hand, 2 iterations and 4 h per iteration were required by Valdez-Navarro's algorithm. Note that the selection of the tolerance parameters (Tol_{SS} & Tol_{BO}) is problem-specific and will impact the algorithm's computational effort. In this work, both Tol_{SS} & Tol_{BO} were set to 0.0025, which seemed to be adequate as they returned acceptable results in reasonable CPU times.

III.5. Summary

This chapter presented the back-off decomposition algorithm A for the integration of scheduling and control. The proposed methodology is an iterative approach developed to address the computational burden that represents the full resolution of a MIDO problem. In this methodology, parametric uncertainty is approximated through Monte Carlo sampling techniques. The effects of the uncertainty on the system are analyzed and used to calculate back-off terms. Back-off terms are introduced in the formulation to force the system find control profiles, unit operation times and scheduling decisions that significantly diverge from those of an economically attractive but dynamically infeasible point under the presence of uncertainty. A case study illustrating the performance of the algorithm is presented. The results show that the methodology can generate a new operation policy that deals with parametric uncertainty in the model with a degree of probability satisfaction (level of confidence) that is defined by the user through the PDFs and the convergence tolerance criterion considered in the algorithm. Nonetheless, Algorithm A presents some limitations in its calculations. The unit operation times are optimized in a sequential fashion once the control decisions have been estimated from the dynamic cost optimization problem (section III.2.3). Since the unit operation times and control decisions are not optimized sequentially, a better solution may exist if both decisions are explored simultaneously. Also, the scheduling decisions present slight discrepancies in the estimations of key production metrics (e.g., total production, total costs), when compared to the estimations obtained from the dynamic optimization problems. This issue is due to the inability of the current scheduling formulation to estimate certain parameters that are dependent on the quantity of material being processed (unit operation times, unit operation costs, material ratios). The algorithm presented in the next chapter aims to address the issues mentioned above with Algorithm A.

Chapter IV. Back-off Decomposition Algorithm B

In this chapter, the Back-off Decomposition Algorithm B is presented. This algorithm was developed to address the issues of solving the full implementation of the MIDO Problem 1 (section III.1) and the limitations of Algorithm A discussed in the summary of the last chapter (section III.5). One of the key differences between Algorithm A and B is that the optimization of unit operation times and control profiles is performed simultaneously in the latter. Back-off terms are added into the constraints of the dynamic system and utilized to drive the system to a new dynamically feasible and attractive economic solution that can accommodate stochastic parametric uncertainty. Another key difference is the introduction of correlations into the scheduling formulation that allow for a dynamic calculation of unit operation times, unit operation costs and better estimations for the material transfer ratios between units. The scheduling plan is thus expected to offer more accurate estimations for the dynamic operation of the process, i.e., it captures better the dynamic operation of the process. The case study and instances used to analyse the performance of the algorithm are presented and their results discussed in this chapter. The information presented in this chapter has been submitted for publication and is currently under review[3].

IV.1. Methodology

In the back-off decomposition Algorithm B, the MIDO Problem 1 (section III.1) is decomposed and reformulated as follows: (1) A parametric sensitivity analysis that generates the correlations needed to construct the scheduling problem. The correlations are used to capture the effects of the back-off terms on the dynamics of the process into a formulation that the scheduling problem can accept. (2) A scheduling problem, performed to obtain scheduling decisions for the plant and which can determine unit operation times and material transfer between units under the consideration of stochastic events (accounted for through the correlations developed in step 1). (3) A dynamic optimization problem, which determines the optimal control regimes and unit operation times of the process in the presence of the back-off terms imposed on the process constraints. (4) Monte Carlo simulations, which are used to generate sufficient statistical data to update the calculation of the back-off terms. The set of problems is solved in an iterative fashion until an optimal and feasible scheduling and control solution that can accommodate the uncertainty in the system is identified. The graphical representation of the algorithm is shown in **Figure 8**.

In the new proposed algorithmic framework both unit operation times and control decisions are now are obtained from a single dynamic optimization formulation, which is an innovation with respect to previous studies published in the literature[1,2,13]. The simultaneous approach proposed here for the dynamic optimization of the process allows the system to search for unit operation times and control actions that can accommodate stochastic parametric uncertainty, while improving the economics of the system. In addition, the effects of the back-off terms were not taken into consideration by the scheduling decisions in previous methods. To make our proposed approach more effective, a parametric sensitivity analysis is performed to generate correlations, which are key inputs for the scheduling problem to determine the sensitivity of the back-off terms to the scheduling decisions. Each of the steps considered in the present back-off framework is described next.

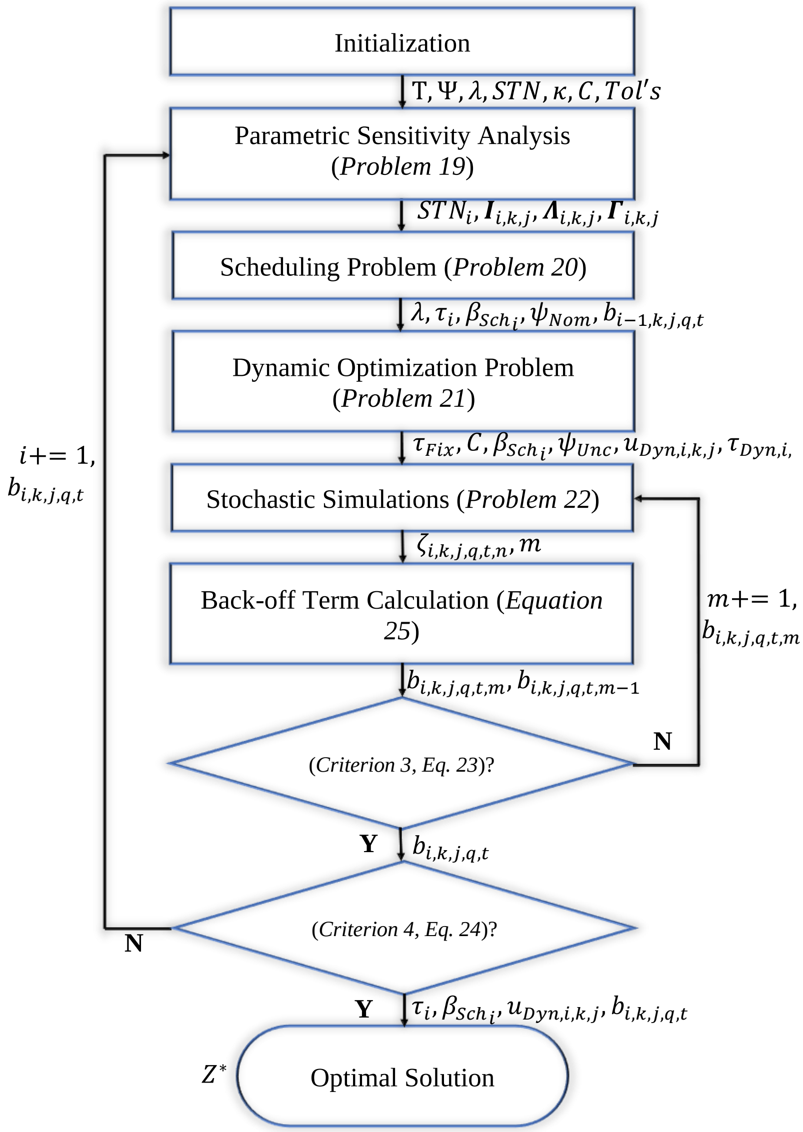

Figure 8. Back-off Decomposition Algorithm B Diagram.

IV.1.1. Initialization

As shown in **Figure 8**, the algorithm starts with the definition of the sets T (Unit Operation Times set) and Ψ (Parameters set). The probabilistic distribution functions (PDF) and their corresponding parameters (η) that describes each parameter included in ψ_{Unc} must be defined, i.e., for an uncertain parameter w: $\psi_{Unc_w} = PDF_w(\eta_w)$. Also, parameters needed for the specification of the State-Task Network (STN) must be provided (i.e., tasks, states and their characteristics). Similarly, basic parameters used in the formulation of the scheduling problem (i.e., E, H, P) also need to be defined. Moreover, the tolerance (stopping) criteria used in the framework must be defined, i.e., Tol_{SS} for Criterion 3 (Eq. 23) in section IV.1.5 & Tol_{BO}, for Criterion 4 (Eq. 24) in section IV.1.6. Also, fixed costs (C) should be defined by the user. The iteration index for the algorithm is initialized, i.e., $i = 0$. Furthermore, the set κ used in the parametric sensitivity analysis must be defined *a priori*. This set provides the range of operating conditions and processing material quantities that can be used as degrees of freedom to improve the dynamic operation of each task assigned to a processing unit.

IV.1.2. Parametric Sensitivity Analysis

The parametric sensitivity analysis step is performed to generate and process the data of the operation of each unit individually for the set of operating conditions specified in the set κ. This analysis is performed under consideration of stochastic parametric uncertainty, which is represented through the back-off terms. This step enables the construction of explicit correlations that are used in the scheduling formulation at each i^{th} iteration. Accounting for the explicit relations between the set κ and the correlated variables, i.e., unit operation costs, unit operation times and the incoming/outgoing material relations, would often result in nonlinear functions that may lead to MINLP formulations in the scheduling problem. To circumvent this issue, linear correlations are considered in this work to reduce the complexity of the scheduling formulation; these correlations capture the sensitivity of these variables under uncertainty for each unit considered in the scheduling problem.

Figure 9 presents a summary of the parametric sensitivity analysis and its relationship with the scheduling problem. As shown in this figure, a parametric sensitivity analysis is performed on each individual unit to generate data on the behavior of the unit operation times, unit operation costs and the variations in material inputs and outputs as a function of the set κ. Linear correlations are then constructed and used as input parameters to the scheduling formulation.

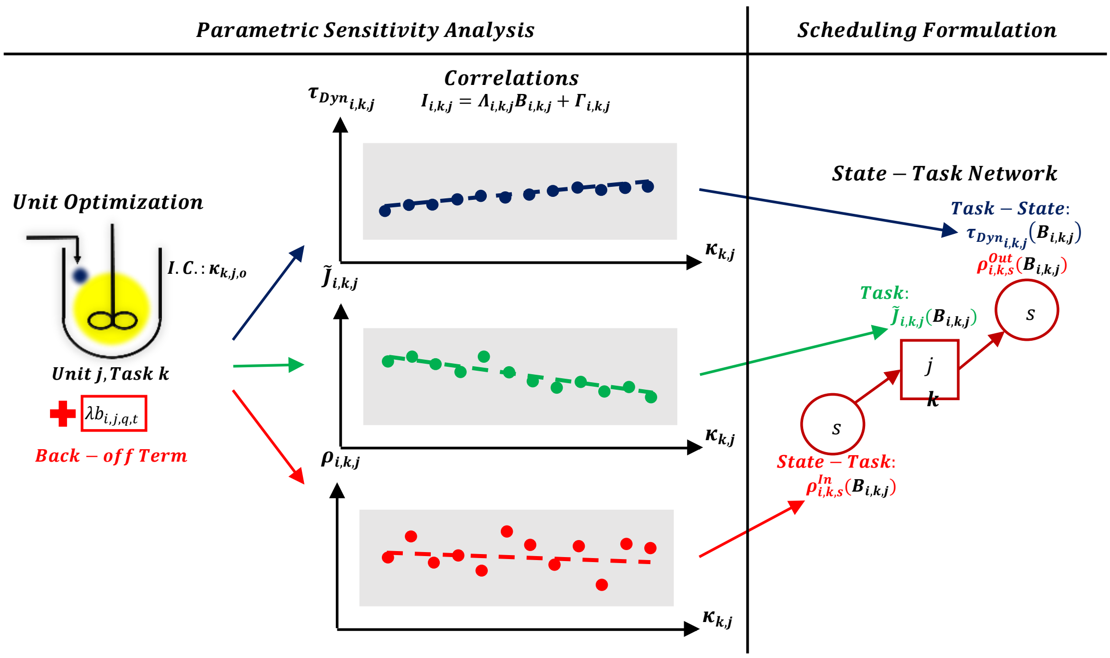

Figure 9. Parametric Sensitivity Analysis and Scheduling Relation Diagram.

As shown in **Figure 8**, the key inputs to perform the parametric sensitivity analysis step are the back-off terms ($\lambda b_{i,k,j,q,t}$), the set of nominal model parameters (ψ_{Nom}) and the mechanistic model of the each unit and process (task) of the chemical plant. In the case $i = 0$ (initial iteration step), the back-off terms ($\lambda b_{i,k,j,q,t}$) are set to 0, i.e., the system is analyzed at nominal conditions. Each parametric analysis involves the solution of a series of optimization problems that aim to find the optimal control actions and the operation times that minimize the operating costs of each unit j, while realizing task k, for the condition $\kappa_{k,j,o}$, where $\kappa_{k,j,o} \in \kappa_{k,j}$. The set κ should include values uniformly distributed between κ_{Max} and κ_{Min} with a user-selected number of elements. Note that the number of elements has a direct impact on the computational cost and the accuracy of the correlations. The optimization problem considered for this step is as follows:

$$\max_{u_{i,k,j,t,o},\tau_{Dyn_{i,k,j,o}}} Z_{SA}\left(x(t)_{i,k,j,o}, \dot{x}(t)_{i,k,j,o}, u(t)_{i,k,j,o}, \Psi, \kappa_{k,j,o}, t, \tau_{Dyn_{i,k,j,o}}, C, \tilde{J}_{i,k,j,o}\right) \quad (19)$$

$s.t.$

$$f_p\left(x(t)_{i,k,j,o}, \dot{x}(t)_{i,k,j,o}, u(t)_{i,k,j,o}, \kappa_{k,j,o}, \psi_{Nom}, t, \tau_{Dyn_{i,k,j,o}}\right) = 0, \quad \forall\, t, p \in N_p, k \in N_{Pr}, j \in N_E \quad (19a)$$

$$h_q\left(x(t)_{i,k,j,o}, \dot{x}(t)_{i,k,j,o}, u(t)_{i,k,j,o}, \kappa_{k,j,o}, \psi_{Nom}, t, \tau_{Dyn_{i,k,j,o}}\right) \pm \lambda b_{i-1,k,j,q,t} \le 0, \forall\, t, q \in N_q\,, k \in N_{Pr}, j \in N_E \quad (19b)$$

$$\rho^{In}_{i,k,j,o} = g^{In}\left(x(t)_{i,k,j,o}|_{t=0}\right) \forall\, k \in N_{Pr}, j \in N_E \quad (19c)$$

$$\rho^{out}_{i,k,j,o} = g^{Out}\left(x(t)_{i,k,j,o}|_{t=\tau_{Dyn_{i,k,j}}}\right) \forall\, k \in N_{Pr}, j \in N_E \quad (19d)$$

$$\tilde{J}_{i,k,j,o} = g^{Cost}\left(x(t)_{i,k,j,o}, u(t)_{i,k,j,o}, t, \tau_{Dyn_{i,k,j,o}}\right) \forall\, t, k \in N_{Pr}, j \in N_E \quad (19e)$$

$$\tau_{Min_{k,j}} \le \tau_{Dyn_{i,k,j,o}} \le \tau_{Max_{k,j}},\ \tau_{Dyn_{i,k,j,o}} \in \tau_i^*,\ \forall j \in N_{E_k}, k \in N_{Pr} \quad (19f)$$

where

$\psi_{Nom} \in \Psi, \tau_i^* \in \mathrm{T}, \kappa_{k,j,o} \in \kappa_{k,j}$

Note that problem 19 and problem 21 must share similar objectives, e.g., maximize profits or minimize makespan.

For each task (k) and unit (j), the result from each optimization problem describes a particular combination between the unit operation times ($\tau_{Dyn_{i,k,j,o}}$), unit operation costs ($\tilde{J}_{i,k,j,o}$) and the variations in material inputs ($\rho^{In}_{i,k,j,o}$) and outputs ($\rho^{out}_{i,k,j,o}$) with respect to a discrete realization in the set $\kappa_{k,j}$ (i.e., $\kappa_{k,j,o}$, where $\kappa_{k,j,o} \in \kappa_{k,j}$) under stochastic parametric uncertainty, which is represented through the back-off terms ($b_{i-1,k,j,q,t}$), as shown in Eq 191b). Note that $\kappa_{k,j,o}$ represents a user-defined realization in the range of values to be considered for the set κ for unit j and task k. Eq. (19a) represents the mechanistic process model. Eq. (19b) represents the process constraints of the unit, and they involve production goals, safety requirements, quality concerns and/or feasibility limitations. Note that Eq. (19b) enforces the system to find the control decisions and unit operation times backed-off from the nominal values. A more detailed discussion on the back-off terms and their calculation is provided in section IV.1.7. Eq. (19c) and Eq. (19d) represent the calculation of the relations of material that enters or leaves the unit, which is fractional (mass, volume, mol) and is user-specified. Eq. (19e) represents the calculations of the cost to operate the unit j to realize task k. Eq. (19f) represents the bounds on the processing unit times ($\tau_{Dyn_{i,k,j,o}}$). The data gathered from the solution of these optimization problems for each $\kappa_{k,j,o}$ in $\kappa_{k,j}$ (i.e., $\tau_{Dyn_{i,k,j,o}}$, $\tilde{J}_{i,k,j,o}$, $\rho^{In}_{i,k,j,o}$ and $\rho^{out}_{i,k,j,o}$) is used to identify linear correlations between these variables and $\kappa_{k,j}$. These correlations can be represented as follows:

$$\boldsymbol{I}_{i,k,j} = \boldsymbol{\Lambda}_{i,k,j}\kappa_{k,j} + \boldsymbol{\Gamma}_{i,k,j} \tag{19g}$$

where $\boldsymbol{I}_{i,k,j} = \left[\tilde{J}_{i,k,j}\ \tau_{Dyn_{i,k,j}}\ \rho^{In}_{i,k,j}\ \rho^{out}_{i,k,j}\right]^T$ is the vector of the variables of interest for each unit j performing task k; $\boldsymbol{\Lambda}_{i,k,j} = \left[\Lambda^{C}_{i,k,j}\ \Lambda^{\tau}_{i,k,j}\ \Lambda^{\rho^{In}}_{i,k,j}\ \Lambda^{\rho^{out}}_{i,k,j}\right]^T$ is the vector of the correlation slopes for each unit j realizing task k whereas $\boldsymbol{\Gamma}_{i,k,j} = \left[\Gamma^{C}_{i,k,j}\ \Gamma^{\tau}_{i,k,j}\ \Gamma^{\rho^{In}}_{i,k,j}\ \Gamma^{\rho^{out}}_{i,k,j}\right]^T$ is the vector of the correlation intercepts for each unit j realizing task k. Note that the relations of incoming and outgoing material in the scheduling problem are calculated at the midpoint between the maximum and minimal load, i.e., $\left(\kappa_{Min_{k,j}} + \kappa_{Max_{k,j}}\right)/2$. This is performed to avoid the need to solve complex (nonlinear) scheduling formulations, thus becoming input parameters. The resulting values of these ratios, i.e., $\delta^{In}_{i,k,j,s}$ and $\delta^{Out}_{i,k,j,s}$, are then used in the state-task network representation considered in the scheduling formulation, as shown in **Figure 9**. $\delta^{In}_{i,k,j,s}$ represents the state-task relation (i.e., inputs of a task) for each state s feeding unit j that will realize task k. $\delta^{Out}_{i,k,j,s}$ represents the task-state relation (i.e., outputs of a task) for each state s exiting unit j that realized task k.

IV.1.3. Scheduling Problem

The continuous-time formulation and the State-Task Network (STN) are used in this work to formulate a scheduling problem for chemical batch plants. The parametric sensitivity analysis described above allows the scheduling model to consider the unit operation costs ($\tilde{J}_{i,k,j}$) and unit operation times ($\tau_{Dyn_{i,k,j}}$) based on the set κ. In particular, the set κ is often represented as the quantity of material processed ($B_{i,k,j,e}$) by a unit j realizing task k at event e in the scheduling problem. The continuous-time scheduling formulation proposed in this work is represented in the following MILP problem:

$$\max_{W_{i,k,e},Y_{i,j,e},B_{i,k,j,e},d_{i,s,e}} Z_{Sch}\left(p_s, d_{i,s,e}, \delta_{i,k,j,s}^{In}, \delta_{i,k,j,s}^{Out}, E_i, \Gamma_i, \Lambda_{\mathrm{i}}, H, \tilde{J}_{i,k,j,e}\right) \quad (20)$$

$$s.t.$$

$$\textstyle\sum_{k\in N_{Pr_j}} W_{i,k,e} \leq Y_{i,j,e}, \forall\, j \in N_E, e \in E \quad (20a)$$

$$W_{i,k,e} B_{Min_{k,j}} \leq B_{i,k,j,e} \leq W_{i,k,e} B_{Max_{k,j}}, \forall\, k \in N_{Pr}, j \in N_{E_k}, e \in E \quad (20b)$$

$$ST_{i,s,e} = ST_{i,s,e-1} - d_{i,s,e} + \textstyle\sum_{k\in N_{Pr_s}} \sum_{j\in N_{E_k}} \left(\delta_{i,k,j,s}^{Out} B_{i,k,j,e-1} - \delta_{i,k,j,s}^{In} B_{i,k,j,e}\right), \quad \forall\, s \in S, e \in E \quad (20c)$$

$$ST_{i,s,e} \leq O_s, \forall\, s \in S, e \in E \quad (20d)$$

$$Q_{i,j,e} = Q_{i,j,e-1} + \textstyle\sum_{k\in N_{Pr_j}} B_{i,k,j,e} - \sum_{k\in N_{Pr_j}} \sum_{s\in S_k^{out}} \delta_{i,k,j,s}^{Out} B_{i,k,j,e-1}, \quad \forall\, j \in N_E, e \in E \quad (20e)$$

$$Q_{i,j,e_f} = 0, \forall\, j \in N_E, e \in E \quad (20f)$$

$$\textstyle\sum_{e\in E} d_{i,s,e} \geq r_s \;\; \forall\, s \in S \quad (20g)$$

$$T_{i,k,j,e}^F = T_{i,k,j,e}^S + \Lambda_{i,k,j}^\tau B_{i,k,j,e} + \Gamma_{i,k,j}^\tau W_{i,k,e}, \forall\, k \in N_{Pr_j}, j \in N_E, \; e \in E \quad (20h)$$

$$T_{i,k,j,e}^F \geq T_{i,k,j,e}^S, \forall\, k \in N_{Pr_j}, j \in N_E, e \in E \quad (20i)$$

$$T_{i,k,j,e+1}^S \geq T_{i,k,j,e}^F, \forall\, k \in N_{Pr_j}, j \in N_E, e \in E \quad (20j)$$

$$T_{i,k,j,e+1}^F \geq T_{i,k,j,e}^F, \forall\, k \in N_{Pr_j}, j \in N_E, e \in E \quad (20k)$$

$$T_{i,k,j,e+1}^S \geq T_{i,k,j,e}^S, \forall\, k \in N_{Pr_j}, j \in N_E, e \in E \quad (20l)$$

$$T_{i,k,j,e}^F \leq H, \forall\, k \in N_{Pr_j}, j \in N_E, e \in E \quad (20m)$$

$$T_{i,k,j,e}^S \leq H, \forall\, k \in N_{Pr_j}, j \in N_E, e \in E \quad (20n)$$

$$T_{i,k,j,e_f}^F = H, \forall\, k \in N_{Pr_j}, j \in N_E, e \in E \quad (20o)$$

$$T_{i,k,j,e+1}^S \geq T_{i,l,j,e}^F - H(1 - W_{i,l,e}) \quad \forall\, j \in N_E, k \in N_{Pr_j}, l \in N_{Pr_j}, k \neq l, e \in E, e \neq e_f \quad (20p)$$

$$\tilde{J}_{i,k,j,e} = W_{i,k,e} \Gamma_{i,k,j}^C + B_{i,k,j,e} \Lambda_{i,k,j}^C, \forall\, k \in N_{Pr_j}, j \in N_E, \; e \in E \quad (20q)$$

$$\tilde{J}_T = \textstyle\sum_e^E \sum_k^{N_{Pr}} \sum_j^{N_{E_k}} \tilde{J}_{i,k,j,e} \quad (20r)$$

$$where$$

$$\delta_{i,k,j,s}^{In} \in \delta_i^{In}, \delta_{i,k,j,s}^{Out} \in \delta_i^{Out}, \forall\, k \in N_{Pr_j}, j \in N_E, \forall\, s \in S$$

$$\Gamma_{i,k,j}^\tau, \Gamma_{i,k,j}^C \in \Gamma_i, \forall\, k \in N_{Pr_j}, j \in N_E$$

$$\Lambda_{i,k,j}^\tau, \Lambda_{i,k,j}^C \in \Lambda_i, \forall\, k \in N_{Pr_j}, j \in N_E$$

$$p_s \in P, r_s \in R, \forall\, s \in S$$

where the objective function may be set to maximize the profits obtained from producing a variety of products by the implementation their corresponding recipes into a given chemical batch plant. Note that other types of objective functions can be used in this framework (e.g., minimization of turnaround time). Note that Problem 3 and Problem 4 should consider objective functions with the same objective to achieve coherent solutions (e.g., minimize costs, maximize profits, etc.). p_s is the sale price of state s and E is the number of event points utilized in the scheduling formulation. E is user defined and can be determined iteratively by starting with a small number and adding increments until no more improvements on the objective function are detected[42,43]. Eq. (20a) represents the allocation constraints where $W_{i,k,e}$ is set to zero unless task k at event e is taking place in unit j at the i^{th} iteration; also, this constraint depends on the value of $Y_{i,j,e}$, which can be either 0 or 1 and represents the assignment of unit j at event e at the i^{th} iteration. Eq. (20b) represents the capacity constraint, where $B_{i,k,j,e}$ is the material holdup of unit j while processing task k at event e. Eq. (20c) represents the state material balances. $ST_{i,s,e}$ is the total quantity of state s at event e whereas $\delta^{In}_{i,k,j,s}$ and $\delta^{Out}_{i,k,j,s}$ represent the fraction of state s that enters or leaves task k, respectively. Both $\delta^{In}_{i,k,j,s}$ and $\delta^{Out}_{i,k,j,s}$ are obtained from the parametric sensitivity analysis and represent inputs (i.e., model parameters) to the scheduling formation, as described in section IV.1.2. Eq. (20d) ensures that the material quantity of state s at event e is no greater than its capacity O_s. Eq. (20e) is the material balance inside each unit j at each event e, where $Q_{i,j,e}$ is the quantity of material inside the unit. This balance is conditioned by Eq. (20f) on the last event point, e_f, which sets the material balance equal to zero: no remaining material inside any of the processing units is allowed at the end of the operation. Eq. (20g) represents the market demand constraints, where r_s is the total demand of state s and $d_{i,s,e}$ is the demand satisfied of such state at event e. Note that Eq. (20g) enforces that market requirements r_s provided by the user must be met by the present scheduling formulation. Eq. (20h) represents the time at which unit j stops whichever task it is realizing, and is based on the time unit j needs to realize task k at event e, depending on the linear correlations with the parameters $\Lambda^{\tau}_{i,k,j}$ (slope) and $\Gamma^{\tau}_{i,k,j}$ (intercept) and the quantity of material processed ($B_{i,k,j,e}$) under the condition of the utilization of the unit ($W_{i,k,e}$). Eq. (20h) to Eq. (20o) are time logic constraints, i.e., they enforce that the realization of task k in unit j at event e does not overlap with another task, following the sequence of processes as specified, while delimiting and specifying the time of realization (specifying a start and an end) through the time axis and enforcing that no unit can be operated beyond the time horizon H. Eq. (20p) ensures that a unit that can perform multiple tasks is only used to perform a task at a time. Eq. (20q) is used to calculate the dynamic unit operating costs, which depend on the linear correlations obtained from the parametric sensitivity analysis with the parameters $\Lambda^{C}_{i,k,j}$ (slope) and $\Gamma^{C}_{i,k,j}$ (intercept) and the quantity of material processed ($B_{i,k,j,e}$) under the condition of the utilization of the unit ($W_{i,k,e}$). Eq. (20r) determines the total operating cost of the plant.

IV.1.4. Dynamic Optimization Problem

The solution to Problem 20 (Z^{*}_{Sch}) returns the set of scheduling decisions β_{Sch_i} ($B_{i,k,j,e}$, $W_{i,k,e}$ and $Y_{i,j,e}$) obtained at the i^{th} iteration step. The set β_{Sch_i} represents the key inputs required for the formulation of the dynamic optimization problem that aims to find the control actions and unit operation times that optimize the economics of the batch plant. This problem is formulated as follows:

$$\max_{u_{i,k,j,t},\tau_{Dyn_{i,k,j}}} Z_{Dyn}\left(x(t)_{i,k,j}, \dot{x}(t)_{i,k,j}, u(t)_{i,k,j}, \Psi, t, \tau_{Dyn_{i,k,j}}, C, \beta_{Sch_{i,k,j}}\right) \quad (21)$$

$s.t.$

$$f_p\left(x(t)_{i,k,j}, \dot{x}(t)_{i,k,j}, u(t)_{i,k,j}, \psi_{Nom}, t, \tau_{Dyn_{i,k,j}}, \beta_{Sch_{i,k,j}}\right) = 0, \quad \forall\, t, p \in N_p, k \in N_{Pr}, j \in N_E \quad (21a)$$

$$h_q\left(x(t)_{i,k,j}, \dot{x}(t)_{i,k,j}, u(t)_{i,k,j}, \psi_{Nom}, t, \tau_{Dyn_{i,k,j}}, \beta_{Sch_{i,k,j}}\right) \pm \lambda b_{i-1,k,j,q,t} \leq 0, \quad \forall\, t, q \in N_q, k \in N_{Pr}, j \in N_E \quad (21b)$$

$$\tau_{Min_{k,j}} \leq \tau_{Dyn_{i,k,j}} \leq \tau_{Max_{k,j}},\ \tau_{Dyn_{i,k,j}} \in \tau_i,\ \forall, k \in N_{Pr}, k \in N_{Pr}, j \in N_E \quad (21c)$$

where
$\psi_{Nom} \in \Psi,\ \tau_{Dyn_i} \in \tau_i,\ \tau_i \in \mathrm{T}$

where Z_{Dyn} represents an objective function that considers the transient operation of each task in the plant; $\beta_{Sch_{i,k,j}}$ are the scheduling decisions for unit j performing task k and $\beta_{Sch_{i,k,j}} \in \beta_{Sch_i}\ \forall\ k \in N_{Pr}, j \in N_E$. Eq. (21a) represents the mechanistic model of the processes included in the batch plant. The introduction of the back-off terms in Eq. (21b) enforces the optimization to find control decisions and unit operation times backed-off from the nominal values that ensure dynamic feasibility. The back-off terms $\lambda b_{i,k,j,q,t}$ represent the deviation of the system from the nominal point under stochastic realizations in the uncertain model parameters at the inequality constraint q of unit j performing task k at time t at the i^{th} iteration. Note that the back-off terms include a multiplier-factor λ, which can be thought as the level confidence given to each of the back-off terms (i.e., the amount of output variability that needs to be accommodated due to uncertainty). A more detailed discussion on the back-off terms and their calculation is provided in section IV.1.7. Note that back-off terms at $i = 0$ are set to zero (as there is no information from the plant that could be used for their calculation). Eq. (21b) represents the process constraints for each unit, and they involve production goals, safety requirements, quality concerns and/or feasibility limitations. Eq. (21c) represents the bounds on the processing unit times ($\tau_{Dyn_{i,k,j}}$). $B_{i,k,j,e}$ is an input to problem (4) and represents the initial amount of material that needs to be processed. Problem (4) aims to find optimal control profiles and unit operation times that can process $B_{i,k,j,e}$ in the presence of uncertainty. Note that $B_{i,k,j,e}$ may differ at the end of operation for some specific processes depending on the nature of their own operation, e.g., semi-batch processes involving inlet streams that may continuously feed material into the unit.

IV.1.5. Stochastic Simulations

The solution from Problem 21, Z_{Dyn}^{*}, returns the optimal control profiles ($u_{Dyn_{i,k,j}}$) and unit operation times ($\tau_{Dyn_{i,k,j}}$) in the presence of back-off terms that were estimated in a preceding iteration. At this step, both $u_{Dyn_{i,k,j}}$ and $\tau_{Dyn_{i,k,j}}$ are defined as model parameters (i.e., fixed in the calculations) and used to evaluate the system under the influence of n stochastic realizations in the uncertain parameters ($\psi_{Unc_{i,n}}$). The quantity of material to be processed on each unit ($B_{i,k,j,e}$) depend on the performance of preceding tasks under uncertainty using fixed control profiles and processing times obtained from problem (21). Accordingly, $B_{i,k,j,e}$ are only used to initialize problem (22) since it is not known a priori the quantity of material that would be transferred from preceding tasks due to uncertainty in the model parameters. The probability functions that were used to describe ψ_{Unc} in the initialization step (section IV.1.1) are used to generate the random values of each realization in the uncertain parameter set using Monte Carlo sampling techniques. The statistical data generated from this step is used to compute the back-off terms, as described in the next section. The problem considered for this step is as follows:

$$\max \ \zeta_{i,k,j,q,t,n,m} = h_q\left(x(t)_{i,k,j,n,m}, \dot{x}(t)_{i,k,j,n,m}, u_{Dyn_{i,k,j}}, \psi_{Unc_{i,n}}, t, \tau_{Dyn_{i,k,j}}\right) \qquad (22)$$

$$s.t.$$

$$f_p\left(x(t)_{i,k,j,n,m}, \dot{x}(t)_{i,k,j,n,m}, u_{Dyn_{i,k,j}}, \psi_{Unc_{i,n}}, t, \tau_{Dyn_{i,k,j}}\right) = 0, \qquad (22a)$$

$$\forall\, t, p \in N_p, k \in N_{Pr}, j \in N_E$$

where

$$\psi_{Unc_{i,n,m}} \in \Psi, \tau_{Dyn_{i,k,j}} \in \tau_i, \tau_{Dyn_i} \in \tau_i, \ \tau_i \in \mathrm{T}$$

Problem 22 is solved n times where n represents a specific random realization in the set of uncertain parameters ψ_{Unc}. As shown in **Figure 8**, this problem is solved in batches of size L (i.e., $n \in [0, L)$) until a user-defined criterion (Eq. 23) is met. The total number of iterations (N_{MC}) necessary to satisfy criterion I (Eq. 23) is unknown *a priori*. Hence, the index m is introduced here to keep track of the number of iterations needed to satisfy criterion I (Eq. 23), i.e., $m \in [0, N_{MC})$. Note that problem 22 is solved for LN_{MC} times. Note that problem 22 is a feasibility problem, i.e., only the mechanistic process model constraint (22a) is enforced whereas the maximum deviation in the process constraints (h_q) is assessed (ζ) under random realizations in the uncertain parameters.

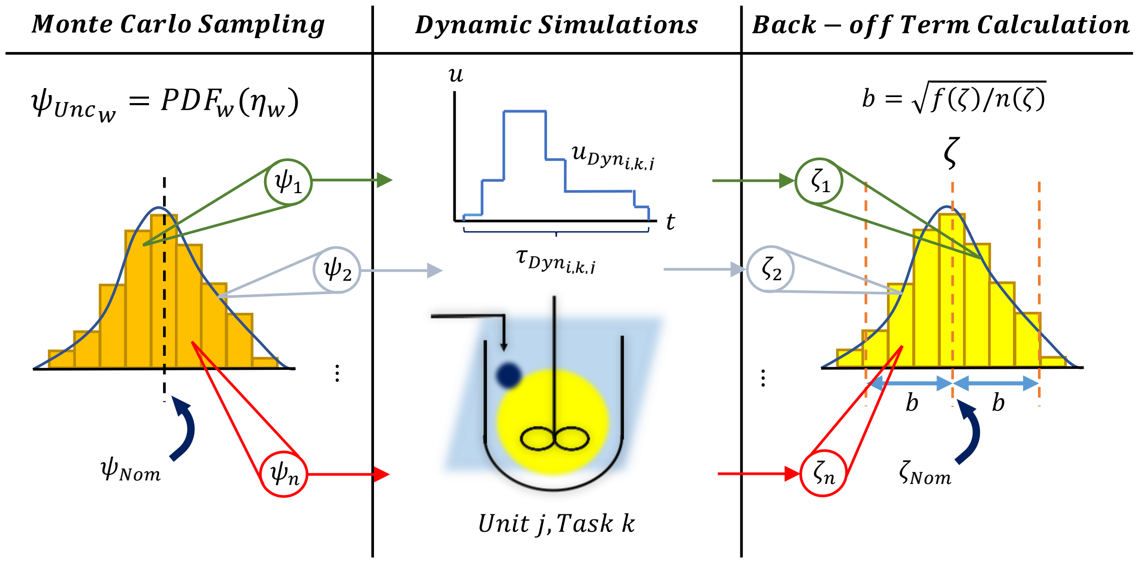

Figure 10. Stochastic Simulations and Back-off Term Calculation Diagram.

A graphical representation of a single iteration in m is shown in **Figure 10**. In the Monte Carlo Sampling section, the uncertain parameters are described by their corresponding PDF's. Random realizations of the uncertain parameters are selected and used to evaluate the dynamic feasibility of the system under those conditions, i.e., solving an instance of Problem 22. The values of the observed system variables (ζ) are then stored after each solution of Problem 22. The back-off term calculation is performed with the gathered information at the end of each m iteration (considering the information of previous batches).

Criterion 3 (Eq. 23) quantifies the deviations in the back-off terms between the m^{th} and $m-1^{th}$ data populations. Each population m is composed of $L * m$ realizations in the uncertain parameters, i.e., population m contains the information generated at the m^{th} iteration and the information generated in previous iterations. This criterion terminates the back-off calculation loop (inner loop in **Figure 8**) when the errors between back-off terms for iteration m and those for iteration $m-1$ satisfy a user defined tolerance (Tol_{SS}).

Criterion 3 (Eq. 23) is only active for $m \geq 2$. Note that this criterion must be satisfied for all the back-off terms considered in the formulation. The specific procedure to estimate the back-off terms is presented in section IV.1.7.

$$\left|1 - b_{i,k,j,q,t,m-1}/b_{i,k,j,q,t,m}\right| \leq Tol_{SS}\ \forall\ \mathrm{t}, k \in N_{Pr}, \mathrm{j} \in N_E, \mathrm{q} \in N_q \tag{23}$$

where $b_{i,k,j,q,t,m}$ is the back-off term calculated at the m^{th} iteration of the back-off calculation. Note that index m is dropped from the notation of the back-off terms outside of the inner loop shown in **Figure 8** since it is only used to distinguish the back-off terms calculated in between batches of size L, i.e., $b_{i,k,j,q,t,m} \rightarrow b_{i,k,j,q,t}$ when Criterion 3 (Eq. 23) has been satisfied. Note that Eq. (23) is only active for $m \geq 2$. Also, this criterion must be satisfied for all the back-off terms considered in the formulation. Additional information on this step and the specific procedure to estimate the back-off terms are provided in section III.2.5.

IV.1.6. Algorithm Termination

The algorithm is terminated when the difference between successive back-off terms calculated at the current and previous iteration is below a user-defined tolerance (Tol_{BO}), i.e.,

$$\left|1 - b_{i-1,k,j,q,t}/b_{i,k,j,q,t}\right| \leq Tol_{BO}, \forall\ \mathrm{t}, k \in N_{Pr}, \mathrm{j} \in N_E, \mathrm{q} \in N_q \tag{24}$$

Otherwise, the algorithm proceeds to the next iteration (as shown in **Figure 8**). Due to the iterative nature of the back-off approach, and its dependence on the random statistical parameters, local solutions are not guaranteed while using this method. However, the methodology provides a scheduling and control scheme that exhibits dynamic feasibility thus providing higher satisfaction of the process constraints in the presence of stochastic realizations in the system parameters.

IV.1.7. *Back-Off Term Calculation*

In this work, a back-off term ($\lambda b_{i,k,j,q,t}$) is the representation of the deviation of the system in the values of the observed variables of the of the q^{th} constraint function at a time t for a unit j performing task k at the i^{th} iteration for stochastic realizations in the uncertain parameters. A back-off term is composed by the amount of variability (λ) considered in the q^{th} constraint and the deviation of the system caused by uncertainty ($b_{i,k,j,q,t}$). λ is a user-defined parameter and it could have different values for each q constraint, i.e., λ_q.

The introduction of $\lambda b_{i,k,j,q,t}$ into the constraints h_q is expected to drive the process to search for a solution capable of accommodating a λ level of uncertainty, backed off from the optimal solution using nominal values in the uncertain model parameters[18]. Note that there is a back-off term for each inequality h_q at each time point t. The standard deviation on the back-off terms is calculated as follows:

$$b_{i,k,j,q,t} = \sqrt{\frac{1}{N_{MC}}\sum_{n=1}^{N_{MC}}\left[\zeta_{i,k,j,q,t,n} - \frac{1}{N_{MC}}\sum_{n=1}^{N_{MC}}\zeta_{i,k,j,q,t,n}\right]^2} \quad \forall\ t, k \in N_{Pr}, j \in N_{E_{i,q}}, q \in N_{q_i} \tag{25}$$

where $\zeta_{i,k,j,q,t,n}$ are obtained from the solution of problem 22. This expression assumes that the data used for the back-off terms calculation, a set of $\zeta_{i,k,j,q,t,n}$, follow a statistical distribution that can be represented by a normal distribution. Eq. (25) represents the calculation of the normal standard deviation for discrete random variables and is used for the calculation of the deviation of a process subjected to stochastic uncertainty for the back-off terms calculations in this work at each iteration m. Note that Eq. (25) may need to be adapted in case that the data set of $\zeta_{i,k,j,q,t,n}$ does not follow a normal distribution.

IV.2. Case Study

The performance of the methodology of Algorithm B has been evaluated using a slightly modified version of the case study described in the previous chapter (section III.3). **Table 5** lists the chemical reaction parameters that were modified. For this case study, it is assumed that k_{AB}, k_{BC}, k_{BDE} and k_{2DF} are the uncertain parameters in the system that follow a normal distribution with a standard deviation equivalent to 5% of their expected (nominal) values (reported in **Table 4**).

Table 5. Case Study – Modified chemical reaction parameters.

Variable	Value	Description
$\boldsymbol{E_{A\rightarrow B}}$ $[\boldsymbol{K}]$	5000	Activation Energy in RI
$\boldsymbol{E_{B\rightarrow C}}$ $[\boldsymbol{K}]$	8000	Activation Energy in RI

The profit function for the dynamic optimization problem considered for this case study is as follows:

$$Z_{Dyn} = \sum_{p}^{N_{Pr}} \hat{c}_p + \check{c}_{OPcosts};\ N_{Pr} = \{RI, RII, FI, SI\} \tag{28}$$

where $\check{c}_{RI}$, $\check{c}_{RII}$, $\check{c}_{FI}$ & $\check{c}_{SI}$ represent the operation costs and revenues generated (if any) by each of the processes (tasks). $\check{c}_{OPcosts}$ is a cost function that evaluates possible losses in production due to changes in the unit operation times. That is, Eq. (28) represent the calculations of the chemical plant profits. The overall cost of the process RI is:

$$\check{c}_{RI} = \check{c}_{RI_A} + \check{c}_{RI_{Aux}} \tag{29}$$

where $\check{c}_{RI_{Aux,}}$ represents the total cost of the auxiliary services utilized to perform RI, which is directly related to the valve opening fractions ($F_{Hot_j}(t)$ & $F_{Cool_j}(t)$):

$$\check{c}_{RI_{Aux}} = \sum_{j}^{N_{E_{RI}}} \left(p_{F\,Hot} \int_0^{\tau_{Dyn_{i,j}}^{RI}} F_{Hot_j}(t)dt + p_{F\,Cool} \int_0^{\tau_{Dyn_{i,j}}^{RI}} F_{Cool_j}(t)dt \right) \tag{30}$$

where $p_{F_{Hot}}$ and $p_{F_{Cool}}$ are the price of using the heating or the cooling lines, respectively. Likewise, $\check{c}_{RI_A}$ is the total cost of the quantity of raw species A loaded to perform RI:

$$\check{c}_{RI_A} = p_A \sum_j^{N_{E_{RI}}} n_{A_{i,j}}(t)|_{t=0} \tag{31}$$

where p_A is the cost per unit of species A, $n_{A_{i,j}}(t)|_{t=0}$ is the initial quantity of species A in unit j that performs RI. The overall cost of the process FI is described by a penalization of waste generated in this unit (i.e., the amounts of unreacted A and the by-product C), i.e.:

$$\check{c}_{FI} = p_{Waste} \sum_j^{N_{E_{FI}}} \left(n_{A_{i,j}}(t) + n_{C_{i,j}}(t)\right)|_{t=\tau^{RI}_{Dyn_{i,j}}} \tag{32}$$

where p_{Waste} is the penalty cost, $n_{A_{i,j}}(t)|_{t=\tau^{RI}_{Dyn_{i,j}}}$ and $n_{C_{i,j}}(t)|_{t=\tau^{RI}_{Dyn_{i,j}}}$ are the quantity of A and C, respectively, at the end of operation of unit j that performed RI. The overall cost of process RII, which is the cost of feeding species D into the reactor through, is:

$$\check{c}_{RII} = \check{c}_{RII_D} = p_{F_{Feed}} \sum_j^{N_{E_{RII}}} \int_0^{\tau^{RII}_{Dyn_{i,j}}} V_{R_{i,j}}(t) C_{F_{i,j}}(t) dt \tag{33}$$

where $p_{F_{Feed}}$ is the unitary cost of species D, $V_{R_{i,j}}(t)$ is the volumetric flow of the feeding of unit j at time t and $C_{F_{i,j}}(t)$ is the concentration of feed of D. The overall cost of process SI ias as follows:

$$\check{c}_{SI} = \check{c}_{SI_{Mixi}} - \check{c}_{SI_E} \tag{34}$$

In Eq. (34) $\check{c}_{SI_E}$ represents the revenues for selling species E, i.e.:

$$\check{c}_{SI_E} = p_{Pure} \sum_j^{N_{E_{RII}}} n_{E_{i,j}}(t)|_{t=\tau^{RII}_{Dyn_{i,j}}} \tag{35}$$

where p_{Pure} is the unitary price of E and $n_{E_{i,j}}(t)|_{t=\tau^{RII}_{Dyn_{i,j}}}$ is the quantity of E in unit j at the end of operation. $\check{c}_{SI_{Mix}}$ is a penalty cost for the quantity produced of the mixture of D and F in process SI:

$$\check{c}_{SI_{Mix}} = p_{Impure} \sum_j^{N_{E_{RII}}} \left(n_{D_{i,j}}(t) + n_{F_{i,j}}(t) \right) |_{t=\tau^{RII}_{Dyn_{i,j}}} \tag{36}$$

where p_{Impure} is the unitary penalty cost of the mixture and $n_{D_{i,j}}(t)$ and $n_{F_{i,j}}(t)$ is the quantity of D and F in unit j at time t, respectively.

The objective function (Eq. 28) which includes the penalty cost $\check{c}_{OPcosts}$ of possible losses in production, based on unit operation times, $\check{c}_{OPcosts}$ (Eq. 37), which compares the unit operation times calculated in the current iteration with those calculated in the previous iteration. This function penalizes the unit operation times and is mainly used as an incentive to find optimal unit operation times.

$$\check{c}_{OPcosts} = \left(\sum_k^{N_{Pr}^{Dyn}} \sum_j^{N_{E_k}} \sum_a^{N_{Prod}} p_a D_{a_{i,j}}(t)|_{t=\tau^k_{Dyn_{i,j}}} \right) * \left(\xi\left(N_{Pr}^{Dyn}\right) - \sum_k^{N_{Pr}^{Dyn}} \left(\frac{\sum_j^{N_{E_k}} \tau^k_{Dyn_{i,j}}}{\sum_j^{N_{E_k}} \tau^k_{Dyn_{i-1,j}}} \right) \right) \tag{37}$$

where N_{Pr}^{Dyn} represents the set of processes that consider the dynamic operation (i.e., their unit operations times are optimized); $\xi(A)$ is the cardinal number of a set (cardinality). N_{Prod} is the set of substances that are sold generate a revenue, where a is an element of said set. $D_{a_{i,j}}(t)|_{t=\tau^k_{Dyn_{i,j}}}$ is the quantity of substance a produced at unit j performing task k at the end of operation. p_a is the selling price of substance a. The arrange for this case study as follows:

$$\check{c}_{OPcosts} = p_{Pure} \sum_j^{N_{E_{RII}}} D_{E_{i,j}}(t)|_{t=\tau^{RII}_{Dyn_{i,j}}} \left(\xi\left(N_{Pr}^{Dyn}\right) - \sum_k^{N_{Pr}^{Dyn}} \left(\frac{\sum_j^{N_{E_k}} \tau^k_{Dyn_{i,j}}}{\sum_j^{N_{E_k}} \tau^k_{Dyn_{i-1,j}}} \right) \right) \tag{38}$$

where N_{Pr}^{Dyn} represents the set of processes that consider the dynamic operation (i.e., their unit operations times are optimized), whereas $\xi(A)$ is the cardinal number of a set (cardinality). Substance E is the only one that produces an income.

Note that the scheduling formulation uses a similar profit function to guide the optimization:

$$Z_{Sch} = \check{c}_{ST} - \tilde{J}_T \tag{39}$$

where $\check{c}_{ST}$ represents the overall sum of costs and revenues for the flow of material in the plant (states) and $\tilde{J}_T$ is the total operation costs (as described in section IV.1.3). The term $\check{c}_{ST}$ represents the overall sum of costs and revenues for the flow of material in the plant (states):

$$\check{c}_{ST} = \sum_e^E \sum_s^S d_{i,s,e}\, p_s \tag{40}$$

where p_s represents the cost or profit for each state s, where $p_s < 0$ is for the consumed materials and wastes, $p_s > 0$ is for the final products and $p_s = 0$ is for intermediary states (p_s is a user-defined parameter); $d_{i,s,e}$ is the market demand of materials at each event point e.

The objective of this case study is to identify a scheduling and control strategy that can maximize the batch plant profits in the presence of uncertainty in the model parameters.

IV.3. Results

The decomposition back-off algorithm B has been implemented to address the optimal scheduling and control of the batch plant presented in **Figure 5¡Error! No se encuentra el origen de la referencia.**. Pyomo 5.6.6[73], running in Python 3.7, was used to solve the scenarios presented in this work. A computer with an Intel® Core™ i7-8700 CPU @ 3.2 GHz and 16 GB of RAM was employed. CPLEX was used for the MILP scheduling problem and ipopt with ma57 HSL linear solver, for the dynamic optimization problem and the stochastic simulations. The Differential-Algebraic Equations (DAE)[72] module from Pyomo was utilized for the discretization of the differential equations. Orthogonal collocation on finite elements was chosen as the discretization method. For the present case study, 40 finite elements and 3 collocation points for each dynamic process were considered. Note that these parameter values returned acceptable solutions in an average CPU time of ~2 s per simulation.

IV.3.1. Scenario 1: Effect of Constraint Variability Limits.

The purpose of this scenario is to assess scheduling and control decisions using different variability limits on the process constraints (i.e., different values of λ). In the present scenario, only one unit is available per process, i.e., one unit for RI, RII, FI and SI, respectively. Three instances were considered for this scenario: A) Nominal Case – no back-off term implementation ($\lambda = 0$); B) Uncertainty in the reaction kinetic parameters, i.e., k_{AB}, k_{BC}, k_{BDE} and k_{2DF} (see supplementary material) using $\lambda = 2$; C) Similar to B) but setting $\lambda = 3$. Note that the solution at the iteration zero ($i = 0$) corresponds to the solution of the optimization of the system at the nominal conditions in the operation, i.e., the nominal solution is the solution of the Problem 19, Problem 20 and Problem 21 when all the back-off terms are equal to zero.

This nominal solution is used as a benchmark to compare the solutions of instances B and C described above.

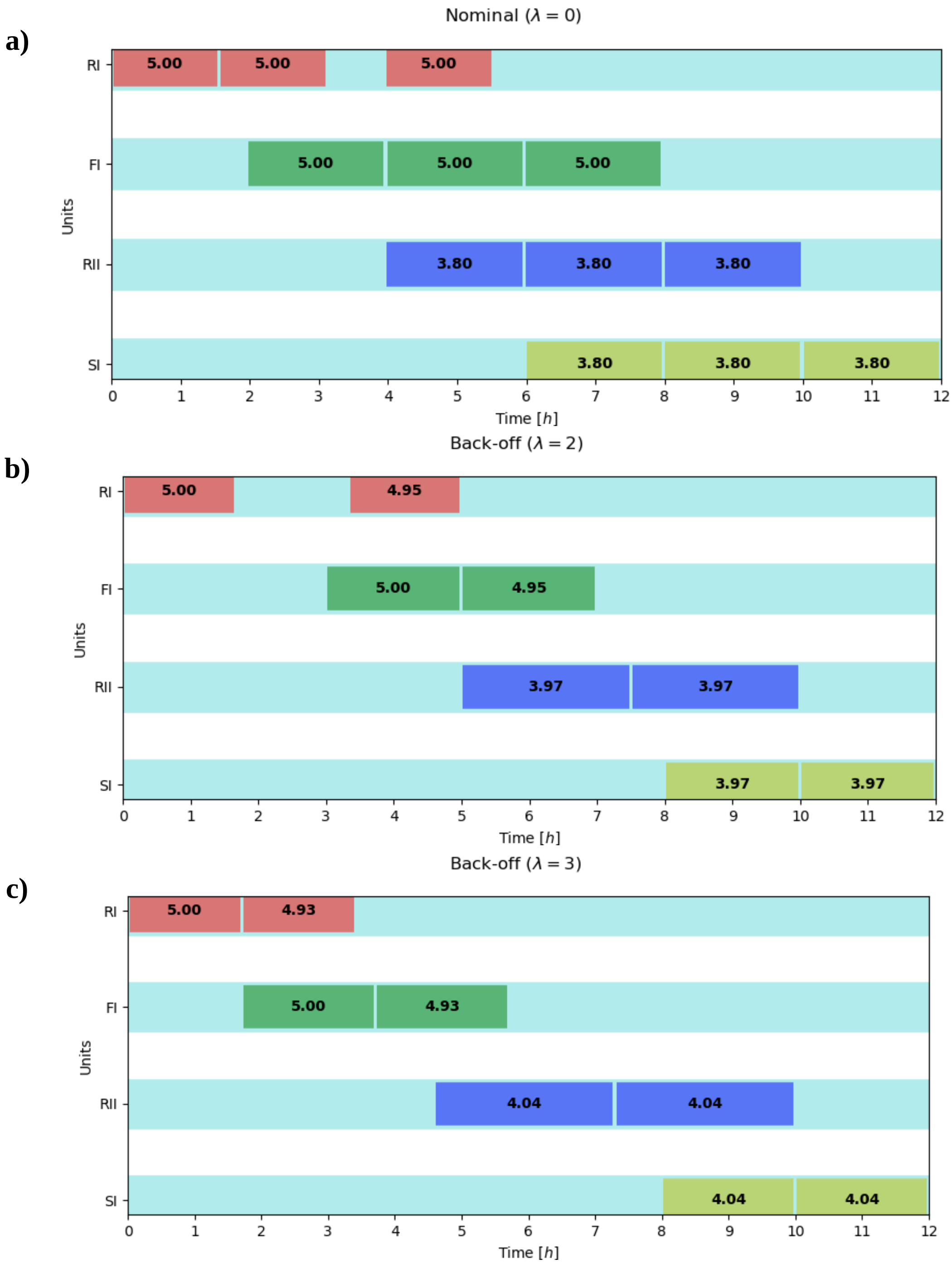

Figure 11. Algorithm B – Scheduling Plans
(a) Nominal solution ($\lambda = 0$), **(b)** back-off method ($\lambda = 2$) and **(c)** back-off method ($\lambda = 3$). Each unit allocation is portrayed with their corresponding material processing quantity in m^3.

Figure 11 shows the scheduling solutions for each of the instances considered for the present scenario. As shown in this figure, the main difference between the scheduling decisions is the number of total allotted jobs, i.e., 3 jobs per process for the nominal condition (instance A) and 2 jobs per process for the backed-off solutions (i.e., instances B and C). A job in this scenario consists in the allocation of one unit per task where the material follows the sequence portrayed in **Figure 11¡Error! No se encuentra el origen de la referencia.**, i.e., when the material completes the basic processing sequence. This difference is mainly due to the influence of the back-off terms over the unit operation times for RII since larger unit operation times are needed in this unit to accommodate increasing values of λ (as shown in **Table 6**). This is because the system is required to satisfy the backed-off constraints (i.e., uncertainty in the reaction kinetic parameters) which makes the system slightly more constrained, by demanding a higher quality of the product, while the system aims for a regime that maximizes the plant profits. Note that the unit operation times for RI follow a similar trend, i.e., their unit operation times increases with increasing values of λ.

Note that the scheduling solutions reported in **Figure 11** are not unique; it can be visually checked that some process allocations allow different permutations that yield the same performance and profits (i.e., the instant at which certain operations may take place could be slightly different as there are gaps along the time axis that the time horizon allows). As shown in **Table 6**, there is a reduction in the total profits when comparing the nominal case (A) with the other instances (B and C) as the value of λ increases. Nonetheless, when the average profits per scheduled job (Total Net Profit/# of Jobs) are compared, an increase can be observed. Since larger values of λ force the process constraints to slightly modify the production; that is, to accommodate the stochastic uncertainty in the reaction kinetic parameters, product quality constraints are required to ensure a higher product quality (with reduced wastes). This results in a slight increase in the averaged net profits.

Regarding the plant operational costs, the dominant terms in Eq. (30) are the auxiliary service fees for reaction RI (the cost $\check{c}_{RI}$) and the reactant feeding in RII ($\check{c}_{RII}$) since they directly depend on the control profiles and the corresponding unit operation time. Consequently, the operation costs calculated through the linear correlations developed for the scheduling formulation (see section IV.1.2 & IV.1.3) impact the plant profits. As shown in **Table 6**, the current back-off formulation can maintain the average operational costs (Total Operation Cost/# of jobs) at similar values but with a very slight decrease as λ increases.

Table 6. Algorithm B – Results for the First Scenario.

Case	$Nominal\ (\lambda = 0)$	$Back - Off\ (\lambda = 2)$	$Back - Off\ (\lambda = 3)$
$\boldsymbol{RI}$ UOT ($\boldsymbol{\tau_{RI}}$)	$1.56\ h$	$1.66\ h$	$1.72\ h$
$\boldsymbol{RII}$ UOT ($\boldsymbol{\tau_{RII}}$)	$2.02\ h$	$2.5\ h$	$2.74\ h$
Avg. CPU Time per iteration	$\sim 8.01\ h$	$\sim 10.4\ h$	$\sim 11.86\ h$
Total Operation Cost	$32{,}336.04\ m.u.$	$21{,}226.84\ m.u.$	$21{,}082.34\ m.u.$
Avg. Cost per job	$10{,}778.68\ m.u.$	$10{,}613.42\ m.u.$	$10{,}541.17\ m.u.$
Total Net Profit	$54{,}095.58\ m.u.$	$41{,}613.42\ m.u.$	$44{,}129.72\ m.u.$
Avg. Profit per job	$18{,}031.86\ m.u$	$20{,}806.71\ m.u.$	$22{,}064.86\ m.u.$
Jobs	3	2	2
Total Iterations	1	8	7

Figure 12 shows the profiles of the controlled (temperature) and manipulated (auxiliary services valve opening) variables for process RI, whereas **Figure 13** illustrates the profiles of the controlled variables (species D feed) and manipulated variables (inlet valve opening) for RII. Note that the unit operation times are different depending on the level of variability (λ) considered in the analysis. Also, the control profiles are non-trivial and exhibit significant differences that are necessary to achieve the backed-off production constraints under different levels of constraint satisfaction. The differences observed in the control profiles, combined with the corresponding unit operation times, are the main factors that cause the increments in the average net profits and the decrease in the average costs. Note that the temperature constraint in **Figure 12** refers to the final temperature value constraint of the reactor's temperature for RI.

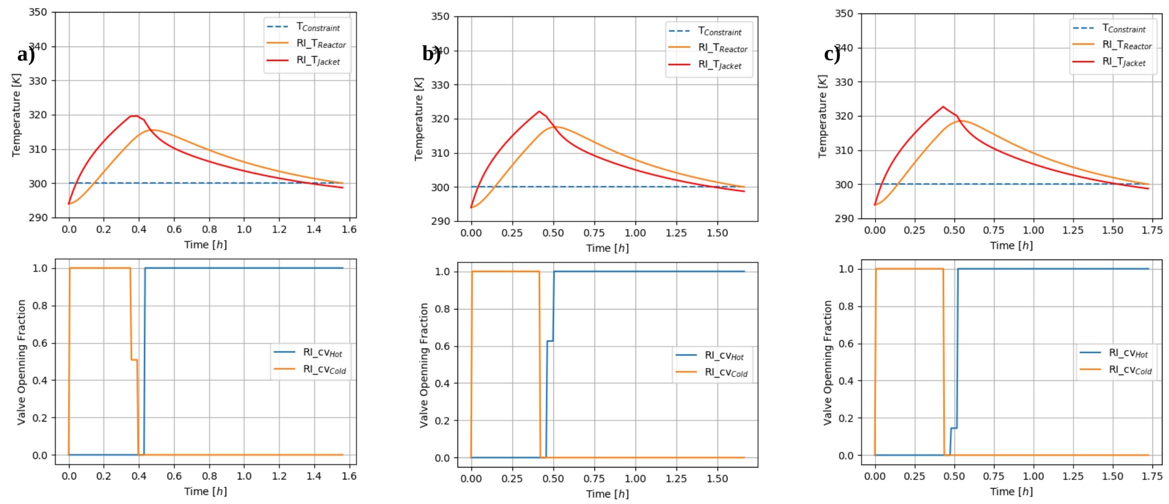

Figure 12. Algorithm B – Controlled variable (Reactor Temperature) and control decisions (Aux. Serv. Valves) for RI in the first scenario.
(a) Nominal solution ($\lambda = 0$), **(b)** back-off method ($\lambda = 2$) and **(c)** back-off method ($\lambda = 3$). cv_{Hot} and cv_{Cold} represent the valve's opening of the hot and cold auxiliary service, correspondingly.

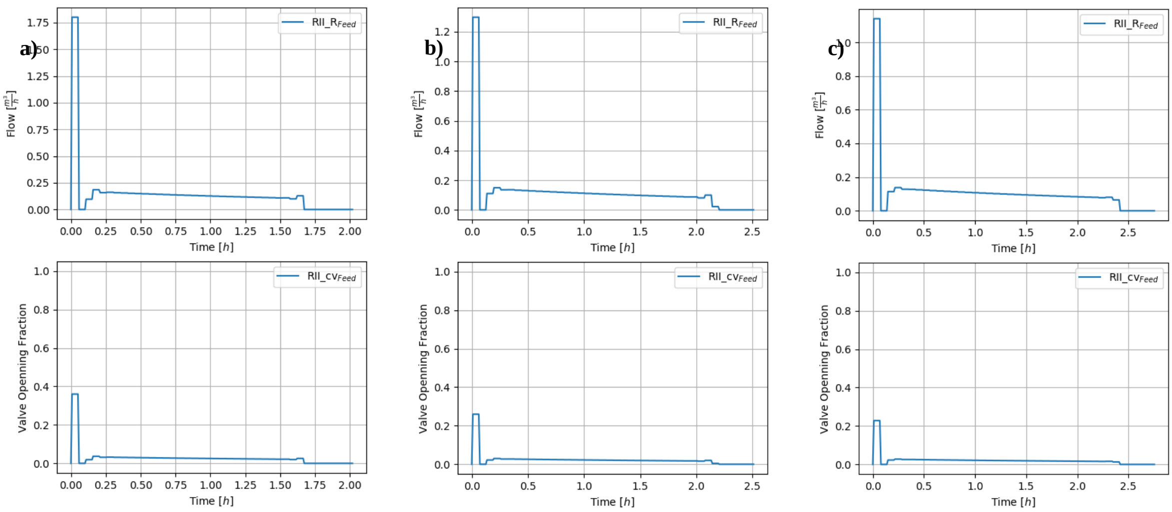

Figure 13. Algorithm B – Controlled variable (Species D inflow) and control decisions (Inlet valve) for RII in the first scenario.
(a) Nominal solution ($\lambda = 0$), **(b)** back-off method ($\lambda = 2$) and **(c)** back-off method ($\lambda = 3$). cv_{Feed} corresponds to the feeding stream valve opening.

As shown in **Figure 12**, increasing the value of λ implies that auxiliary services must be used for an extended period of time to achieve the corresponding production goals. Consequently, a slight increase in the operation costs for process RI is observed as the value of λ is increased. This behaviour is compensated by a slight decrease in the costs of process RII when large values of λ are considered. This is because the overall consumption of species D in process RII (which is fed to the reactor to realize the reaction) is slightly lower as λ increases (compared to $\lambda = 0$). This combination of events drives down the total operation cost per job by approximately 1.5% ($\lambda = 2$) and 2% ($\lambda = 3$) when compared to the nominal case ($\lambda = 0$). While the average costs per job remain at somewhat similar values, the production of species E increases as λ increases, as the back-off production constraint demands a higher quantity of E, thus yielding higher profits: 15.4% for $\lambda = 2$ and 22.4% for $\lambda = 3$ when compared to $\lambda = 0$. As shown in **Table 6**, more CPU time per iteration is needed: closer to a 30% for $\lambda = 2$ and almost 50% for $\lambda = 3$. In summary, as the value of λ increases, an increase in the performance per job is observed (measured by average per job): there is an increment in the net profits while slightly decreasing operation costs. The downside is the loss of 1 job when uncertainty is considered. Nonetheless, the main benefit of implementing this back-off methodology is the that the batch plant can obtain a reliable quality of the product under stochastic parametric uncertainty at the expense of larger unit operation times.

IV.3.2. Scenario 2: The Effect of Multiple Units.

This scenario aims to assess the performance of the back-off methodology B under a more complex scenario involving more than one available processing unit for the dynamic processes involved in the chemical batch plant. The considerations made for this scenario are as follows: I) The plant consists of two reactors with different capacities for RI, two reactors with different capacities for RII, one filter for FI and one separator for SI. II) Uncertainty in the parameters k_{AB}, k_{BC}, k_{BDE} and k_{2DF} (see supplementary material), with $\lambda = 2$. The capacities for the processes with multiple units (RI & RII), are presented in **Table 7**. Note that the unit capacities for RI and RII are different than those considered for scenario 1.

Table 7. Algorithm B – Reactor material capacities for the second scenario.

Unit	RI_1	RI_2	RII_1	RII_2
Capacity [m^3]	4	5	5	6

Figure 14 shows the scheduling obtained for this scenario. Compared to scenario 1 of section IV.3.1, instance B), the scheduling for this scenario presents 2 jobs that exchange processing material: First job consists of the first unit allocations of R1_1 and RI_2, the first allotment of FI and SI, and the unique allotment of RII_1. The second job consists of the remaining unit allocations. Another key difference is that certain units are not operating at their full capacity. Because the unit operation costs are assumed to follow a linear behavior, the system perceives the same performance by either allocating the units with a full or partial loads. Similarly to scenario 1 (instance B), the scheduling decisions are not unique for the present scenario. Note that the time required to achieve the production constraints in both reactions slightly decreases as the total mass of reactants increases, as seen for RII_1 and RII_2 in **Figure 14**, where both reactors perform process RII with different processing quantities.

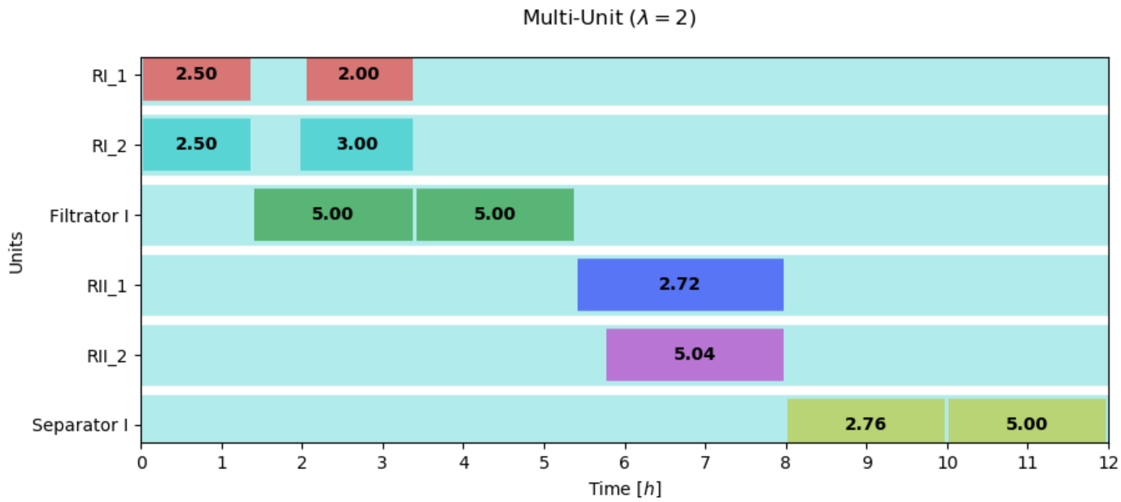

Figure 14. Algorithm B – Scheduling solution for the back-off method ($\lambda = 2$) – multi unit case.
Each unit allocation is portrayed with their corresponding material processing quantity in m^3.

Table 8 reports the results obtained for the present scenario. Comparing the average job cost of this scenario with the total job cost of instance B) in scenario 1, there is an increase of an 8.26%. The increase in costs is due to the variation in operation costs between both units available for each reaction process: one unit is slightly more expensive to operate when the operation of both units is compared when processing same quantities of material. Also, the increase in costs is due to the units not operating at their full capacity since there is no incentive to process material that will not result in a desired product. Note that a larger Time Horizon may allow to operate more units at full capacity that may result in more profitable operation. On the other hand, the net profits decreased by 6.1% compared to instance B) in scenario 1, for the same reasons that caused the cost increment.

Table 8. Algorithm B – Results for the Second Scenario.
Case $\boldsymbol{BO}\ (\boldsymbol{\lambda = 2}) - \boldsymbol{Multi\ Unit.}$

Unit	**1**	**2**
$\boldsymbol{RI}$ UOT ($\boldsymbol{\tau_{RI}}$)	1st: $1.39\ h$ 2nd: $1.35\ h$	1st: $1.39\ h$ 2nd: $1.43\ h$
$\boldsymbol{RII}$ UOT ($\boldsymbol{\tau_{RII}}$)	$2.6\ h$	$2.24\ h$
Avg. CPU Time per iteration	$\sim 35.74\ h$	
Total Operation Cost	$22{,}980.84\ m.u.$	
Avg. Cost per job	$11{,}490.42\ m.u.$	
Total Net Profit	$39{,}074.22\ m.u.$	
Avg. Profit per job	$19{,}537.11\ m.u$	
Jobs	2	
Total Iterations	7	

Compared to the solutions obtained for scenario 1 (instance B), the second scenario needs at least twice the CPU time per iteration to arrive to a solution, which is due to the increase in the interactions between units under uncertainty. That is, this scenario considers 6 units that interact among themselves and that are subject to stochastic uncertainty (4 RI reactors feeding 2 RII reactors). This implies that the propagation of uncertainty occurs among a larger number of processing units thus leading to a significant increase in the number of random realizations needed to estimate the back-off terms (see sections IV.1.5 and IV.1.7). Hence, more CPU time is needed to compute the corresponding back-off terms at each iteration step. Note that in scenario 1 only one unit is considered for RI and RII.

As shown in **Table 8**, the total net profit for the present scenario is 6.1% lower than that obtained from scenario 1 (instance B). The number of available units, the variations in unit capacities, the number of stochastic parameters and the level of stochastic uncertainty considered, collectively, make the present scenario a complex and challenging problem to solve. In this particular case, the problem aims to find an optimal scheduling and control strategy that can accommodate a large number of stochastic realizations in the parameters involving a relatively large number of processing units. Hence, the consideration of the availability of multiple units capable of realizing the same task have resulted in a more conversative scheduling and control solution than that obtained from scenario 1 (instance B). However, this is problem-specific and directly depends on the way that uncertainty manifests within the system.

IV.4. Summary

This chapter presented an extended and improved version of the decomposition back-off algorithm presented in Chapter III for integration of scheduling and control. The proposed methodology implements correlations that allow for a search of scheduling decisions that better reflect the effects that the back-off terms have on the dynamics of the system. Parametric uncertainty is still approximated through Monte Carlo sampling techniques. The effects of the uncertainty on the system are analyzed and used to calculate back-off terms. Back-off terms are introduced in the formulation to force the system to find control profiles, unit operation times and scheduling decisions that are dynamically feasible under the presence of uncertainty. A case study illustrating the performance of the algorithm is presented. The results show how uncertainty has a strong effect on the solutions and how having multiple units with uncertain parameters interacting quickly becomes computationally expensive. Nonetheless, the scheduling decisions make better estimations of the dynamic operation of the plant thanks to the linear correlations considered in the present framework in the presence of stochastic realizations in the uncertain parameters.

Chapter V. Conclusions & Recommendations

This chapter summarizes the contributions reported in this thesis. Several points regarding improvements to the proposed methodologies and future research opportunities are also presented.

Two new back-off decomposition algorithms for the integration of scheduling and control of multi-unit, multi-product chemical batch plants under stochastic uncertainty were presented. The key idea in both algorithms is to introduce a formulation that seeks for the optimal unit processing times and control decisions using back-off terms, which reflect process variability under uncertainty. In Algorithm B, correlations are introduced to allow the scheduling decisions to account for the backed-off dynamics of the system. It was shown that uncertainty in the model does significantly affect the scheduling plan and the control decisions, i.e., the backed-off solutions differ from the nominal solution.

Both algorithms possess a certain degree of freedom as many parameters are user-defined and can be tuned to allow the algorithm to search for dynamically feasible solutions. Due to the stochastic nature of the process, certain realizations and combinations of the uncertain parameters may never yield a dynamically feasible solution with the current plant design. A sensitivity analysis can be used to help define the zones of unfeasibility, so that an appropriate course of action can be taken to prevent most of the unfeasible scenarios from happening (e.g., change design parameters, add extra constraints, redesign a unit, etc.). The stochastic simulation step represents the current bottleneck of the proposed algorithms due to the high number of scenarios that need to be explored for the generation of data for the calculation of the back-off terms. Improvements on the propagation of uncertainty must be explored to reduce the total CPU time, especially as the issue aggravates as the number of uncertain elements increases. Although the fact that the convergence criterion of the stochastic simulation step can be relaxed, a negative impact in the quality of the back-off terms to correctly represent the uncertainty of the system should be expected.

The introduction of the back-off terms causes a decrease in the overall process economics when the backed-off case is compared to the nominal case. However, the backed-off scheduling plan and backed-off control decisions can accommodate uncertainty. Moreover, as the level of variability considered (λ) in the back-off terms increases, the overall process revenue also increases but remains lower than the revenue obtained for the nominal case. Also, because of the presence of the back-off terms the quality of the product increases in order to offset the effects of the parametric uncertainty. The implementation of the operation regimes found by the proposed approach would allow for specification of products with a quality equal or slightly higher than that is required by the market to accommodate parametric uncertainty. Note that the maximum value λ can take is problem specific. The present methodology can be used by the enterprise management to decide if the potential loss in profits due to the production of an over specified quality product who needs to be produced to accommodate parametric uncertainty is acceptable or if costly and time-consuming laboratory testing may be need to be performed to determine better estimates for the uncertain parameters thereby reducing the uncertainty effects in the batch plant.

The results from the case studies show that the present methodologies remain computationally attractive and can improve the profits of chemical batch plants. This is due to the optimization of unit operation times that can generate control decisions able to deal with stochastic parametric uncertainty, while obtaining a scheduling plant that is able to accommodate the corresponding process dynamics. In particular, the results of Algorithm B show that it does become quite computationally taxing and returns conservative solutions when: I) the number of uncertain parameters considered is considerably increased, II) a larger number of units (each with a different capacity) available per process are considered, and III) there is a high degree of interaction among the units considered in the batch plant configuration (i.e., a unit allocation is between 2 or more jobs). To decrease the computational time required to achieve a solution, the convergence criterions can be relaxed at the expense of a loss in the quality of the solution.

The main advantage that the second methodology (Algorithm B) has over the first methodology (Algorithm A) is that its scheduling decisions better reflect the dynamics of chemical processes of the plant under uncertainty thanks to the introduction of linear correlations that allow the calculation of unit operation times and costs as a function of the quantity of processed material. Thus, the information exchange between layers is more coherent, at the expense of increased CPU costs. Another advantage is that control decisions and unit operation time are obtained simultaneously (Algorithm A does this task sequentially). Overall, both methodologies offer trade-offs in terms of the size of the problem and the available computational resources.

V.1. Future research opportunities

The research findings from this thesis also led to new potential research avenues. Improvements over the current proposed methodologies and future research paths are listed as follows:

- Extend the present approaches to chemical plants operating in a continuous domain or plants that are constituted by a mixture of batch and continuous processes. This may require a reformulation of certain aspects and change of certain considerations (i.e., a scheduling formulation that can accommodate continuous, batch and semi-batch operations). The ideas developed by Andrés-Martínez & Ricardez-Sandoval[15] about a switched system could be taken as the basis to deal with the dynamics of the system when determining optimal control profiles and a *dynamic* scheduling.
- In this work, back-off term calculations assume that the parametric uncertainty can be described using Gaussian probabilistic density distributions. It would be interesting to see how the quality of the solution is affected by considering different probabilistic density distributions. The case of mixed distributions would also be interesting to analyze (i.e., each parameter is described by a different function). The study of Kimaev & Ricardez-Sandoval[66] suggests that the implementation of polynomial chaos expansions may be an alternative way to simulate the propagation of uncertainty in the model.
- Currently, linear correlations are used on the second proposed algorithm (algorithm B), as linear equations portrayed very well the behavior of the variables of interest of the case study. Future research can consider non-linear correlations; while this may improve the accuracy in the scheduling decisions, this addition may also increase the complexity in the calculations since the scheduling problem becomes an MINLP.
- The consideration of another hierarchical layer of management in integration methodologies would allow the development of better solutions for chemical systems. With the addition of the Planning layer, plans can be made such that the plant operation can account for the uncertainties that exist in long time periods. Integrating the Design layer would allow for the conception of technologies that can deal with known process uncertainties (proactive actions). Nonetheless, problem formulations will have to account for the challenges that each layer faces individually (e.g., optimal equipment

sizing) and for the challenges caused by considering multiple layers (e.g., decisions that need to respond to many time horizons that may greatly differ in time scales).

- ✐ Machine learning methods have been successfully implemented into a wide range of applications involving the discovery of new functional materials[74–76]; conventional and emerging chemical and energy systems[76–80]; process scheduling[81–84] and process control[85–87]. The application of these methods to address the optimal integration of scheduling and control decisions is still lacking. Thus, it is recommended for future work to investigate the potential application of these methods to perform optimal scheduling and control of systems under stochastic uncertainty.
- ✐ While scheduling and control strategies have been implemented for a wide variety of chemical engineering applications, there are still a lack of studies involving emerging systems such as CO_2 capture[88–91], chemical looping combustion[92,93] and process intensification systems[9] such as catalytic distillation columns[41]. Scheduling and control studies can therefore provide new insights into the optimal operation of those systems. Therefore, this work recommends this subject as a potential area of future research.

Bibliography

1. Valdez-Navarro, Y. I. & Ricardez-Sandoval, L. A. A Novel Back-off Algorithm for Integration of Scheduling and Control of Batch Processes under Uncertainty. *Ind. Eng. Chem. Res.* **58**, 22064–22083 (2019).
2. Rodríguez Vera, H. U. & Ricárdez-Sandoval, L. A. Integration of Scheduling and Control Under Stochastic Parametric Uncertainty with Varying Unit Operation Times for Chemical Batch Plants: A Back-Off Approach. *IFAC-PapersOnLine* **54**, 578–583 (2021).
3. Rodríguez Vera, H. U. & Ricardez-Sandoval, L. A. Integration of Scheduling and Control for Chemical Batch Plants under Stochastic Uncertainty: A Back-off Approach. *Ind. Eng. Chem. Res.(Under Revision)*
4. Rafiei, M. & Ricardez-Sandoval, L. A. New frontiers, challenges, and opportunities in integration of design and control for enterprise-wide sustainability. *Comput. Chem. Eng.* **132**, (2020).
5. Pistikopoulos, E. N. & Diangelakis, N. A. Towards the integration of process design, control and scheduling: Are we getting closer? *Comput. Chem. Eng.* **91**, 85–92 (2016).
6. Dias, L. S. & Ierapetritou, M. G. From process control to supply chain management: An overview of integrated decision making strategies. *Comput. Chem. Eng.* **106**, 826–835 (2017).
7. Davis, J., Edgar, T., Porter, J., Bernaden, J. & Sarli, M. Smart manufacturing, manufacturing intelligence and demand-dynamic performance. *Comput. Chem. Eng.* **47**, 145–156 (2012).
8. Grossmann, I. Enterprise-wide optimization: A new frontier in process systems engineering. *AIChE J.* **51**, 1846–1857 (2005).
9. López-Guajardo, E. A. *et al.* Process intensification 4.0: A new approach for attaining new, sustainable and circular processes enabled by machine learning. *Chem. Eng. Process. - Process Intensif.* 108671 (2021) doi:10.1016/j.cep.2021.108671.
10. Dias, L. S. & Ierapetritou, M. G. Integration of planning, scheduling and control problems using data-driven feasibility analysis and surrogate models. *Comput. Chem. Eng.* **134**, (2020).
11. Baldea, M. & Harjunkoski, I. Integrated production scheduling and process control: A systematic review. *Computers and Chemical Engineering* vol. 71 377–390 (2014).
12. Engell, S. & Harjunkoski, I. Optimal operation: Scheduling, advanced control and their integration. *Comput. Chem. Eng.* **47**, 121–133 (2012).
13. Koller, R. W. & Ricardez-Sandoval, L. A. A dynamic optimization framework for integration of design, control and scheduling of multi-product chemical processes under disturbance and uncertainty. *Comput. Chem. Eng.* **106**, 147–159 (2017).
14. Koller, R. & Ricardez-Sandoval, L. Integration of Design, Control and Scheduling: A Dynamic Optimization Framework for Multi-product Chemical Processes under Disturbances and Uncertainty. *Comput. Aided Chem. Eng.* **40**, 2077–2082 (2017).
15. Andrés-Martínez, O. & Ricardez-Sandoval, L. A. A switched system formulation for optimal integration of scheduling and control in multi-product continuous processes. *J. Process Control* **106**, 94–109 (2021).
16. Mitra, K., Gudi, R. D., Patwardhan, S. C. & Sardar, G. Resiliency issues in integration of scheduling and control. *Ind. Eng. Chem. Res.* **49**, 222–235 (2010).
17. Beal, L. D. R., Petersen, D., Grimsman, D., Warnick, S. & Hedengren, J. D. Integrated

scheduling and control in discrete-time with dynamic parameters and constraints. *Comput. Chem. Eng.* **115**, 361–376 (2018).
18. Koller, R. W., Ricardez-Sandoval, L. A. & Biegler, L. T. Stochastic back-off algorithm for simultaneous design, control, and scheduling of multiproduct systems under uncertainty. *AIChE J.* **64**, 2379–2389 (2018).
19. Zhuge, J. & Ierapetritou, M. G. Integration of scheduling and control with closed loop implementation. *Ind. Eng. Chem. Res.* **51**, 8550–8565 (2012).
20. Chu, Y. & You, F. Integration of scheduling and dynamic optimization of batch processes under uncertainty: Two-stage stochastic programming approach and enhanced generalized benders decomposition algorithm. *Ind. Eng. Chem. Res.* **52**, 16851–16869 (2013).
21. Chu, Y. & You, F. Integrated scheduling and dynamic optimization of complex batch processes with general network structure using a generalized benders decomposition approach. *Ind. Eng. Chem. Res.* **52**, 7867–7885 (2013).
22. Gutiérrez-Limón, M. A., Flores-Tlacuahuac, A. & Grossmann, I. E. MINLP formulation for simultaneous planning, scheduling, and control of short-period single-unit processing systems. *Ind. Eng. Chem. Res.* **53**, 14679–14694 (2014).
23. Chu, Y. & You, F. Integrated planning, scheduling, and dynamic optimization for batch processes: MINLP model formulation and efficient solution methods via surrogate modeling. *Ind. Eng. Chem. Res.* **53**, 13391–13411 (2014).
24. Rafiei, M. & Ricardez-Sandoval, L. A. A trust-region framework for integration of design and control. *AIChE J.* **66**, 1–15 (2020).
25. Rafiei, M. & Ricardez-Sandoval, L. A. Integration of design and control for industrial-scale applications under uncertainty: a trust region approach. *Comput. Chem. Eng.* **141**, 107006 (2020).
26. Rafiei, M. & Ricardez-Sandoval, L. A. Stochastic Back-Off Approach for Integration of Design and Control under Uncertainty. *Ind. Eng. Chem. Res.* **57**, 4351–4365 (2018).
27. Palma-Flores, O., Rafiei, M. & Ricardez-Sandoval, L. A. On the implementation of nonlinear model predictive control for simultaneous design and control using a back-off approach. *IFAC-PapersOnLine* **53**, 11551–11556 (2020).
28. Bernal, D. E., Carrillo-Diaz, C., Gómez, J. M. & Ricardez-Sandoval, L. A. Simultaneous design and control of catalytic distillation columns using comprehensive rigorous dynamic models. *Ind. Eng. Chem. Res.* **57**, 2587–2608 (2018).
29. Kookos, I. K. & Perkins, J. D. An algorithm for simultaneous process design and control. *Ind. Eng. Chem. Res.* **40**, 4079–4088 (2001).
30. Kookos, I. K. & Perkins, J. D. Chapter B2 The back-off approach to simultaneous design and control. *Comput. Aided Chem. Eng.* **17**, 216–238 (2004).
31. Patilas, C. S. & Kookos, I. K. A novel approach to the simultaneous design & control problem. *Chem. Eng. Sci.* **240**, 116637 (2021).
32. De Carvalho, R. F. & Alvarez, L. A. Simultaneous Process Design and Control of the Williams-Otto Reactor Using Infinite Horizon Model Predictive Control. *Ind. Eng. Chem. Res.* **59**, 15979–15989 (2020).
33. Pistikopoulos, E. N. Uncertainty in process design and operations. *Comput. Chem. Eng.* **19**, 553–563 (1995).
34. Kookos, I. K. & Perkins, J. D. Control structure selection based on economics: Generalization of the back-off methodology. *AIChE J.* **62**, 3056–3064 (2016).
35. Patilas, C. S. & Kookos, I. K. A quadratic approximation of the back-off methodology for the control structure selection problem. *Comput. Chem. Eng.* **143**, 107114 (2020).
36. Patilas, C. S. & Kookos, I. K. Algorithmic Approach to the Simultaneous Design and Control Problem. *Ind. Eng. Chem. Res.* **60**, 14271–14281 (2021).

37. Shobrys, D. E. & White, D. C. Planning, scheduling and control systems: Why cannot they work together. *Comput. Chem. Eng.* **26**, 149–160 (2002).
38. Jaskó, S., Skrop, A., Holczinger, T., Chován, T. & Abonyi, J. Development of manufacturing execution systems in accordance with Industry 4.0 requirements: A review of standard- and ontology-based methodologies and tools. *Comput. Ind.* **123**, (2020).
39. Dias, L. S. & Ierapetritou, M. G. Integration of scheduling and control under uncertainties: Review and challenges. *Chem. Eng. Res. Des.* **116**, 98–113 (2016).
40. Ricardez Sandoval, L. A., Budman, H. M. & Douglas, P. L. Simultaneous design and control of processes under uncertainty: A robust modelling approach. *J. Process Control* **18**, 735–752 (2008).
41. Liñán, D. A. & Ricardez-Sandoval, L. A. Optimal design and dynamic transitions of multitask catalytic distillation columns: A Discrete-Steepest Descend Framework. *Chem. Eng. Process. - Process Intensif.* 108655 (2021) doi:10.1016/j.cep.2021.108655.
42. Ierapetritou, M. G. & Floudas, C. A. Effective continuous-time formulation for short-term scheduling. 1. Multipurpose batch processes. *Ind. Eng. Chem. Res.* **37**, 4341–4359 (1998).
43. Floudas, C. A. & Lin, X. Continuous-time versus discrete-time approaches for scheduling of chemical processes: A review. *Comput. Chem. Eng.* **28**, 2109–2129 (2004).
44. Méndez, C. A., Cerdá, J., Grossmann, I. E., Harjunkoski, I. & Fahl, M. State-of-the-art review of optimization methods for short-term scheduling of batch processes. *Comput. Chem. Eng.* **30**, 913–946 (2006).
45. Kondili, E., Pantelides, C. C. & Sargent, R. W. H. A general algorithm for short-term scheduling of batch operations-I. MILP formulation. *Comput. Chem. Eng.* **17**, 211–227 (1993).
46. Barbosa-Póvoa, A. P. F. D. & Pantelides, C. C. Design of multipurpose plants using the resource-task network unified framework. *Comput. Chem. Eng.* **21**, (1997).
47. Pinto, T., Barbósa-Póvoa, A. P. F. D. & Novais, A. Q. Design of multipurpose batch plants: A comparative analysis between the STN, m-STN, and RTN representations and formulations. *Ind. Eng. Chem. Res.* **47**, 6025–6044 (2008).
48. Biegler, L. T. *Nonlinear Programming: Concepts, Algorithms, and Applications to Chemical Processes*. (2010). doi:10.1137/1.9780898719383.
49. Edgar, T. F. & Himmelblau, D. M. Optimization of chemical process. *Mcgraw Hill Chem. Eng. Ser.* 672 (2001) doi:10.1016/j.compchemeng.2009.11.012.
50. Bhatia, T. & Biegler, L. T. Dynamic optimization in the design and scheduling of multiproduct batch plants. *Ind. Eng. Chem. Res.* **35**, 2234–2246 (1996).
51. Mishra, B. V., Mayer, E., Raisch, J. & Kienle, A. Short-term scheduling of batch processes. A comparative study of different approaches. *Ind. Eng. Chem. Res.* **44**, 4022–4034 (2005).
52. Nyström, R. H., Franke, R., Harjunkoski, I. & Kroll, A. Production campaign planning including grade transition sequencing and dynamic optimization. *Comput. Chem. Eng.* **29**, 2163–2179 (2005).
53. Nyström, R. H., Harjunkoski, I. & Kroll, A. Production optimization for continuously operated processes with optimal operation and scheduling of multiple units. *Comput. Chem. Eng.* **30**, 392–406 (2006).
54. Flores-Tlacuahuac, A. & Grossmann, I. E. Simultaneous cyclic scheduling and control of a multiproduct CSTR. *Ind. Eng. Chem. Res.* **45**, 6698–6712 (2006).
55. Flores-Tlacuahuac, A. & Grossmann, I. E. Simultaneous cyclic scheduling and control of tubular reactors: Parallel production lines. *Ind. Eng. Chem. Res.* **50**, 8086–8096

(2011).
56. Touretzky, C. R. & Baldea, M. Integrating scheduling and control for economic MPC of buildings with energy storage. *J. Process Control* **24**, 1292–1300 (2014).
57. Touretzky, C. R. & Baldea, M. A hierarchical scheduling and control strategy for thermal energy storage systems. *Energy Build.* **110**, 94–107 (2016).
58. Zhuge, J. & Ierapetritou, M. G. Integration of scheduling and control for batch processes using multi-parametric model predictive control. *AIChE J.* **60**, 3169–3183 (2014).
59. Du, J., Park, J., Harjunkoski, I. & Baldea, M. A time scale-bridging approach for integrating production scheduling and process control. *Comput. Chem. Eng.* **79**, 59–69 (2015).
60. Zhuge, J. & Ierapetritou, M. G. A Decomposition Approach for the Solution of Scheduling Including Process Dynamics of Continuous Processes. *Ind. Eng. Chem. Res.* **55**, 1266–1280 (2016).
61. Kelley, M. T., Pattison, R. C., Baldick, R. & Baldea, M. An efficient MILP framework for integrating nonlinear process dynamics and control in optimal production scheduling calculations. *Comput. Chem. Eng.* **110**, 35–52 (2018).
62. Caspari, A., Tsay, C., Mhamdi, A., Baldea, M. & Mitsos, A. The integration of scheduling and control: Top-down vs. bottom-up. *J. Process Control* **91**, 50–62 (2020).
63. Simkoff, J. M. & Baldea, M. Production scheduling and linear MPC: Complete integration via complementarity conditions. *Comput. Chem. Eng.* **125**, 287–305 (2019).
64. Zhuge, J. & Ierapetritou, M. G. An integrated framework for scheduling and control using fast model predictive control. *AIChE J.* **61**, 3304–3319 (2015).
65. Santander, O. & Baldea, M. Control-aware batch process scheduling. *Comput. Chem. Eng.* **152**, 107360 (2021).
66. Kimaev, G. & Ricardez-Sandoval, L. A. A comparison of efficient uncertainty quantification techniques for stochastic multiscale systems. *AIChE J.* **63**, 3361–3373 (2017).
67. Li, Z. & Ierapetritou, M. G. Process scheduling under uncertainty using multiparametric programming. *AIChE J.* **53**, 3183–3203 (2007).
68. Narraway, L. T., Perkins, J. D. & Barton, G. W. Interaction between process design and process control: economic analysis of process dynamics. *J. Process Control* **1**, 243–250 (1991).
69. Bansal, V., Sakizlis, V., Ross, R., Perkins, J. D. & Pistikopoulos, E. N. New algorithms for mixed-integer dynamic optimization. *Comput. Chem. Eng.* **27**, 647–668 (2003).
70. Mehta, S. & Ricardez-Sandoval, L. A. Integration of Design and Control of Dynamic Systems under Uncertainty: A New Back-Off Approach. *Ind. Eng. Chem. Res.* **55**, 485–498 (2016).
71. Palma-Flores, O. & Ricardez-Sandoval, L. A. Simultaneous design and nonlinear model predictive control under uncertainty: A back-off approach. *J. Process Control* **110**, 45–58 (2022).
72. Zavala, V. M. pyomo. dae: A Modeling and Automatic Discretization Framework for Optimization with Differential and Algebraic Equations. (2016).
73. Hart, W. E. *et al. Pyomo — Optimization Modeling in Python*. vol. 67 (Springer International Publishing, 2017).
74. Guan, Y. *et al.* Machine Learning in Solid Heterogeneous Catalysis: Recent Developments, Challenges and Perspectives. *Chem. Eng. Sci.* **248**, 117224 (2021).
75. Hayashi, Y. *et al.* Application of machine learning to a material library for modeling of relationships between material properties and tablet properties. *Int. J. Pharm.* **609**, 121158 (2021).

76. Vivanco-Benavides, L. E., Martínez-González, C. L., Mercado-Zúñiga, C. & Torres-Torres, C. Machine learning and materials informatics approaches in the analysis of physical properties of carbon nanotubes: A review. *Comput. Mater. Sci.* **201**, (2022).
77. Chaffart, D. & Ricardez-Sandoval, L. A. Optimization and control of a thin film growth process: A hybrid first principles/artificial neural network based multiscale modelling approach. *Comput. Chem. Eng.* **119**, 465–479 (2018).
78. Kimaev, G. & Ricardez-Sandoval, L. A. Nonlinear model predictive control of a multiscale thin film deposition process using artificial neural networks. *Chem. Eng. Sci.* **207**, 1230–1245 (2019).
79. Mao, J., Miao, J., Lu, Y. & Tong, Z. Machine learning of materials design and state prediction for lithium ion batteries. *Chinese J. Chem. Eng.* **37**, 1–11 (2021).
80. Lu, Z. Computational discovery of energy materials in the era of big data and machine learning: A critical review. *Mater. Reports Energy* **1**, 100047 (2021).
81. Chen, R., Yang, B., Li, S., Wang, S. & Cheng, Q. An Effective Multi-population Grey Wolf Optimizer based on Reinforcement Learning for Flow Shop Scheduling Problem with Multi-machine Collaboration. *Comput. Ind. Eng.* **162**, 107738 (2021).
82. Yamashiro, H. & Nonaka, H. Estimation of processing time using machine learning and real factory data for optimization of parallel machine scheduling problem. *Oper. Res. Perspect.* **8**, 100196 (2021).
83. Wang, H., Yan, Q. & Zhang, S. Integrated scheduling and flexible maintenance in deteriorating multi-state single machine system using a reinforcement learning approach. *Adv. Eng. Informatics* **49**, 101339 (2021).
84. Zhang, L., Deng, Q., Lin, R., Gong, G. & Han, W. A combinatorial evolutionary algorithm for unrelated parallel machine scheduling problem with sequence and machine-dependent setup times, limited worker resources and learning effect. *Expert Syst. Appl.* **175**, 114843 (2021).
85. Singer, G. & Cohen, Y. A framework for smart control using machine-learning modeling for processes with closed-loop control in Industry 4.0. *Eng. Appl. Artif. Intell.* **102**, 104236 (2021).
86. Alhajeri, M. S., Wu, Z., Rincon, D., Albalawi, F. & Christofides, P. D. Machine-learning-based state estimation and predictive control of nonlinear processes. *Chem. Eng. Res. Des.* **167**, 268–280 (2021).
87. Wang, P., Yang, Y. & Shayesteh, N. Process modeling in laser powder bed fusion towards defect detection and quality control via machine learning : The state-of-the-art and research challenges. *J. Manuf. Process.* **73**, 961–984 (2022).
88. He, Z., Sahraei, M. H. & Ricardez-Sandoval, L. A. Flexible operation and simultaneous scheduling and control of a CO2 capture plant using model predictive control. *Int. J. Greenh. Gas Control* **48**, 300–311 (2016).
89. Bui, M. *et al.* Dynamic modelling and optimisation of flexible operation in post-combustion CO2 capture plants-A review. *Comput. Chem. Eng.* **61**, 245–265 (2014).
90. Abdul Manaf, N., Cousins, A., Feron, P. & Abbas, A. Dynamic modelling, identification and preliminary control analysis of an amine-based post-combustion CO2 capture pilot plant. *J. Clean. Prod.* **113**, 635–653 (2016).
91. Patrón, G. D. & Ricardez-Sandoval, L. An integrated real-time optimization, control, and estimation scheme for post-combustion CO2 capture. *Appl. Energy* **308**, 118302 (2022).
92. Lucio, M. & Ricardez-Sandoval, L. A. Dynamic modelling and optimal control strategies for chemical-looping combustion in an industrial-scale packed bed reactor. *Fuel* **262**, 116544 (2020).
93. Francisco Morgado, J., Cloete, S., Morud, J., Gurker, T. & Amini, S. Modelling study

of two chemical looping reforming reactor configurations: looping vs. switching. *Powder Technol.* **316**, 599–613 (2017).