Junior Management Science 5(1) (2020) 50-80

Junior Management Science

journal homepage: www.jums.academy

Stochastic Optimization of Bioreactor Control Policies Using a Markov Decision Process Model

Quirin Stockinger

Technische Universität München

Abstract

Biopharmaceuticals are the fastest-growing segment of the pharmaceutical industry. Their manufacture is complicated by the uncertainty exhibited therein. Scholars have studied the planning and operation of such production systems under some uncertainties, but the simultaneous consideration of fermentation and resin yield uncertainty is lacking so-far. To study the optimal operation of biopharmaceutical production and purification systems under these uncertainties, a stochastic, dynamic approach is necessary. This thesis provides such a model by extending an existing discrete state-space, infinite horizon Markov decision process model of upstream fermentation.

Tissue Plasminogen Activator fermentation and chromatography was implemented. This example was used to discuss the optimal policy for operating different fermentation setups. The average per-cycle operating profit of a serial setup was 1,272 $; the parallel setup produced negative average rewards. Managerial insights were derived from a comparison to a basic, titer maximizing policy and process sensitivities. In conclusion, the integrated stochastic optimization of biopharma production and purification control aids decision making. However, the model assumptions pose room for further studies.

Keywords: Markov decision process; biopharmaceuticals production; fermentation uncertainty; chromatography resin; stochastic performance decay.

1. Introduction

There are currently 316 active biopharmaceutical ingredients available on the market (Walsh, 2018). Within the pharmaceutical industry, so-called biopharmaceuticals represent the fastest-growing segment, generating more than 160 bn€ in annual revenue (Otto et al., 2014). These biopharmaceuticals are not the result of chemical synthesizing, but their manufacture involves the fermentation of bacterial or mammalian cells. To capture the produced proteins, the fermentation is followed by a process of purification. However, the control of biopharmaceutical production and purification systems is highly complex because of the inherent technical uncertainties (Farid et al., 2005). This contribution studies the optimal, simultaneous control of biopharmaceutical fermentation and purification processes.

Biopharma production, on a high level, is a two-stage process: during the upstream process (USP), a cell culture is grown in sequential, volume-increasing bioreactor media wherein the fermentation environment is most commonly controlled by way of feeding substrates such as nutrients into the reactors (fed-batch fermentation) and changing the temperature, pressure, and pH-value. The culture's protein production is induced in the final bioreactor of the sequence, i.e., the "production reactor," by changing the nutrient concentration in the medium accordingly. When enough product has been fermented, the protein of interest must be separated from the medium during the downstream process (DSP). The most common technique of purification is chromatography (Liu et al., 2014), but it is also the "most expensive part of the downstream process" (Nweke et al., 2018, p. 992). Because chromatography resins' capacity to bind proteins is uncertain (Farid et al., 2005) and deteriorates over time (Jiang et al., 2009), the exchange of spent chromatography resin constitutes operational complexity in the DSP. While there may be additional filtration steps and intermediate storage steps, the upstream and downstream processes are highly interdependent, and their simultaneous control poses significant room for research.

The need for optimization of biopharmaceutical production is an ongoing topic of discussion in practice. A recent report by consultancy McKinsey & Company, for example, cites "finding ways to improve the performance of the production

DOI: https://dx.doi.org/10.5282/jums/v5i1pp50-80

process" as one of the critical operational considerations that biopharmaceutical manufacturers must ponder to succeed in the face of adverse market realities such as changing payor behavior and competition from biosimilars (Otto et al., 2014, p. 6). Even though small scheduling errors can already lead to huge financial downside from lost batches due to the high value of proteins (Schmidt, 1996), operations research techniques are still only sparsely adopted in the industry (Martagan et al., 2018).

Nevertheless, academic interest in the study of controlling biopharmaceutical production has been highly active. The planning of up- and downstream capacities and processes, for example, has been discussed in literature, including chromatography column sizing (Allmendinger et al., 2014), maintenance of spent resins (Liu et al., 2014), production under chromatography yield (Liu et al., 2016) and product titer uncertainties (Lakhdar et al., 2006), and purification capacities (Siganporia et al., 2014). The control of upstream fermentation has largely focused on maximizing product concentration in a single bioreactor (Banga et al., 1997; Pandian and Noel, 2018; Peroni et al., 2005;Rocha et al., 2014; Saucedo and Karim, 1997). While these endeavors model cell-level kinetics, they often do not consider the financial trade-offs of system-level decision making, i.e., integrated decisions across all involved process steps. However, as early as 1996 (Schmidt, 1996), and more recently (Martagan et al., Accepted/In press; Martagan et al., 2016; Martagan et al., 2018), the literature on the system-level control of production and purification equipment has aimed at filling this void. Some scholars argue that purely maximizing the protein concentration during fermentation may not yield economically optimal results when considering the associated operating and purification costs (Martagan et al., 2018).

While Martagan et al. (Accepted/In press) studied simultaneous system-level control of protein production and purification processes, they abstracted upstream decision making to a single decision of how much protein to produce. Furthermore, they did not consider the issue of chromatography resin performance decay, as introduced by Liu et al. (2014). Nevertheless, its consideration is relevant for practice because resin material is a major driver of downstream operating costs (Farid, 2007). To the best of my knowledge, no existing paper has simultaneously considered the biopharmaceutical fermentation and optimal resin exchange schedule sub-problems under uncertain fermentation and stochastic performance decay. Existing literature, therefore, doesn't conclusively answer some outstanding questions about the influence of the aforementioned uncertainties. What is the optimal, simultaneous control policy for the USP and DSP? When is chromatography resin exchanged under stochastic decay? When under different minimum allowed resin capacities and different resin costs? How does the consideration of two parallel production reactors change the optimal policy?

This work provides a dynamic stochastic model spanning both upstream and downstream operations which is used to answer these questions. Using this dynamic program, this contribution analyzes a practice-representative production and purification process using the example of the recombinant protein Tissue Plasminogen Activator (TPA), a prominent product of the early biopharma industry (Datar et al., 1993) which is still relevant today (Johnston, 2010). On a system-level, decisions about the production of TPA and its purification must be made simultaneously. During the USP, decisions about how long to grow the culture before converting it into its protein-producing state and when to harvest the TPA from the medium are considered. A linear accumulation of TPA in the production medium and possible batch failure due to contamination are assumed. The states of the observed production reactors are assumed to represent physiological states during the culture's lifecycle, e.g., growth, production, and decay, but cell-level kinetics are not modeled. Because there is no intermediate storage, and both process steps are highly interdependent, harvesting of TPA constitutes two necessarily simultaneous decisions: to harvest the production medium and to accept the medium into the first chromatography step. For the purification of a single batch, five purification cycles in the first chromatography column are assumed to take place within one decision epoch. After each purified batch, the performance of the resin, i.e., what fraction of the TPA in the medium it can bind, deteriorates stochastically. This gives rise to the need for maintenance actions related to the exchange of spent resin. Maintenance activities are assumed to take one decision epoch due to their short duration. Because the first chromatography column can be regarded as the bottleneck of a multi-step chromatography process, only the first chromatography step is considered in the presented model. Analogously, only the production reactor of a seed-train is considered as its bottleneck. To test hypotheses about the parallelization of production, the provided model is extended by a second, parallel production reactor in the same seed-train and scenarios are analyzed. Sets of states for which the same control actions are optimal are discussed because of the approach's demonstrated value in previous research (Martagan et al., 2018).

By building on the theoretical foundations of Schmidt (1996), this work contributes a framework for the system-level study of simultaneous decision making in the USP and DSP under uncertainty. Contrary to existing models, it allows for the study of the optimal operation of parallel production reactors and chromatography resin maintenance. Furthermore, it explicitly models protein accumulation during production. This thesis also contributes optimal control policies for the production and purification of TPA. By studying resin exchange policies under varying process conditions, this thesis builds on Liu et al. (2014) understanding of what influences the carrying out of costly maintenance activities. Furthermore, this thesis argues for the business case of the stochastic optimization of integrated production and purification control compared to simple upstream titer maximization. The derived decision spaces aid managerial and operational decision making under the uncertain environment of biopharmaceutical production.

The remainder of this thesis is organized as follows:

Chapter 2 reviews the problem context and surveys existing literature on biopharmaceutical production. Chapter 3 details the problem characteristics and this work's research questions. Chapter 4 outlines Markov decision processes (MDPs) as the solution approach and introduces the case study of TPA production. Chapter 5 describes the developed model and Chapter 6 the results of the numerical case study. Chapter 7 concludes this work and provides an outlook for future research endeavors.

2. Review of Literature and Research

To the best of my knowledge, no existing paper studies the simultaneous, system-level control of upstream fed-batch fermentation and downstream resin exchange. In this chapter, existing literature and research on the optimization of biopharmaceutical production schedules in the light of uncertainty are reviewed. First, an introduction to the biopharmaceutical production process and their academic and economic relevance are provided and the literature review methodology summarized (Section 2.1). Next, the control of fed-batch fermentation processes based on cell-level kinetics is reviewed (Section 2.2). This is extended by a survey of the issue of planning biopharma production capacities and activities, with a focus on the inherent stochasticity of the process (Section 2.3). In Section 2.4, the middle ground between the two prior abstraction levels, optimal control policies of integrated production systems under biological uncertainties, is reviewed. Concluding this literature overview, the concrete research gaps in existing research are identified and this thesis is motivated (Section 2.5).

2.1. Context of literature research

This section provides a brief overview of the biopharmaceutical market and defines common production processes. Additionally, the methodology of the literature review is described.

2.1.1. Biopharmaceutical production

Since the first biopharmaceutical drug being commercialized in 1982, the market has continuously grown (Figure 1). Between 2014 and 2018 alone, 129 distinct biopharmaceuticals have been commercialized across the United States and European Union. Taking the 58 withdrawn active ingredients into account, currently, 316 biopharmaceutical active ingredients are available (Walsh, 2018).

Regardless of the topic's academic and commercial relevance, the lack of a uniform definition of terms has been lamented (Rader, 2005). Within the pharmaceutical industry, two terms need to be distinctly defined to allow for a concise discussion of the topic. Ordinary "drugs" are manufactured by "chemical (non-biological) means and involving small molecules", whereas "biopharmaceuticals" are "manufactured by biotechnology methods and involving complex biological molecules" (Rader, 2005). Biopharmaceuticals may be "produced from cultures of eukaryotic or prokaryotic cells, isolated from natural sources, or made by synthetic methods" (Jagschies et al., 2018, p. 59).

The production process of biopharmaceuticals, on a high level, is separated into two sequential, highly interdependent phases: an upstream process (USP) and a downstream process (DSP). During the USP, the active pharmaceutical ingredient (API) is synthesized by cultivating the living cells under controlled nutrient conditions, i.e., fermentation. Upon reaching the required API quantity, recovery and filtration of the API must take place in the DSP before packaging and shipping of final products (Jagschies et al., 2018, p. 76). This is necessary because of impurities produced alongside the API during fermentation. The purification can directly follow the fermentation process or be postponed (Siganporia et al., 2014). One of the most common yet highly cost-intensive process steps during product recovery is chromatography (Jiang et al., 2009; Liu et al., 2014). During chromatography, the product of interest is separated from the medium and impurities based on their physiological differences. For this task, so-called chromatography columns hold resins which either bind the product of interest or the impurities (Martagan et al., Accepted/In press). Additional steps between USP and DSP, like intermediate storage and filtration, are possible. However, these are not in-scope for this contribution.

Each phase may consist of a series of bioreactors and chromatography steps, respectively. During the USP, bioreactors of increasing volumes may be used to grow an initially small amount of cell culture to commercial production scale. Such a "train" of sequentially interconnected bioreactors may be referred to as a "seed train" (Jagschies et al., 2018, p. 632). During the DSP, between two and six sequential chromatography steps may be required to meet the purity demands of the desired product (Martagan et al., 2018). A schematic of a biopharmaceutical production process using a single seed-train of three bioreactors of increasing volumes and three sequential chromatography steps is provided in Figure 2. To improve the utilization of the purification equipment, different process set-ups are possible. Appendix 1, for example, shows two serial upstream seed-trains feeding into one downstream process. It may, however, be more capacity- and cost-feasible to inoculate multiple parallel production reactors from one seed-train (Jagschies et al., 2018, 653f), visualized in Appendix 2.

Bioreactors can be operated in two modes: batch processing and continuous processing. Batch processing allows distinct production periods and lot sizes while continuous production allows continuous harvesting (Siganporia et al., 2014). Within the batch process category, two prominent operating modes exist: fermenters-batch and fed-batch. Of these, fed-batch fermentation has been the most popular mode historically. Under this production paradigm, batches of high-value products such as APIs are fermented in bioreactors under a controlled environment while being continuously fed nutrients (Banga et al., 1997). Contrasting continuous production, the product is harvested at the end of the fed-batch production (Siganporia et al., 2014)(Siganporia

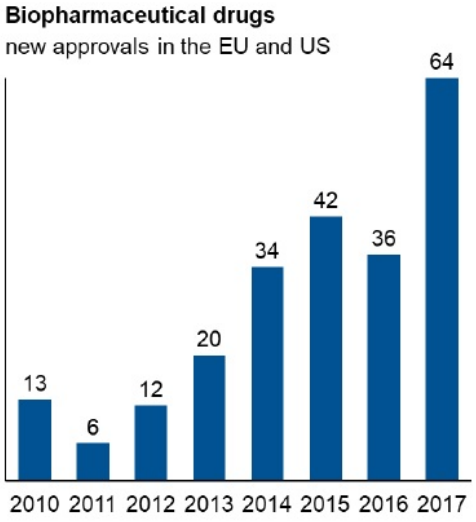

Figure 1: The number of newly approved biopharmaceutical drugs, including those with identical ingredients, in the EU and the US shows a positive, historical growth trend (Walsh, 2014; Walsh, 2018)

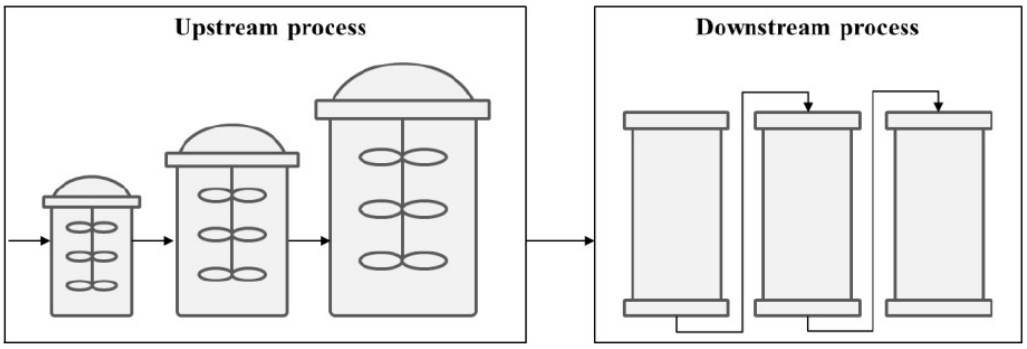

Figure 2: Schematic view of a biopharmaceutical production process with a singular serial seed-train and three sequential chromatography columns

et al., 2014). While continuous fermentation processes are regarded as more economical in some cases, the pharmaceutical, biotechnology, and food industries predominantly use batch processes. Especially for the production of Monoclonal Antibodies (MABs), one of the fastest growing product category in biopharma (Liu et al., 2016), fed-batch has been established as the most prolific production paradigm (Siganporia et al., 2014). Its popularity is mainly due to its lower process complexity, recent increases in production titer sizes, avoidance of over-feeding and high levels of sterility (Freitas et al., 2017; Rani and Rao, 1999; Siganporia et al., 2014).

2.1.2. Methodology of the literature survey

Determining optimal control of fed-batch fermentation has been the subject of academic research for at least twenty years (Lee et al., 1999; Rani and Rao, 1999). For the purpose of this contribution's merit, only literature published in or accepted to be published in peer-reviewed journals and conference proceedings with at least 2nd quartile Scimago or A Jourqual3 rankings were reviewed. Primary sources for collecting relevant literature were references in seminal papers (such as Liu et al. (2014); Martagan et al. (2016); and Schmidt (1996)), Web of Science, EBSCO Host Business Source Complete, and Google Scholar. Databases were searched based on keywords from the most relevant papers. Based on the context described above (Chapter 1), existing literature was categorized along the following five characteristics:

- Planning level: The level of abstraction with which the biopharmaceutical production process is studied.
- Decision space: How the locus of decision-making is defined, i.e., what the set of necessary decisions is.
- Uncertainties: The process-specific uncertainties and their implications which are considered.
- Solution method: The mathematical approach to modeling and, finally, solving the identified problem.

- Financial trade-offs: Whether the financial value of decision making along the production process is considered.

These criteria allow the precise discussion of the existing academic literature and, later on, the proposed approach of this work. Among existing literature, three abstraction levels were identified on which fed-batch biopharmaceutical production is generally studied: (1) biomass kinetics within one bioreactor, (2) scheduling and capacity planning on facility-level, and (3) operating an integrated system of bioreactors and chromatography steps. Following this motivation of the biopharmaceutical production control problem and its study, the next three sections will review the existing literature along the three introduced planning levels: Bioreactor control (Section 2.2), Production and capacity planning (Section 2.3), Production system operations (Section 2.4). Across these levels, problem characteristics will be discussed and, finally, a worthwhile gap in the academic conversation will be concluded.

2.2. Biorreactor control

A large amount of uncertainty during the production of biopharmaceutical products stems from the unpredictability of living organisms' behavior. Cell growth, production rate, and contamination are highly non-linear in nature (Pandian and Noel, 2018). It is therefore paramount to control bioreactors effectively by taking the optimal decision in each situation, i.e., adhering to an optimal policy. In this section, existing literature on fed-batch bioreactor modeling and determination of control policies is surveyed. The non-linear nature of biopharmaceutical production processes rules out conservative linear models (Pandian and Noel, 2018). Control models used for this task are consequently based on Dynamic Programming (DP) (Rocha et al., 2014). DP, as defined by Bellman (1957b), models a dynamic problem as a set of consecutive transitions from state to state. Therefore, bioreactor control researchers have described fed-batch fermentation processes predominantly as control problems, e.g., open loop control (Rani and Rao, 1999). These control problems are often constrained by differential equations modeling the physiological kinetics of the biomass within a reactor (Banga et al., 1997). Models of biomass kinetics can be either structured, i.e., "explicitly describe intracellular processes," or unstructured, i.e., rely on "concentrations of nutrients and metabolites" (Xing et al., 2010, p. 208). Unstructured models often incorporate Monod-type bacterial growth models (Xing et al., 2010). Under the lens of process control, the most common control objective is to find the nutrient feed rate policy which maximizes product concentration at the terminal time or overall reactor productivity. Consequentially, the decision space is often limited to the feed rate of nutrients over the course of fermentation making it a singular control problem. Banga et al. (1997) and Banga et al. (2005), for example, formulate open-loop control problems, modeling the kinetics of one fed-batch fermenter with the goal to maximize the yield of penicillin and ethanol, respectively, by way of controlling the substrate feed-rate. Ponte et al. (2018) describe fed-batch production of the recombinant fungus ROL and obtain optimal feeding trajectories. Skolpap et al. (2004) and Skolpap et al. (2008) model the control of the Monod-type kinetics of α-amylase in a fed-batch fermenter, and include the switching time from batch-mode to fed-batch mode as a decision variable. Peroni et al. (2005) propose an approximate dynamic program to determine the feed-rate profile and the fermentation end time in invertase production of Saccharomyces Cerevisiae.

Due to the computational challenge of determining the control parameters in higher order differential equations, stochastic optimization routines have gained popularity for solving these control problems. While Xing et al. (2010); Skolpap et al. (2004); and Skolpap et al. (2008) use Markov chain Monte Carlo simulation, Banga et al. (2005) implement control vector parameterization, stochastic approaches, such as random search, genetic algorithms, and differential evolution, as well as dynamic hill climbing, to estimate kinetic model parameters. Other stochastic heuristics include model predictive control using evolutionary computation (Ashoori et al., 2009; Freitas et al., 2017), particle swarm optimization (Liu et al., 2013), and, more recently, reinforcement learning (Pandian and Noel, 2018). Peroni et al. (2005) use a neural network for the implementation of an initial approximation and iterative improvements thereof. Earlier, a feedforward neural network was used by Chaudhuri and Modak (1998) to model the same problem.

In general, most surveyed models are deterministic and don't consider process uncertainties such as variable production rates and spontaneous cell death due to contamination. Delvigne et al. (2006), however, propose a Markov chain model of the concentration gradients to complement a deterministic kinetic model. Attempts to include stochastic cell behavior in dynamic control problems have also been made, for example, by Saucedo and Karim (1998); Saucedo and Karim (1997), who propose an MDP for modeling the concentration of ethanol produced by way of an optimal feed-rate policy.

Process engineering works towards the optimal utilization of existing process capabilities (Rocha et al., 2014). Therefore, maximum productivity of the reactor, rather than economically-optimal operation is a common objective (Peroni et al., 2005). However, some scholars even at the bioreactor control level, tie optimization to a financial tradeoff between the value of a maximized product concentration and the cost of nutrients fed into the reactor (Ponte et al., 2018; Saucedo and Karim, 1998).

Process control of batch and fed-batch biopharmaceutical production to achieve maximum productivity has been an area of active academic discussion. However, maximizing production titers during the upstream process is only one way of optimizing production. Biopharmaceutical production, as discussed above, is an interconnected system of upstream fermentation and downstream purification. Given the high costs of all involved process steps, maximizing protein production in the USP may not be an optimal policy for an economically

incentivized decision maker. The next section deals with capacity and production planning literature on the facility-level which allows the consideration of downstream processes in decision making.

2.3. Production and capacity planning

The novelty and complexity of the biopharmaceutical industry, combined with the incumbent focus on spreadsheet-based planning, led to unrealized savings from more sophisticated planning models (Lakhdar and Papageorgiou, 2008). Recently, mathematical optimization of production and capacity planning has been subject to rigorous academic research spanning both up- and downstream processes (Liu et al., 2016). This section reviews past research on this abstraction level that deals with biopharmaceutical production, it is summarized in Table 1.

Due to the high level of abstraction in medium-term planning, earlier work on biopharmaceutical production planning often considered the production process as a black box with decisions focusing on production schedules irrespective of downstream operations. Gatica et al. (2003) and Lakhdar et al. (2006), for example, determined production quantities and campaign durations for a not closer specified biopharma production process. More recently, decision making has expanded to include purification operations as well. Siganporia et al. (2014) modeled capacity decisions in sequential upstream production and downstream purification. Liu et al. (2014), building on Lakhdar et al. (2005), studied production and maintenance planning covering both the USP and DSP, with a special focus on the issue of downstream maintenance work, i.e., when decayed chromatography resins are scheduled to be replaced. Therein, chromatography resins deteriorate after each purified batch, leading to the resin binding less of the available product from the medium in following batches. To restore the performance of the chromatography process, maintenance activities related to the exchange of the used resin are necessary. Liu et al. (2016) expanded their prior work to include sizing and sequencing decisions in both process steps for MAB production. Focusing on a DSP decision space, chromatography capacities and their operations were studied by Allmendinger et al. (2014). Furthermore, Allmendinger et al. (2014); Liu et al. (2014), and Liu et al. (2016) simulated scenarios with parallel production setups being harvested into one or more purification suites.

Due to the inherently uncertain nature of biopharmaceutical production, a stream of literature has specifically focused on dealing with this. Gatica et al. (2003), for example, modeled clinical trial success scenarios in their Mixed Integer Linear Program (MILP) determining capacity plans. Lakhdar et al. (2006) and Lakhdar and Papageorgiou (2008) expanded previous work on multi-period planning and scheduling (Lakhdar et al., 2005) by considering uncertain production titers. Liu et al. (2016), in a modification of their previously deterministic MILP (Liu et al., 2014), considered uncertainties in both the production titer and the resin yield during chromatography purification by means of triangularly distributed stochastic parameters when determining production plans. The impact of uncertain production titer on chromatography decisions was also considered by Allmendinger et al. (2014).

For planning problems under uncertainty, literature has mostly focused on stochastic programming approaches such as 2-stage programming (Lakhdar and Papageorgiou, 2008), Chance-Constrained Programming (CCP) (Lakhdar et al., 2006, Liu et al., 2016), and scenario-based programs (Gatica et al., 2003). Allmendinger et al. (2014) formulated the closed-loop control of a process economic model which was optimized using evolutionary algorithms. Deterministic planning problems are generally modeled as MILPs and solved using standard approaches (Lakhdar et al., 2005; Liu et al., 2014; Siganporia et al., 2014).

As optimal decision making regarding the use of financial and capital assets is one of the principal goals in production and capacity planning, it's no surprise that this is also true in biopharmaceuticals planning. In this stream of planning research, the financial trade-offs consider (1) maximizing sales revenue while incurring minimum capacity investments and operating costs or (2) minimizing costs per sold product. Gatica et al. (2003) and Siganporia et al. (2014), for example, studied a capacity planning problem's trade-off between capacity investments and operating costs and revenues generated from selling the produced amount. Lakhdar et al. (2005); Lakhdar et al. (2006), and Lakhdar and Papageorgiou (2008) studied the classic production planning conundrum of satisfying demand under operating profit maximization. Liu et al. (2014), on the other hand, studied the financial trade-offs of production and maintenance simultaneously. They specifically focus on the trade-off between reduced purification yield from reusing the same resin and incurring costly resin maintenance. Liu et al. (2016) focused on the trade-off between sales revenues and operating costs under consideration of uncertain titers and purification yields, minimizing the total costs of goods. In the latter category of financial trade-offs, Allmendinger et al. (2014) aimed at minimizing the cost of goods per gram of sold product.

In conclusion, biopharmaceutical production planning literature seems to be playing catch-up regarding consideration of operational uncertainty. Due to the high economic impact of resin maintenance, performance decay and the related scheduling of resin exchanges constitute a further need for academic attention. Furthermore, a priori planning and scheduling policies, are not able to guide ad-hoc operations of systems of bioreactors and chromatography columns especially under the influence of biopharmaceutical process uncertainties. The next two sections focus on existing approaches which aim at closing this gap and conclude with the still remaining gap as the focus of this work.

2.4. Production system operations

Recently, academia's widespread focus on concentration maximization within singular bioreactors and its disregard of uncertainties and downstream process implications have

Table 1: Biopharma planning literature

Reference	Decision space	Uncertainties	Model
Lakhdar et al. (2005)	Campaign sequence and duration, production quantities	n/a	MILP
Siganporia et al. (2014)	Capacity plans, out-sourcing decisions	n/a	MILP
Liu et al. (2014)	Maintenance and production plans	n/a	MILP
Gatica et al. (2003)	Product portfolio, capacity planning	Clinical trial out-come	Scenario-based MILP
Lakhdar et al. (2006)	Production plans	Production titer	CCP MILP
Lakhdar and Papageorgiou (2008)	Production plans, sales and backlog profiles	Production titer	2-Stage MILP
Allmendinger et al. (2014)	Capacity plans, chromatography operations	Production titer	Closed-loop control
Liu et al. (2016)	Sequencing and sizing, production plans	Production titer, resin yield	CCP MILP

been questioned by management scholars, such as Martagan et al. (2016). They argue for the need for harvesting policies maximizing discounted financial profit under the consideration of system-level operations. This section surveys existing literature on system-level approaches to the optimization of biopharmaceuticals production.

The decision space in system-level biopharmaceutical production optimization has expanded. Originally, decision making was exclusively concerned with the operations of an upstream seed-train under a known average downstream protein yield (Schmidt, 1996). Schmidt studied the optimal scale-up and harvest policy on a seed-train consisting of sequential, in volume increasing bioreactors. Therefore, the model allowed actions involved in the inoculation of a prepared bioreactor, the feeding of nutrients to grow the culture, the transfer of the medium from a smaller reactor to the next larger one, and the facilitation of protein production and harvesting. 20 years later, Martagan et al. (2016) researched upstream harvesting decisions under explicit purification dynamics. In their model, a decision maker chooses between continuing fermentation and harvesting the protein produced thus far. Later, they studied decisions about chromatography pooling windows under explicit knowledge about upstream product yield (Martagan et al., 2018. Therein, they studied optimal chromatography pooling. Their model aimed at determining which chromatography "lanes", i.e., the amount of protein and impurity flowing through the column per time unit, to "pool", i.e., which lanes to capture from the chromatographic separation, under the consideration of the trade-off between impurity levels and product yield. Most recently, Martagan et al. (Accepted/In press) integrated the simultaneous determination of the upstream production quantity and downstream technology and pooling window. Herein, they modeled a decision maker's interdependent choice about which chromatography technique to use at a given time, which chromatography lanes to pool, and how much protein to ferment. It is interesting to note the level of abstraction to which decisions in the different process steps are studied. While Schmidt (1996) modeled detailed production process operations, excluding downstream decision making, Martagan et al. (2016) only modeled two different actions at each time interval (continue or harvest) and Martagan et al. (Accepted/In press) reduced upstream decisions to the selection of what amount of protein to produce.

Due to the stochastic nature of the dynamic control biopharmaceutical manufacturing, literature at the system abstraction level has exclusively modeled decision making using MDPs. However, the considered uncertainties differ. Schmidt (1996) accounted for the stochasticity of protein production through living organisms. Continuation of culture growth and protein production was assumed not deterministic but subject to random upsets such as contamination. Martagan et al. (2016) considered the uncertainty in the accumulation of MABs and impurities, as well as the arrival of random shocks, such as sudden failure or increased impurity accumulation, during production. Martagan et al. (2018) and Martagan et al. (Accepted/In press) studied the uncertainty of how much protein and impurity remain in the medium at the beginning of each chromatography step. Each piece of literature on system-level decision making in biopharmaceutical production considers the financial trade-off of the decision maker's actions. Schmidt (1996) model, for example, was aimed at aiding process change decisions. Therefore, he modeled the financial impact of harvesting and continuing fermentation. On the one hand, operating revenues were based on an average protein concentration per harvested liter of medium and an average purification yield. On the other hand, each undertaken action incurred a volume-dependent cost for the fed nutrients and a fixed cost. By studying average operating profits, Schmidt (1996) was able to deduct the financial impact of process parameter

changes, such as a reduction of the contamination risk or an increase of the productive phase of the culture, for example, due to investment in process improvements. Similarly, Martagan et al. (2016) considered the financial trade-off between incurring operating costs for continuing fermentation to increase the potential payoff from selling the accumulated MABs and realizing the potential profits by harvesting before batch failure occurs. Martagan et al. (2018) studied protein yield and product purity. Inherent therein is the financial trade-off between the fixed revenue earned from satisfying customer orders and the costs incurred at each chromatography step. In their model, fixed revenue is earned if at least as much product as the customer's yield requirement is delivered, otherwise a yield penalty cost is incurred, given that purity requirements are met. In their most recent work, Martagan et al. (Accepted/In press) integrate both considerations (costs of continued fermentation and profits associated with chosen chromatography operations) into a simultaneously considered financial trade-off.

Research on the system-operations abstraction level and on all other discussed levels is summarized in the following section. It is also there, where the gap is identified which this contribution aims at closing.

2.5. Concluding remarks and identified research gap

After an introduction to the biopharmaceutical production industry in Section 2.1, Section 2.2 reviewed control theory literature on fermentation reactor control constrained by culture kinetics. Existing research on this topic has often focused on the maximization of fermentation productivity with little regard for the financial trade-offs between costly fermentation and associated purification operations. Section 2.3 introduced planning level considerations and the notion of chromatography resin performance decay in the work of Liu et al. (2014). This linear, deterministic planning model, among other decisions, determined maintenance schedules for the replacement of spent chromatography resins. While Liu et al. (2016) model did not consider resin decay, it modeled the uncertainty of the resin's yield. Uncertain, decaying resin yield, however, has not yet been considered.

Section 2.4 reviewed existing literature on the control of fermentation and purification systems. Although research has explored the control problem of interconnected fermentation and purification systems (Martagan et al., Accepted/In press), a paucity of further study of the problem appears to exist. Existing academic work on the topic seems to either overly simplify fermentation operations (Martagan et al., Accepted/In press; Martagan et al., 2016) or exclude control of purification processes (Schmidt, 1996). Furthermore, when to incur chromatography raw material costs, i.e., when to exchange spent resin, has not yet been considered on the system abstraction-level.

Due to the interdependence of up- and downstream decision making in the complex production of highly valuable products, optimal policies for simultaneous operation of protein production and chromatography maintenance under fermentation and resin decay uncertainties constitute a worthwhile endeavor for academia and practice. This contribution, therefore, aims at closing this gap by providing optimal decision policies for the control of an integrated biopharmaceutical production system. The next chapter summarizes the concrete characteristics of this identified problem and introduces the research questions of this contribution.

3. Problem Statement & Research Questions

Following a summary of the setting of the problem studied in this contribution (Section 3.1), this chapter introduces the research questions this contribution aims at answering (Section 3.2).

3.1. Problem context

As per the identified gap in biopharmaceutical production control literature, this contribution aims at the simultaneous optimization of two interconnected sub-problems, i.e., the upstream fermentation policy and the downstream resin exchange schedule. Building on existing literature (see Sections 2.2-2.4), the following problem setting is considered:

Uncertainties: Both, product fermentation and purification are stochastic in nature. These uncertainties are hypothesized to play a critical role in the optimal control of a biopharmaceutical production system. Due to the inherently non-linear nature of the living organisms used in biopharmaceutical production, the physiological states of the observed fermentation culture are highly uncertain. Although fed-batch fermentation is argued to be highly sterile, a risk of batch failure, e.g., due to contamination, persists and is hypothesized to influence optimal decision making (Martagan et al., 2016; Schmidt, 1996). As the performance decay of chromatography resins has thus far only been assumed to be deterministic (Liu et al., 2014) but resin yields constitute a relevant uncertainty (Farid et al., 2005), maintenance decisions under stochastic performance decay are studied.

Planning level: The identified problem is examined on a system-level. The optimal control of the production reactor(s) depends on the state of the fermentation process as observed by a hypothetical production operator. While these states are informed by culture physiologies (Martagan et al., 2016; Schmidt, 1996); explicit culture kinetics are not in-scope for this work. Such approximations are considered "sufficient to describe the evolution of the [fermentation] system through time" (Schmidt, 1996, p. 607). Analogously, the operation of the product purification, while depending on the physiological details of the chromatography technique, are not considered herein.

Decision space: The decision space of interest is the simultaneous control of the highly coupled USP and DSP. The control of the production reactors in a predetermined upstream seed-train configuration to produce a predetermined protein, including decisions regarding the preparation of the reactors, the culture's growth, continuation of production or harvesting the accumulated product are studied in this contribution as these realistically constitute the object of daily

decision making for production operators. This contribution considers only the production reactor of a predetermined USP. For one, it's mainly interested in the production phase and harvesting decisions, and these take place in the production reactor, rather than during the volume-scale up from inoculum flasks to the production volume across the preceding seed-train. As a second argument for this decision, the production reactors of a seed-train can be considered as the bottleneck of that seed-train. For multiple parallel production reactors, the USP decision space consequently spans all bottleneck reactors. The parallel production reactors are considered the bottleneck of parallel production because their operation limits the batch throughput; furthermore, parallel production reactors inoculated by a single seed-train are considered more economical than entire parallel seed-trains (Jagschies et al., 2018, p. 653). The control of the purification process, on the other hand, focuses on the first chromatography step for simplicity, in line with previous work (Liu et al., 2014). Consequently, policies for the first chromatography column with a predetermined resin and capacity must be determined. Decisions include accepting the harvest from the upstream process and exchanging used resin to restore chromatography performance.

Financial trade-offs: The financial trade-off of both sub-problems are integrated and considered simultaneously. During the USP, the financial trade-off between continuing fermentation by incurring operational costs and "locking in" value by harvesting before the whole batch is lost due to sudden failure from contamination has been identified as the most relevant (Martagan et al., 2016; Schmidt, 1996). Downstream decisions about resin maintenance are made in the face of the financial trade-off between scheduling the costly chromatography resin exchange and postponing the exchange, thus accepting a lower financial payoff from the next harvest. On the system-level, a controller, therefore, must integrate the decisions of both sub-problems. This integration considers the financial trade-off of when to "lock in" the revenue from already produced protein before losing the whole batch and the implied cost of reducing the resin performance by purifying a harvested batch. On the one hand, accepting a harvest for purification secures the revenue from the fermented protein but, on the other hand, also accelerates the necessity of costly resin exchange. Meaning, whenever a batch is purified, the need for exchanging the resin is moved closer to the present.

These problem characteristics, compared to existing literature, are summarized in Appendix 3. A formal statement of the problem is provided in Appendix 4.

3.2. Research questions

This contribution poses several questions which existing literature has yet to address but which could contribute to academia's and practice's understanding of this complex environment. This section derives these questions from the identified gaps.

Schmidt (1996) modeled the simultaneous control of sequential upstream bioreactors and Martagan et al. (Accepted/In press) considered both up- and downstream operations but reduced protein production to a single decision. Because of the benefit of simultaneously optimizing up- and downstream operations demonstrated by Martagan et al. (Accepted/In press), there exists paucity to answer the following research question.

> RQ1: What is the optimal, simultaneous control policy for the USP and DSP?

Furthermore, while Liu et al. (2014) determined optimal resin maintenance schedules, only deterministic decay was considered, even though resin yields constitute a source of uncertainty (Allmendinger et al., 2014; Farid et al., 2005; Liu et al., 2016). However, maintenance decisions under uncertain remaining resin capacities, haven't been considered. They can constitute a relevant endeavor for practical decision making because of the high financial impact of chromatography material (Farid, 2007).

> RQ2: When is the chromatography resin exchanged under stochastic decay?

> RQ3: What is the influence of changes to process parameters such as the resin cost and the minimum viable capacity on the optimal policy?

Lastly, because parallel fermentation can increase the batch throughput of a biopharmaceutical production process, seed-trains may be set up to inoculate multiple parallel production reactors to increase utilization of purification equipment (Jagschies et al., 2018, p. 652). While Allmendinger et al. (2014) and Liu et al. (2016) modeled different ratio setups between upstream seed-trains and downstream purification, the operations-level implications of multiple parallel production reactors have yet to be studied, giving the following research question interest and relevance.

> RQ4: How does the consideration of two parallel production reactors change the optimal policy from RQ1?

The related, non-directional hypotheses are summarized in Appendix 5. How this contribution aims at answering these questions is described in the following chapter (Chapter 3). Afterward, the respective analyses are carried out (Chapter 4).

4. Solution Approach

In this thesis, a stochastic dynamic programming approach is developed which allows the study of simultaneous decision making regarding the fed-batch fermentation process, the timing of harvesting, and the maintenance of chromatography resin.

This chapter argues for the use of stochastic dynamic programming over purely stochastic approaches to study the bioreactor system control problem (Section 4.1), reviews the theoretical background of Markov decision processes (Section 4.2), and introduces the numerical case study used to study the identified problem (Section 4.3).

4.1. Dynamic vs. stochastic programming

After reviewing the problem characteristics in the previous chapter (Section 3.1), the curious reader may ponder the question: "under which mathematical framework could one study such a complex environment?" This section argues for MDP as the most fitting modeling technique in this context.

Before discussing possible approaches, the model requirements of the proposed context are reviewed. From the problem context, one can derive the following three requirements of the model approach:

- Regular observation: The model must allow regular observation of a complex system under uncertainties. A hypothetical production system controller must be able to observe the production and purification processes and the cell cultures, periodically.
- Decisions based on state: To study optimal decision making in the control of biopharmaceutical production and purification systems, the model must allow a hypothetical system controller to form a decision based on their observed information. In other words, one observation must provide enough information for the choice of the next decision.
- Required optimality: Due to the significant financial trade-off underlying decisions during biopharmaceutical production and purification, the model should provide an optimal prescription on what actions to take.

Two mathematical modeling techniques may constitute candidates for studying the research questions this contribution poses: stochastic programming and dynamic programming.

Stochastic programming deals with decision problems which can be formulated as follows: "Some decisions must be made today, but important information will not be available until after the decision is made" (King and Wallace, 2012, p. 2). In principle, the problem studied in this thesis can be formulated to fit this description. However, two arguments discourage the use of stochastic programming in this context. Firstly, as Powell (2014) notes, stochastic programs are practically restricted to two decisions instances, so-called two-stage approximations. In two-stage programs, a decision x_t in time t under uncertainty is made, followed by scenario information becoming known and a second decision instance in which decisions for all following epochs are made $x_{t+1} \cdots x_T$ (Powell, 2014). Due to the high dimensionality of scenario trees, which must capture all historical information up to the current decision epoch, stochastic programs are virtually only applicable for problems which can be approximated as two-stage problems. Because biopharmaceutical production systems are monitored and operated in discrete time intervals (Martagan et al., 2016) over sometimes week-long production campaigns (Schmidt, 1996), it seems evident that a two-stage approximation is not feasible. Secondly, as multi-stage stochastic programs (two-stage programs included) are approximations, Powell (2014, p. 111) further postulates that the optimal solution to such an approximation is "(with rare exceptions) not an optimal policy." As this work is interested in finding an optimal policy to study its research questions, stochastic programming can be eliminated as a viable candidate for modeling the identified problem.

Dynamic programs, on the other hand, can be argued to fit the requirements of the presented problem context. Firstly, dynamic programs deal with sequential decision problems and, therefore, allow the study of multi-stage decision making. Due to their dynamic nature, these problems don't approximate multi-stage problems using scenarios but allow modeling them stage-by-stage. Secondly, properly modeled dynamic programs are Markovian (Powell, 2014) and, therefore, aren't history dependent. These first two characteristics allow the modeling of a decision problem, in which an observer sequentially takes in information about the controlled system and can inform a control decision solely based on the last available state description. Thirdly, for subclasses of dynamic programs, it can be shown that an optimal policy exists. For example, using the Banach fixed-point theorem, it can be shown that an optimal policy exists for the total discounted reward problem of an infinite horizon MDPs (Saucedo and Karim, 1997) and relative value iteration converges to the optimal value function of the average cost problem of an infinite horizon MDP (Gupta et al., 2015). Markov decision processes are dynamic programs with stochastic state transitions. Section 4.2 provides more detail on their theoretical background.

This section concludes that dynamic programming meets the requirements of the problem identified in this thesis and, furthermore, MDPs, as a special case of dynamic programs, provide optimal policies for the control of stochastic systems.

4.2. Markov decision processes

This section introduces the notion of Markov decision processes, stochastic dynamic programs which are commonly used to model the control of a complex system which evolves according to a controlled Markov process.

MDPs are tools for analyzing dynamic systems in which state transitions are stochastic and can be influenced by the actions taken by a controller. As such, MDPs combine characteristics of DP and Markov chains (Tijms, 2003, p. 233). The theoretical foundations of MDPs lie in Bellman (1957a) and Howard (1960) for the interested reader.

An MDP is a partly controlled Markov process, i.e., a process following the Markov property (cf. Markov (1954)), in which the transition probability to the next state is solely dependent on the process's current state and the controller's chosen action. Puterman (2014, p. 17–20) mathematically formulates such a process in terms of the 5-tuple ($\mathcal{T}$,$\mathcal{S}$, $\mathcal{A}$, $\mathcal{P}$, $\mathcal{R}$):

- a set of N decision epochs $\mathcal{T}$, over which the system evolves. If N is a finite number, the MDP is referred to as a finite horizon MDP, otherwise as an infinite horizon MDP.

- a set of possible states $\mathscr{S}$, which the system may occupy at any epoch $t \in \mathscr{T}$
- a set of actions $\mathscr{A}$, which the controller of the system can choose. In any given state $s \in \mathscr{S}$, only a subset $\mathscr{A}_s \subseteq \mathscr{A}$ may be permissible. Therefore, $\mathscr{A} = \bigcup_{s \in \mathscr{S}} \mathscr{A}_s$
- a set of transition probabilities P of dimension $|\mathscr{S}| \times |\mathscr{S}| \times |\mathscr{A}|$, which determines the probabilities $p(s, s', a)$ of the system progressing to state s' after choosing action $a \in \mathscr{A}_s$ in state s
- a rewards function R which maps $\mathscr{S} \times \mathscr{A} \rightarrow \mathbb{R}$. In addition to progressing the system into its next state, taking an action $a \in \mathscr{A}_s$ in state s results in an immediate reward $r(s, a)$ for the controller. Depending on the sign of the reward, this may be interpreted as an income (if positive) or a cost (if negative) (Saucedo and Karim, 1998).

Generally, the output of an MDP is a policy π, i.e., the contingency plan determining the action to take in every state of the system in a specified decision epoch. A policy $\pi$is therefore a sequence of decision rules $(d_1, d_2, \cdots, d_{N-1})$ where $d_t \in D$, $\forall t \in (1, 2, \cdots, N-1)$ for $N \leq \infty$. Such a policy is called stationary, if $d_t = d$, $\forall t \in \mathscr{T}$, if all decisions are independent of the current decision epoch. Stationary policies are integral to interpreting infinite-horizon MDPs (Puterman, 2014). An optimal action optimizes a predetermined performance index. Mainly two optimality criteria exist for the solution of MDPs: maximum total expected discounted rewards and average expected rewards. In the case of the total discounted rewards problem of an infinite horizon MDP, the function $v_\lambda^\pi(s)$, which maps a value to any given state s under policy π and the discount factor λ, takes the form of the Bellman's equation

$$v_\lambda^\pi(s) \equiv \sup_{a \in A_s} \left\{ r(s,a) + \sum_{s' \in S} \lambda(s, s', a) v(s') \right\} \tag{1}$$

which recursively determines the discounted, expected reward of choosing the optimal action for any given state s. Therefore, the optimal policy π^* for a given discount rate λ can be determined by solving the following equation:

$$v_\lambda^{\pi^*}(s) \equiv \sup_{\pi \in \Pi} v_\lambda^\pi(s) \tag{2}$$

An optimal policy regarding the total expected reward criterion consequently maximizes the total expected reward, whereas an optimal policy under the average reward criterion maximizes the average reward. Extending Bellman's work in which these MDPs were to be solved backward iteratively by value iteration, Howard (1960) proposed the policy iteration algorithm for solving infinite horizon MDPs. MDPs allow modeling the control of complex systems which evolve over time following a Markov process. As such, Boucherie and van Dijk (2017) provide examples of MDPs' application in highly complex industries, such as healthcare, transportation, and financial modeling.

A prerequisite of the applicability of MDPs to a given problem, however, is the definition of a state space such that all relevant information is captured in each state and the stochastic process over them becomes Markovian. A Markovian process's state in the epoch epoch $t + 1$ solely depends on its state in epoch t, i.e. the process is path-independent or memory-less. As the state of a bioreactor can be described by the biological and thermodynamic characteristics of its content, and these characteristics are path-independent, it has been argued that the process of a bioreactor is Markovian (Schmidt, 1996). A respective assumption is reasonable for the dynamics of chromatography resins performance decay (Liu et al., 2014). Therefore, the requirements of a Markov decision process in this context are met.

4.3. TPA production and purification

To study the research questions introduced in Section 3.2, this contribution analyses the case study of recombinant Tissue Plasminogen Activator (TPA) production and purification. This section argues for this decision and introduces the numerical case study. To draw managerial conclusions on the determined optimal policies, their sensitivity to process changes is analyzed and they are compared to a basic, upstream titer maximizing policy.

TPA is a recombinant protein with medical applications in the dissolution of blood clots. It was considered a "flagship product of the young biotechnology production industry" (Datar et al., 1993, p. 349), but remains an important protein today (Johnston, 2010). TPA can be produced by mammalian, i.e., Chinese hamster ovary, cells or bacterial, i.e., E.coli, cells (Datar et al., 1993).

The case study of TPA was chosen due to two reasons. Firstly, its importance led to wide coverage in academic literature. Because reliability and availability of specific numerical information are scarce in a secretive industry such as biopharmaceuticals (Schmidt, 1996), being able to rely on previously peer-reviewed publications provides a level of data validity. Secondly, TPA is an economically and curatively important product of the biopharma industry to this day. It generates yearly revenues of c.22 m€ in the US and its use adds 0.75 quality-adjusted life years by reducing disability after ischemic strokes (Johnston, 2010).

Schmidt (1996) summarizes the production of TPA as follows: mammalian cells are attached to inexpensive microcarriers (0 $/g) in a medium which includes the costly nutrient Fetal Calf Serum (FCS), because TPA exhibits anchorage-dependent growth. Furthermore, TPA production only starts after the culture has finished growing, because TPA exhibits non-growth dependent production. Therefore, a protein producing phase is induced following a growth phase. The medium required for the continuation of the growth phase (12.8 $/L) is more expensive than that which is used during protein production (2 $/L) because of the higher FCS concentration of the former. The medium is exchanged every 12 hours and a growing culture is switched into its TPA producing state by exchanging growth for production medium. Schmidt (1996) assumed that exchanging the medium is

equivalent to harvesting all TPA from the medium. By doing this, however, his model is constrained to considering an average TPA concentration of 15 mg/L instead of the time-dependent accumulation of up to 33.5 mg/L of TPA reviewed by Datar et al. (1993).

This thesis models fed-batch dynamics in which the medium is only harvested at the end of the batch. After each cycle, the spent medium in the bioreactor is exchanged with new medium without reducing the TPA concentration, i.e., new medium is added to the reactor while the TPA concentration is growing continuously. In practice, this can be accomplished, for example, feeding new FCS into the spent medium or by separating the spent medium from the TPA without harvesting. The production reactor of a three-reactor seed-train can have a volume of c.160 L and the value of TPA is c.24,000 \$/g (Schmidt (1996)). As the weekly contamination risk in TPA production is between 5 and 10 % and a week has 14 12-hour production cycles, Schmidt (1996) assumes the probability of a successful state transition of a bioreactor to be 99.3 % for all non-zero and non-one probabilities.

After reaching an economical concentration of TPA in the production medium, the protein is separated from the medium without intermediate storage. Lin et al (1993), for example, described a process of affinity purification during which an anti-TPA MAB is used to ease the binding of TPA during chromatography. In Liu et al. (2014) numerical case study, the cost of chromatography maintenance was assumed to be equivalent to three-quarters of the revenue from one sold production batch. Because the value of a full batch of TPA, i.e., 5.36 g of pure TPA, is c.128,640 \$, and resin material is the main driver of these maintenance costs, chromatography resin is assumed to cost c.96,480 \$. This aligns with the magnitude of reported per resin L-prices and necessary resin volumes for chromatography resins by Allmendinger et al. (2014) and Liu et al. (2016). Every action is assumed to incur 100 \$ of fixed costs, mainly due to labor costs.

5. Mathematical Model Formulation

This chapter presents a stochastic dynamic program to study the system-level control of biopharmaceutical production and purification under fermentation and chromatography performance decay uncertainty. The presented model aims at finding simultaneous optimal policies for fermentation control and purification to maximize average expected operating profit. Section 5.1 provides an overview of the used nomenclature. Section 5.2 introduces the proposed discrete state space, infinite time horizon MDP. Section 5.3 and 5.4 consider two peculiarities of the proposed model: the accumulation of protein in the medium and the decay of the chromatography resin. Section 5.5 introduces the objective function of the model.

5.1. Nomenclature

5.2. MDP formulation

The model proposed in this contribution extends Schmidt (1996) upstream model by a downstream component. Furthermore, it allows for multiple parallel production reactors being harvested into a single downstream process to test the presented hypotheses about the implications of the purification process on optimal policies in the USP seed-train and vice versa (see Section 3.2).

The presented framework consists of the two sub-problems described above and is mathematically formulated in terms of the 5-tuple ($\mathcal{T}$,$\mathcal{S}$, $\mathcal{A}$, $\mathcal{P}$, $\mathcal{R}$) using the notation introduced in Section 4.2:

Decision epochs: In accordance with Schmidt (1996), this contribution proposes a discrete-time, infinite time horizon Markov decision process. Let $\mathcal{T} = \{0, 1, 2 \cdots, N\}$ be the set of discrete, equidistant decision epochs with an infinite N. Each decision epoch $t \in \mathcal{T}$ represents represents an observation of the system. A discrete time representation was chosen as it best represents industry practice. Industry practice for biopharmaceutical production is seldom real time monitoring. While real-time, on-line, monitoring of culture physiologies is the topic of ongoing academic discussion (e.g., Abu-Absi et al. (2011), Wechselberger et al. (2013)), on-line monitoring is often high-priced or limited in practicality and functionality (Lourenço et al., 2012). In practice, therefore, monitoring of biomass physiologies is often carried out off-line (Lourenço et al., 2012). Thereby, samples are either taken automatically or manually from the bioreactor and transferred to a laboratory for analysis. Because this is a time-intensive process, a discrete-time representation of the modeled process makes sense. Martagan et al. (2016), for example, observe measurement intervals of two to three days in practice.

An infinite horizon is modeled to allow for the study of stationary policies.

State space: During upstream production, multiple sequential bioreactors may be used to scale-up a cell culture to production volume. As described above, this work focuses only on the production reactors as the bottleneck of the process. The state space of the upstream fermentation policy problem for one production reactor consists of approximations of the culture's thermodynamic properties and was defined by Schmidt (1996). Let

$$\mathcal{S}^U = \{\text{empty}, \text{ready}, \text{growth}_i, \text{production}_j, \text{upset} : \forall i, j \in \mathbb{N}, \forall i, j \leq n_g, n_p\} \quad (3)$$

define the set of feasible states of a production reactor. Let n_g, n_p be the predetermined, finite number of periods in which the culture exhibits cell growth and protein production, respectively. The indices i and j, therefore, represent the number of cycles since the start of the respective phase. For example, production_{30} is the beginning of the 30^{th} production cycle, i.e., the culture has traversed 29 cycles since it entered the protein production phase. The state $s^U \in \mathcal{S}^U$,

Table 2: Nomenclature used in the presented Markov decision process model

Superscripts	
k	Process component, $k \in \mathscr{K}$
Indices	
i	Growth cycle, $1 \leq i \leq n_g$
j	Production cycle, $1 \leq j \leq n_p$
l	Parallel upstream production reactor, $1 \leq l \leq n_U$
m	Number of the next purified batch, $1 \leq m \leq$ MaxBatch
Sets	
$\mathscr{T}$	Decision epochs
$\mathscr{S}$	State space of the MDP
$\mathscr{S}^k$	State space of process component k
$\mathscr{A}$	Action space of the MDP
$\mathscr{A}^k$	Action space of process component k
$\mathscr{A}_s$	Feasible actions in state s
$\mathscr{A}^k_{s^k}$	Feasible actions in the k-th process component in state s^k
$\mathscr{P}$	Transition probabilities of the MDP
$\mathscr{P}^k$	Transition probabilities of process component k
$\mathscr{K}$	Process components $\mathscr{K} = \mathscr{U} \cup D$, where $\mathscr{U}$ are the parallel upstream reactors and D is the first downstream chromatography step
$\mathscr{U}$	Set of n_U parallel upstream production reactors, $\mathscr{U} = \{U_1, \ldots, U_l, \ldots, U_{n_U}\}$
$\mathscr{G}_p$	Production-competent growth states which can be converted into production states within one decision epoch
Functions	
$\mathscr{R}$	Rewards function, $\mathscr{S} \times \mathscr{A} \rightarrow \mathbb{R}$
π	Policy function, $\mathscr{S} \rightarrow \mathscr{A}$
π^*	Optimal policy function, which maximizes the objective criterion
$\mathscr{V}^\pi(s)$	Value function for policy π, $\mathscr{S} \rightarrow \mathbb{R}$
Parameters	
N	Number of decision epochs
t	Decision epoch, element of T
n_U	Number of parallel upstream production reactors, $n_U = \lvert U \rvert$
$n_g,\ n_p$	Number of growth and production cycles, respectively
Δ	Resin state, in which no further harvests can take place and the resin must be exchanged
capacity_m	Remaining capacity of the resin to bind protein in the m-th harvested batch,%
MaxBatch	Maximum number of batches purifiable using the same chromatography resin, before it must be exchanged
$p(s^k, s'^k, a^k)$	Probability of the k-th process component transitioning from its state s^k in a decision epoch to s'^k in the following decision epoch, if action a^k is taken in it
$p(s, s', a)$	Probability of the system transitioning from state s in a decision epoch to s' in the following decision epoch, if action a is taken
$r(s^k, a^k)$	Reward obtained from choosing action a^k in state s^k in the k-th process component
$r(s, a)$	Reward obtained from choosing action a in state s

(Continued)

Table 2—continued

Parameters	
V	Volume of a production reactor, L
$c_g,\ c_p$	Cost of a liter of growth and production medium, respectively, \$/L
c_b	Cost of a gram of microcarriers, \$/g
fixed(a^k)	Fixed cost of carrying out action a^k in the k-th process component, \$
v_g	Value of a liter of growth medium, \$/L
v_p	Value of a gram of protein, \$/g
x_j	Concentration of protein in the production medium in the j-th production cycle, g/L
x_{max}	Maximum possible concentration of protein in the production medium by the end of the production phase, g/L
c_{resin}	Material cost of exchanging spent resin, \$

therefore, describes a bioreactor in terms of its readiness to be used and the physiology of the contained culture. Accordingly, a reactor may be empty, sanitized and ready for inoculation, its culture may traverse states of growth and states of protein production. Out-of-the-ordinary states such as contamination are summarized as upset. In growth-independent production cultures, an ordered set of production-competent growth states $\mathscr{G}_p$ exists. In these states, the culture exhibits kinetics which allow its conversion into a protein producing state within one cycle. For n_U parallel seed-trains, the state of the upstream sub-problem is represented by the vector $s^U = (s^{U_1}, \cdots, s^{U_l}, \cdots, s^{U_{n_U}}) \in \mathscr{S}^{U_1} \times \cdots \times \mathscr{S}^{U_l} \times \cdots \times \mathscr{S}^{U_{n_U}}$. For simplicity, the assertion

$$\mathscr{S}^{\mathscr{U}} = \mathscr{S}^{\mathscr{U}_l}, \forall\, \mathscr{U}_l \in \mathscr{U} \tag{4}$$

is made, where $\mathscr{S}^{\mathscr{U}_l}$ is the state space of the l-th in the set $\mathscr{U}$ of all $n_{\mathscr{U}}$ parallel production reactors.

The state space of the downstream resin exchange schedule problem consists of the remaining performance of the resin in the chromatography column, i.e., what percentage of the protein in the medium can successfully be bound by the resin during purification. The performance of the resin used during chromatography decays after each purified batch. Based on research on resin performance for the industry-standard Protein A chromatography by Jiang et al. (2009), research by Liu et al. (2014) assumed a small number of batches purified at full performance, followed by a linear performance decay to a minimum viable resin performance. This work formulates Liu et al. (2014) resin decay pattern as the state space of the downstream sub-problem. Let

$$\mathscr{S}^D = \{\text{capacity}_m : m \in \mathbb{N},\ 1 \leq m \leq \text{MaxBatch}\} \cup \Delta \tag{5}$$

define the set of feasible states of a downstream chromatography column. Therein, MaxBatch is the finite, maximum number of batches which can be purified using the same resin before it must be exchanged. For MaxBatch = 10, the capacity of the resin may start at full capacity before the first batch is purified, i.e., capacity_1=100%, and deteriorate over predefined, intermediate steps to its minimum viable capacity, i.e., capacity_{10}, before purification of the tenth batch. If the resin is used at its minimum capacity, $\text{capacity}_{\text{MaxBatch}}$, it deteriorates to a state in which it must be replaced, Δ. The state $s^D \in \mathscr{S}^D$, therefore, is a rational, non-negative number in $[100\%, \text{capacity}_{\text{MaxBatch}}]$, describing the remaining protein binding capacity of the resin at the start of each decision epoch, or Δ.

The state space of a production system with n_U parallel bioreactors and one chromatography column is, therefore, defined as $\mathscr{S} = \mathscr{S}^{U_1} \times \cdots \times \mathscr{S}^{U_{n_U}} \times S^D$, with the vector $s(t) = (s^{U_1}, \cdots, s^{U_{n_U}}, s^D) \in \mathscr{S}$ identifying the state of the system at the beginning of epoch t. The vector (empty,100%), for example, describes a system with a single production reactor in a state, wherein the reactor is empty, and the chromatography resin is at its full performance.

> Assumption 1: The state space is finite. Therefore, $|\mathscr{S}| < \infty$.

Action space: Let $\mathscr{A}^U$ be the set of feasible actions in a production reactor as part of the upstream fermentation sub-problem. The system operator choses an action $a^{U_l}(s^{U_l})$ from the set $\mathscr{A}^{U_l}_{s^{U_l}} \subseteq \mathscr{A}^{U_l}$ of feasible actions to take for the l-th parallel upstream reactor in state s^{U_l}. The upstream action space, $\mathscr{A}^U$, is then the union $\cup_{s,l} \mathscr{A}^{U_l}_{s^{U_l}}$ of all feasible actions in all possible upstream reactor states. Schmidt (1996) action space is adapted to fit the proposed model. As this contribution only considers the production reactors and, therefore, abstracts from the scaling-up of the culture medium from small-scale reactors to the production reactor, actions related to these activities are omitted. The goal of the upstream fermentation control is the production of the optimal amount of protein by way of scheduling the culture growth and production phases within a production reactor. This leads to an action set available for the control of the production reactor which modifies the model of the growth-independent production of TPA by Schmidt (1996). The operator controls a reactor to grow a cell culture and facilitate protein production. To this end, they sanitize an empty reactor for inocula-

tion (prep), feed nutrients into the reactor to grow the culture (addgm), feed a different mix of nutrients into the reactor to convert the growing culture into producing culture and continue this production (addpm), or extract the medium from the bioreactor for harvesting (harvest). In a state of non-production, e.g., upset or growth, harvest is equivalent to dumping the contents of the reactor out. In line with the flexibility allowed by Schmidt (1996), the presented model also includes the combination of harvesting/ dumping and preparation as one action (hprep). The action set is summarized in Table 3. However, not all actions are feasible in every state of a production reactor. Adding production medium, for example, is only feasible for production-competent growth states (thus, inducing production) and all but the final production state (thus, continuing production). Adding growth medium is only feasible in a ready state (thus, inducing growth) and during the growth phase of the culture except for the final growth state.

Let $a^U = (a^{U_1}, \cdots, a^{U_l}, \cdots, a^{U_{n_U}}) \in \mathscr{A}^{U_1} \times \cdots \times \mathscr{A}^{U_l} \times \cdots \times \mathscr{A}^{U_{n_U}}$ be the vector of actions taken in an upstream process of n_U parallel production reactors. Again, for simplicity, the following assertion is made for the action spaces of each production reactor:

$$\mathscr{A}^U = \mathscr{A}^{U_l}, \forall U_l \in U \tag{6}$$

Consequently, let A^D be the set of feasible actions in the downstream resin exchange schedule sub-problem. This contribution incorporates Liu et al. (2014) notion of the maintenance activity required to replace deteriorated resin into the action space of a biopharmaceutical production system operator. At any given state of the chromatography resin, the actions $\mathscr{A}^D_{s^D} \subseteq \mathscr{A}^D$ are feasible and $a^D(s^D) \in \mathscr{A}^D_{s^D}$ denotes the action chosen in the downstream sub-problem. Equivalently to the upstream sub-problem, the action space of the downstream sub-problem, $\mathscr{A}^D$, is the union $\cup^D_s) \mathscr{A}^D_{s^D}$ of the feasible actions in all states of the chromatography resin. An operator must choose between doing nothing (none), accepting the harvest from the upstream process (accept) or exchanging the resin to restore full chromatography performance (exresin). The action space of this sub-problem is summarized in Table 4. Whenever the resin capacity deteriorates past the predetermined minimum viable threshold, $\text{capacity}_{\text{MaxBatch}}$, it must be exchanged, therefore, only action 3 is feasible in the state $s^D = \Delta$.

The action space of the simultaneous USP and DSP control problem presented herein is defined as the set of actions $A = A^{U_1} \times \cdots \times A^{U_{n_U}} \times A^D$ from which the controller of the production system can choose at the beginning of each decision epoch. Let $a(s) \in A_s$ denote the vector of actions taken in system state s representing the joint decision made by the operator. Therein, A_s is the set of feasible actions in states s. The action vector (addpm,exresin) or $(3, 3)$, for example, indicates that, in a system of a single production reactor and one DSP, production medium is added to the production reactor and the resin of the chromatography column is exchanged in the same decision epoch.

As introduced earlier, the simultaneous control of the entire production system is complicated by the inter-dependencies between the up- and the downstream part of the process. Because both processes are assumed to be coupled, i.e., there is no intermediate storage of harvested medium, harvesting a production reactor (harvest or hprep) is only feasible if the chromatography step chooses to accept the medium for purification (accept) in the same decision epoch. Furthermore, in a system of multiple parallel reactors, only one can be harvested at a time.

Transitions: The observed production and purification system evolves over time from state to state, i.e., along the path $\{s(t_0), s(t_1), \cdots, s(t_N)\}$, with $s(t) \in \mathscr{S}, \forall t \in \mathscr{T}$. As described above, the states which a reactor in the upstream fermentation sub-problem occupies over time are defined by approximations of the culture's cell concentrations, thermodynamic quantities and time passed since the start of either growth or production (Schmidt, 1996). A production reactor, therefore, evolves over time according to the physiological rules which govern cell cultures (as captured by cell-level models in Section 2.2) and the actions taken by the controller. In the highly uncertain environment of living cell cultures, however, this evolution is not deterministic.

Let

$$\begin{gathered} p(s^{U_l}, s'^{U_l}, a^{U_l}) = \\ \mathbb{P}\left[s^{U_l}(t+1) = s'^{U_l} | s^{U_l}(t) = s^{U_l}, a^{U_l}(t) = a^{U_l}\right], \\ \forall s^{U_l}, s'^{U_l} \in \mathscr{S}^U, \forall a^{U_l} \in \mathscr{A}^{U_l}_{s^{U_l}}, \\ \forall U_l \in U, \forall t \in T \end{gathered} \tag{7}$$

determine the probability of the l-th production reactor to evolve from state s^{U_l} in one decision epoch t to state s'^{U_l} in the next, $t+1$, if action a^{U_l} is chosen. While this model allows for transition probabilities in the bioreactor to be set one for one, a simplification is made to reduce the number of necessary process parameters: every non-zero, non-one probability is identical for every production reactor U_l. However, in line with Schmidt (1996), the cell culture decline towards the end of its lifecycle manifests by transition probabilities decreasing over the last sixth of the production phase. Using the state and action schemas defined above, Table 5 summarizes the transition probabilities of the upstream fermentation sub-problem.

The transitions of the chromatography step introduce uncertainties to the downstream resin exchange schedule sub-problem. The state of the chromatography step evolves over the predefined set of remaining capacities the resin can have. As studied by Jiang et al. (2009), chromatography resins decay with increased usage and their capacity to bind proteins of interest deteriorates. Accepting a harvest from the upstream part has two consequences: Firstly, an amount of protein is yielded depending on the remaining capacity of the resin and, secondly, the resin's capacity deteriorates to the next capacity step in the following epoch. As the purification of a batch in the first chromatography step is assumed to take five cycles (Liu et al., 2014) and within six to

Table 3: Action space of the control of a production reactor

Action	Abbreviation	code
Do nothing	none	1
Add growth medium	addgm	2
Add production medium	addpm	3
Harvest or dump	harvest	4
Prepare	prep	5
Harvest or dump and prepare	hprep	6

Table 4: Action space for the control of the first chromatography step

Action	Abbreviation	code
Do nothing none	None	1
Accept a harvest for purification	accept	2
Exchange resin	exresin	3

Table 5: Non-zero transition probabilities of an upstream sub-problem, adapted from Schmidt (1996)

s^{U_l}[1]	index[2](s^{U_l}, s'^{U_l},a^{U_l})		
e		none	$p(e,e,1)$
		prep	$p(e,r,5), p(e,u,5) = 1-p(e,r,1)$
r		none	$p(r,u,5) = 1$
		harvest	$p(r,e,4) = 1$
		hprep	$p(r,r,6) = p(e,r,5), p(r,u,6) = 1-p(e,r,5)$
g_i	$1..n_g$	none	$p(g_i,u,1) = 1$
	$1..n_g-1$	addgm	$p(g_i,g_(i+1),2), p(g_i,u,2) = 1-p(g_i,g_(i+1),2)$
	$i \in G_p$	addpm	$p(g_i,p_1,3), p(g_i,u,3) = 1-p(g_i,p_1,3)$
	$1..n_g$	harvest	$p(g_i,e,4) = 1$
	$1..n_g$	hprep	$p(g_i,r,6) = p(e,r,5), p(g_i,u,6) = 1-p(e,r,5)$
p_j	$1..n_p$	none	$p(p_j,u,1) = 1$
	$1..n_p-1$	addpm	$p(p_j,p_(j+1),3), p(p_j,u,3) = 1-p(p_j,p_(j+1),3)$
	$1..n_p$	harvest	$p(p_j,e,4), p(p_j,u,4) = 1-p(p_j,e,4)$
	$1..n_p$	hprep	$p(p_j,r,6) = p(e,r,5), p(p_j,u,6) = p(e,u,5)$
u		none	$p(u,u,1) = 1$
		harvest	$p(u,e,4) = 1$
		hprep	$p(u,r,6) = p(e,r,5), p(u,u,6) = p(e,u,5)$

eight hours (Martagan et al., 2018), harvests aren't blocked by ongoing purification. Yields from chromatography, however, have been considered a source of uncertainty, e.g., by Liu et al. (2016) Therefore, the model proposed in this contribution allows for stochasticity in the performance decay of the chromatography step. Let $p(s^D, s'^D, a^D)$ be the probability of the resin in state s^D decaying to its new state s'^D, if action a^D is taken. This allows modeling of uncertain deterioration: after a purification cycle, resin performance may, for example, reduce to the next lower step, remain at the same performance level or deteriorate by more than one step. Whenever the resin is exchanged, the performance is deterministically restored to the resin's original value, i.e., $p(s^D, \text{cap}_1, 3) = 1$, $\forall s^D \in S^D$. Maintenance activities are not time-intensive (Liu et al., 2014) and are, therefore, assumed to be feasible in one decision epoch. Furthermore, the resin must be exchanged when it deteriorates past its minimum allowed capacity. These dynamics are detailed in Section 5.4. Table 6 details the non-zero transition probabilities.

Transition probabilities within bioreactors (Schmidt, 1996) and within chromatography resins are assumed to be independent. Therefore, let

$$p(s,s',a) = \prod_{k \in \mathscr{K}} p(s^k, s'^k, a^k) \tag{8}$$

define the probability of the system evolving from state s in decision epoch t to state s' in $t+1$, where $\mathscr{K}$ is the set of the

[1] State descriptions have been shortened for readability, i.e., empty to e, ready to r, growth to g, production to p, and upset to u

[2] Index range, in which action a^{U_l} is feasible

biopharmaceutical process "components", i.e., all production reactors and the chromatography column.

> Assumption 2: Transition probabilities are stationary, i.e., do not vary with time. Therefore, $p(s,s',a) = p(s(t),s(t+1),a(t)),\ \forall t \in \mathscr{T}$.

Let $\mathscr{P}$ constitute the set of all transition probabilities $p(s,s',a)$.

Rewards: Production and purification system operations deal with the financial trade-off between operating costs, caused by continued fermentation and chromatography resin exchange, and revenue from purified protein. The rewards obtained from a USP reactor in the state s^{U_l} by choosing the action $a^(U_l)$ are denoted $r(s^{U_l}, a^{U_l)}$. In line with Schmidt (1996), actions incur small fixed costs and volume dependent variable costs or revenues. The cost of adding new growth medium (addgm) is, therefore, the sum of the cost of the growth medium and a fixed cost for carrying out the action. The former is dependent on the two parameters: the volume of the reactor (V in L) and the cost of the growth medium per volume (c_g in \$/L): $r(s^{U_l}$,addgm)$= -c_g\ V$-fixed(addgm). Therein, c_g depends on the concentration of FCS in the growth medium and its cost. Consequently, $r(s^{U_l}$,addpm)$= -c_p\ V$-fixed(addpm) is the cost of inducing or continuing the production by adding nutrients into the spent medium up to a production medium concentration, where c_p is the per-liter-cost of the production medium, depending on its FCS concentration. Preparing an empty reactor (prep) with a growth medium incurs costs for the used growth medium and microcarriers allowing for anchorage-dependent culture growth. Because both costs are volume-dependent, $r(s^{U_l}$,prep$) = -(c_g + c_b)\ V$-fixed(prep) defines the cost of this action, where c_b is the cost of the microcarrier (in \$/L). If the medium is dumped during the growth phase (harvest, hprep), salvage revenue from the spent growth medium is earned. Let, therefore,

$$r(s^{U_l}, a^{U_l},) = \begin{cases} v_g V - \text{fixed(harvest)} & \text{if } a^{U_l} = \text{harvest} \\ (v_g - (c_g + c_b))V \text{fixed(hprep)} & \text{if } a^{U_l} = \text{hprep} \end{cases} \tag{9}$$

where $s^{U_l} \in \{\text{growth}_i :\ 1 \le i \le n_g\}$. If a batch is lost due to doing nothing, negative reward in the amount of the opportunity costs are earned. Let

$$r(s^{U_l}, \text{none}) = \begin{cases} -v_g V & \text{if } s^U \in \{\text{growth}_i :\ 1 \le i \le n_g\} \\ -v_p x_j V & \text{if } s^U \in \{\text{production}_i :\ 1 \le i \le n_p\} \end{cases} \tag{10}$$

determine the opportunity costs of losing a batch due to doing nothing, where v_g is the value of the growth medium in \$/L, v_p is the value of TPA in \$/g, x_j the concentration of TPA in the production medium at the time of losing the batch in g/L, and V the volume of the reactor in L. Rewards from harvesting production medium (harvest, hprep) are only obtained after purification, therefore, these are included in the rewards function of the DSP sub-problem.

Equivalently, the rewards obtained from operating the chromatography step in state s^D by choosing action a^D are denoted $r(s^D, a^D)$, which is defined for all $a^D \in \mathscr{S}^D$ and all $s^D \in \mathscr{S}^D$. The material cost of chromatography resin is significant (Allmendinger et al., 2014; Farid, 2007), therefore the cost of maintenance activities mainly depends on the material costs. Let $r(s^D, \text{exresin}) = -c_{\text{resin}} - fixed(\text{exresin})$ describe the cost of exchanging the chromatography resin. The revenue obtained from purified product depends on the volume of TPA in the medium at the point of harvesting and the remaining binding capacity of the resin. Let

$$r(s^{U_l}, \text{accept}) = \begin{cases} -v_p x_j V \text{capacity}_m - \text{fixed(harvest)} \\ \quad \text{if } \exists! U_l \in \mathscr{U} :\ a^{U_l} = \text{harvest} \\ -v_p x_j V \text{capacity}_m - \text{fixed(hprep)} \\ \quad \text{if } \exists! U_l \in \mathscr{U} :\ a^{U_l} = \text{hprep} \end{cases} \tag{11}$$

define the operating profit obtained from accepting a harvest from exactly one production reactor which depends on the value of the production medium at the time of harvesting and what percentage of that value is captured during purification (capacity$_m$ in %). While Schmidt (1996) implies harvesting every time spent production medium is exchanged for new medium in the reactor (addpm), the presented model only allows harvesting as an action distinct from continuing fermentation which returns the bioreactor to an empty state, to better model the practical realities of fed-batch fermentation as presented by Martagan et al. (2016).

Let the aggregate rewards function

$$\mathscr{R}(s,a) = \sum_{k \in \mathscr{K}} r(s^k, a^k) \tag{12}$$

denote the sum of costs/ revenues of carrying out each action in the action vector a in system state s on their respective reactor/ chromatography column.

> Assumption 3: Like transition probabilities, rewards are stationary and bounded. Therefore, $r(s,a) = r(s(t),a(t)),\ \forall t \in \mathscr{T} and |r(s,a)| \le M < \infty,\ \forall s \in \mathscr{S}, \forall a \in \mathscr{A}(s)$.

By assumption, the reward obtained from harvesting the medium depends strongly on the amount of protein in the medium and the yield performance of the chromatography resin. As these variables constitute important aspects of the presented model, special consideration should be given to their modeling. The following two sections are aimed at this.

5.3. TPA accumulation during production

Protein production in living cell organisms most prominently follows one of two patterns: growth related or non-growth related (Schmidt, 1996). In growth related protein

Table 6: Non-zero transition probabilities of the downstream sub-problem

s^D [3]	$index$ [4]	a^D	$p(s^D, s'^D, a^D)$
cap_m	1..MaxBatch -1	none	$p(\text{cap}_m, \text{cap}_m, 1) = 1$
	1..MaxBatch -1	accept	$\sum_{0\leq s\leq 1} p(\text{cap}_m, \text{cap}_{m+s}, 2)+$ $p(\text{cap}_m, \Delta, 2) = 1$ $p(\text{cap}_m, \text{cap}_m, 2)$
	MaxBatch	accept	$p(\text{cap}_m, \text{cap}_m, 2), p(\text{cap}_m, \Delta, 2) =$ $1 - p(\text{cap}_m, \text{cap}_m, 2)$
	1..MaxBatch	exresin	$p(\text{cap}_m, \text{cap}_1, 3) = 1$

formation, the product is produced as a by-product of culture growth, i.e., during growth states. In non-growth-related protein accumulation, however, proteins are produced only after the cell culture has finished growing and started producing, i.e., during production states. Figure 3 illustrates this difference. Towards the end of their life, cell cultures exhibit decline (Schmidt, 1996)(Schmidt, 1996). This is modelled by the probabilities of a successful transition decreasing over the last sixth of the production phase: $p(p_{30}, p_{31}, 3)$, $\cdots$, $p(p_{35}, p_{36}, 3)$ have the values 0.84, 0.67, 0.5, 0.34, 0.17, 0.05, which were adapted from Schmidt (1996). TPA, the subject of this work's numerical study, is produced in a non-growth-related pattern.

> Assumption 4: The protein concentration x_j monotonically increases in j. Therefore, $x_j \leq x_(j+1)$, $\forall j : 1 \leq j \leq n_p$.

Although stochastic formulations of protein accumulation exist (Martagan et al. (2016)), one may sensibly approximate protein accumulation as a linear relationship as illustrated in Figure 3. Therefore, the proposed model describes a linear growth of TPA in the production medium starting in the first production state and ending with a maximum concentration $x_{\max}$ in the final production state. Let the concentration at the beginning the production cycle j be defined by the equation

$$x_j = \frac{x_{max}}{n_p - 1} \times (j-1), \ \forall j : 1 \leq j \leq n_p, \ \forall n_p > 1 \quad (13)$$

The product of concentration (in g/L) and bioreactor volume (in L), therefore, describes the amount of protein in the reactor (in g).

5.4. Resin performance decay

Chromatography resins decay due to being repeatedly exposed to process conditions; their decay is most commonly monitored based on product yield (Nweke et al., 2018). The implied financial trade-off between reduced revenues from harvesting fermented product and the scheduling of costly resin exchange activities is studied in this contribution. Therefore, modeling resin decay is given special consideration in this section.

The notion of maintenance actions related to resin performance decay is based on Liu et al. (2014), who expanded on prior work with a focus on downstream planning (Allmendinger et al., 2014). Liu et al. (2014) further built on Jiang et al. (2009) study of resin performance decay. They assume that each chromatography step, i.e., the purification of a batch in one chromatography column, takes five cycles. Therefore, they aggregate Jiang et al. (2009) experimental results of resin decay over multiple chromatography cycles to a per-batch decay pattern in their first numerical example. Similar to Liu et al. (2014), this work assumes a resin deterioration over 11 batches, with each batch taking five chromatography cycles. After the eleventh purified batch, the resin must be exchanged. Therefore, after four batches purified at full performance, the resin starts to deteriorate until its minimum capacity is reached before batch 11 (see Figure 4). This decay, however, was previously only deterministic in nature. In practice, this may not be the case and constitute an additional source of uncertainty and operational complexity. Because resin yield is one of the uncertainties which have an impact on production cost and delivery (Farid et al., 2005), Liu et al. (2016) considered the notion of uncertain resin in capacity planning. They modeled resin yield deviations following a triangular distribution. In their numerical case, they assume that resin yields deviate at most 5 % from their standard value depending on the type of resin.

However, because the model in this contribution has a finite, discrete state space, assumptions about the evolution of the resin capacity must be made. The performance decay in this model follows a similar pattern as in the Liu et al. (2014) model but introduces an element of uncertainty by assuming the decay steps to be stochastic rather than deterministic.

> Assumption 5: After a batch is purified, the resin either doesn't decay, decays as expected after one purification, or decays as if two batches were purified. Therefore, $\sum_{0\leq s\leq 2} p(\text{cap}_m, \text{cap}_{m+s}, 2) = 1$, $\forall m <$ MaxBatch and $p(\text{cap}_m, \text{cap}_{m+1}, 2) + p(cap_m, \Delta, 2) = 1$, if $m =$ MaxBatch

The fifth assumption introduces stochasticity in the resin yields. As indicated in Table 6, after the purification of one batch, the yield capacity of the resin may stay the same, dete-

[3] $capacity_m$ has been shortened to cap_m for readability
[4] Index range, in which action a_t^D is feasible

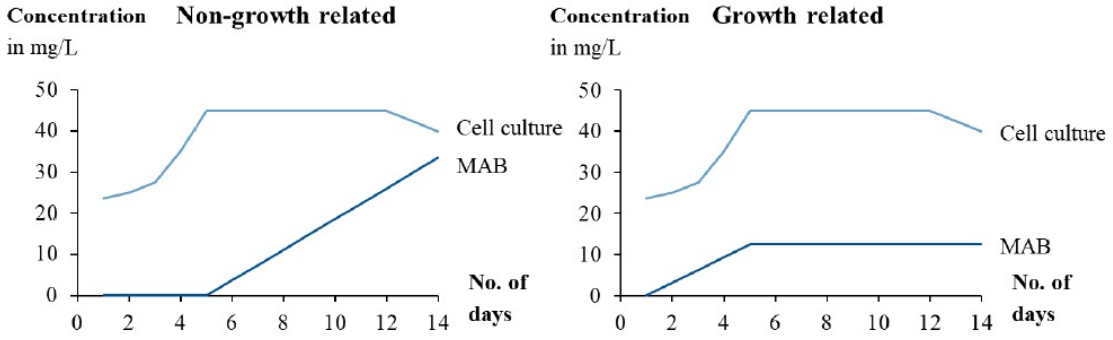

Figure 3: Comparison of non-growth related and growth-related production patterns, adopted from Schmidt (1996)

riorate by one step, or deteriorate by two steps. For example, a resin that yielded 95 % of the product in batch five, may also yield 95 % of batch six, or, more likely, only yield 90 % or even only 85 % of batch six (see some scenarios in Figure 4). In a worst-case scenario (double deterioration after every batch), this resin could reach its minimum capacity of 65% after only two more batches rather than six. For the edge case of the resin being one decay step away from its minimum capacity where a decay by two steps is not feasible, this work assumes that only one decay step is possible. Because of their short duration (Liu et al., 2014), maintenance activities related to restoring the resin capacity take only one decision epoch.

5.5. Optimality objective and optimization

The Markov decision process defined herein aims at determining the optimal stationary policy for operating upstream fermentation, harvesting, and downstream resin exchange. As argued by Martagan et al. (2016), discounting future rewards too strongly on this operations-planning horizon could bias the results. Therefore, the presented problem is studied under average rewards. To achieve this, the model's objective is to maximize the average reward over an infinite time horizon. Contrasting Equation 1, let the average expected reward criterion

$$\nu(s,\pi) \equiv \lim_{N\to\infty} \frac{1}{N} \mathbb{E}_s^{\pi} \left\{ \sum_{t=1}^{N} r(s(t), a^{\pi}(t)) \right\} \tag{14}$$

be defined for all starting states $s \in \mathcal{S}$ and all $\pi \in \Pi$ where Π is the set of all possible policies. Therein, $E_{s^{\pi}}$ is the expected value under policy π of starting in state s, dependent on the reward $r(s(t),a^{\pi}(t))$ of choosing the action $a^{\pi}(t)$ in the state $s(t)$ as prescribed by policy π. Therein, the objective is to find $V^*(s,\pi^*) \equiv sup_{\pi\in\Pi} \mathcal{V}(s,\pi)$. The average reward criterion problem is solved using the relative value iteration algorithm as implemented by Chades et al. (2014) in MatLab (The MathWorks Inc., 2018). While the standard value iteration algorithm does not converge to the optimal value function for average reward problems, the relative value iteration does so (Gupta et al., 2015). For robustness, the total expected discounted reward problem with a discount factor λ close to one is also solved to optimality. In line with Equation 1, let the total expected discounted reward criterion

$$\nu_{\lambda}^{\pi}(s) \equiv sup_{a\in A_s} \left\{ r(s,a) + \sum_{s'\in S} \lambda p(s,s',a)\, \nu(s') \right\} \tag{15}$$

be defined for all $s \in S$. The optimal policy π^* optimizes the objective $\mathcal{V}_{\lambda}^{*}(s,\pi^*) \equiv sup_{\pi\in\Pi} \mathcal{V}_{\lambda}^{\pi}(s,\pi)$. For this, the policy iteration algorithm, also implemented by Chades et al. (2014), is used. All model code is provided in Appendix 7 - Appendix 9.

6. Case Study Results

To investigate the posed research questions, a numerical case study of TPA production and purification, as introduced in Section 4.3, was implemented. To test specific hypotheses, sensitivity analyses were run; the resulting policies were critically analyzed and compared.

This task is complicated by the presented research questions pertaining to two different process setups, or USP to DSP ratios, i.e., how many parallel production reactors are harvested into one downstream process. A basic "1:1" model with one USP production reactor and one DSP, where $\mathcal{K} = \{U, D\}$, and a "2:1" model with two parallel production reactors and one DSP, where $\mathcal{K} = \{U_1, U_2, D\}$ were implemented based on the formulation in Section 5.1. Section 6.1 discusses the optimal control of the 1:1 process setup. Section 6.2 extends Section 6.1 by adding a second parallel production reactor in a 2:1 process setup and the implications thereof are discussed. All case study data is provided in Appendix 12 - Appendix 16.

6.1. Optimal, simultaneous control of a serial USP and the DSP

In this section, the 1:1 setup of one production reactor and one chromatography step is studied. In answering RQ1 and RQ2, the average per-cycle operating profit maximizing policy is presented and discussed. Furthermore, managerial insights are derived from the optimal reward's sensitivity to process parameter changes. Conclusions are also

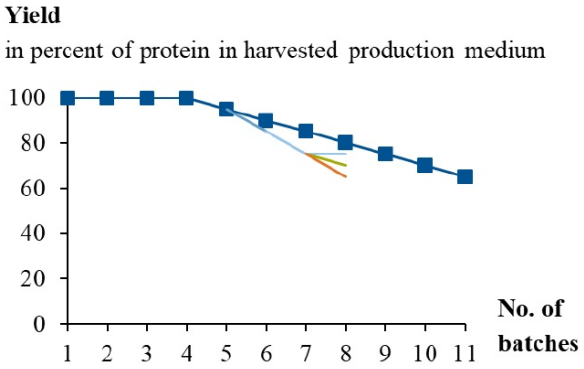

Figure 4: Deterministic performance decay pattern of a candidate chromatography resin over 11 purified batches, with selected stochastic decay scenarios starting after the fifth batch, adapted from Liu et al. (2014)

checked for robustness using a total expected discounted reward problem formulation and a formulation with longer decision epochs (36 instead of 12 hours). First, the resulting optimal policy is described in sub-section 6.1.1 before its theoretical and managerial implications are discussed (subsection 6.1.2) and further insights are derived from analyzing sensitivities (sub-section 6.1.3).

6.1.1. Optimal harvesting and resin exchange control

The average reward-maximizing policy for the simultaneous control of a single USP production reactor and the first chromatography step is depicted in Table 7.

The visualization of the optimal policies in this work is based on the format used by Schmidt (1996). As such, the rows in Table 7 represent the state of the production reactor; the columns represent the state of the chromatography column. Areas represent "decision zones" of states in which the optimal action vector is identical. The number pairs within each decision zone of the table represent the optimal action vector in those states of the system, in accordance with the coding scheme introduced earlier (Table 3 and Table 4). Thereby, the first digit of the ordered tuple is the optimal action to be taken in the production reactor, and the second digit is the optimal action in the chromatography column. For example, if the production reactor is within the last seven cycles of the production phase, and the resin has at least 75 % remaining capacity, it is optimal to harvest the production medium, return the production reactor to a ready state, and purify the harvested TPA, i.e., action vector (6,2) or (hprep, accept). The average reward of the optimal policy is $\mathcal{V}^{\pi^*} = 1,272\$$ of operating profit per 12-hour cycle, leading to the following observation.

> Observation 1: The maximum average per cycle operating profit is positive. Therefore, it may be economically feasible to operate the serial 1:1 production system under the assumed process parameters.

Should the system start with a chromatography resin at full performance, e.g., (empty,100 %), the optimal contingency plan (given no contamination) is to prepare the reactor for inoculation (prep), to grow the culture over five cycles, or 2.5 days, (addgm), to convert the growing culture into a TPA-producing culture by lowering its FCS-concentration in its first production-competent state, g6, (addpm) and to continue TPA-production for 29 cycles, or 14.5 days, (addpm) before harvesting in the 30th cycle and preparing the bioreactor for the next batch (hprep).

For almost all USP states, it is optimal to exchange the chromatography resin only when it is completely depleted (state Δ), i.e., only sometimes is it optimal to prematurely incur resin exchange costs. Between the 27th and 29th production cycle of the USP reactor, for example, resin exchange is optimal even though it's still at as much as 80 % of its full performance. In these cycles, the value of the TPA batch lies between 95,561 \$ and 102,912 \$, but if only 80 % of the TPA were bound, it would be worth only between 76,449 \$ and 82,330 \$. At this point, it is, therefore, optimal to exchange the resin (thus, incurring 96,480 \$ material net of fixed costs) and continue production until the 30th production cycle, where 100 % of the then c.4.4 g of TPA can be purified, realizing 106,587 \$ of operating revenue.

> Observation 2: Premature resin exchange seems to be optimal only sometimes. Robustly, it is optimal for resin performances under 85 %, before the culture reaches the optimal harvesting decision zone.

Policy-related insights remained robust under longer decision epochs and total expected discounted reward maximization. For the total expected discounted reward maximizing policy, policy iteration with a discount factor close to one ($\lambda = 0.99$) was used. The optimal total discounted reward function in the initial state $V^*(s^U = \text{empty}, s^D = 100\%) = 117,395\$$ was obtained. Observations 1 and 2 remained robust with two exceptions: (1) given remaining resin performance of 65% (70%), cultures in their first to fifth (first and second) growth cycle were dumped and the reactor prepared to start a new batch, and (2) resin exchange seemed more risk-seeking, i.e., premature resin exchange was postponed

Table 7: Average reward maximizing policy for the operation of one TPA production reactor

	100%	100%	100%	100%	95%	90%	85%	80%	75%	70%	65%	Δ
e	51 Prepare reactor for inoculation											53
r												
g_1												
...	21 Add growth medium to continue culture growth phase											23
g_5												
g_6												
g_7												
g_8												
p_1												
...												
p_4												
p_5												
p_6												
...	31 Add production medium to convert production-competent growth states, or to continue TPA production											
p_{17}												
p_{18}												
p_{19}												
...												
p_{26}												
p_{27}								33 Add growth medium in reactor, exchange chromatography resin				
p_{28}												
p_{29}												
p_{30}	62											
p_{31}												
...	Harvest production medium and prepare reactor for next batch											13
p_{36}												
u	61 Dump upset medium out											63

to the 65% remaining performance and production cycles 29 and 30. The optimal policies of the respective robustness checks are visualized in Appendix 10 and Appendix 11.

Following this description of the case study results, their theoretical and managerial implications are discussed in the next sub-section.

6.1.2. Theoretical and practical contributions

Fed-batch fermentation and purification of TPA under the assumed process parameters is economically feasible, yielding a positive average operating profit (see Observation 1). This is in line with the findings of Schmidt (1996). The average reward of the presented model, however, is c.88 % smaller than Schmidt's result. This may seem counterintuitive because of the presented model's support for increasing the TPA concentrations up to 33.5 mg/L and higher purification yields. However, because the presented model incorporates batch-ending harvesting dynamics and substantial resin material costs, the lower average reward is reasonable. To offset continuously incurred operating costs, revenues can only be realized at two points in a batch's lifecycle: scrapping it during growth (at its salvage value) or terminating fermentation by harvesting the TPA produced up to that point. Furthermore, harvesting seems to be optimal in the states just before the culture enters decline. The production medium is harvested in the 30th cycle already, given at least 75% resin capacity (Table 7). If the batch was not harvested in this state, the probability of successfully continuing would be reduced to 84%. This observation holds robust under total expected discounted reward maximization. Furthermore, it corroborates the findings of Martagan et al. (2016). In their study of IgG_1 harvesting decisions under uncertain fermentation dynamics, they observe the counterintuitive optimality of harvesting before the culture decline phase at the end of the fermentation time. However, they attribute this to the costs associated with the additionally produced by-products. In the presented study, impurities are not explicitly modeled, yet we observe a similar pattern. During advanced cycles of the production phase, resin exchanges are optimal before the minimum allowed capacity is reached whether resin decay is deterministic or uncertain. The earliest resin exchange is carried out after seven batches have been purified and the resin's performance has deteriorated to 80 %, i.e., in the state (p_{18}, 80 %) and the latest after the eleventh purified batch (Table 7). This is even earlier than previously reported under deterministic resin decay. Liu et al. (2014) found in their numerical study of fermentation and purification schedules that at most ten batches are

purified (leaving capacity at 65 %) before the chromatography resin is exchanged if the maintenance costs are equivalent to three-quarters of a full production batch. Comparing the scenario of stochastic decay to that of deterministic decay, the resin is exchanged after the same number of purified batches in both cases. Under deterministic resin decay[5] i.e., $p(\text{cap}_m, \text{cap}_{m+1}, 2) = 1, \forall m : \ 1 \leq m \leq \text{MaxBatch} - 1$ and $p(\text{cap}_{\text{MaxBatch}}, \Delta, 2) = 1$, we find that the resin is exchanged the earliest after the seventh batch has been purified, i.e., in the state (p_{29}, 80%).

While the effect of stochasticity on the timing of exchanges was limited, it had an effect on the optimal average operating profit obtained under stochastic decay compared to under deterministic decay. The potential downside of a decay by two steps seems to outweigh the potential upside of the resin not decaying, in terms of average operating profits. The sensitivity of the optimal average reward, $\mathscr{V}^*$, and the total expected discounted reward in the initial state, V^* (e, 100%), are compared to the proposed stochastic resin decay in Table 8. Moving from the proposed stochastic decay to a deterministic decay, ceteris paribus, yielded a 3% reduction in $\mathscr{V}^*$, whereas $\mathscr{V}^*$ (e,100%) remained virtually constant. In addition to its theoretical contributions, this case study also provides relevant managerial insights.

Biopharmaceuticals producers seem to benefit from the stochastic optimization of the integrated control of up- and downstream processes. We can compare the optimal policy to a primitive, production titer maximizing policy, π_{MaxTiter}, which might be used in biopharmaceutical manufacturing without decision support systems integrating system-level trade-offs in a worst-case scenario. From this comparison, the value of the presented optimal policy can be quantified. π_{MaxTiter} is provided in Appendix 18. Compared to π_{MaxTiter}, the optimal policy obtained from the presented model outperforms by a wide margin. Titer maximizing production under the assumed process parameters leads to a negative optimal total discounted rewards value function in the initial state, $V^{\pi^*}(s^U = empty, s^D = 100\%) = -76,311\$$. Therefore, while titer maximizing production may sound reasonable under consideration of only the upstream process, it may not be economically feasible on a system-level. The presented optimal policies, on the other hand, provide a business case for profitable production on a system-level.

Furthermore, by studying the sensitivity of average operating profits to changes in the fermentation reliability, managerial implications about the business case of reliability investments can be derived. While a reduction of the per-cycle probability of a successful transition has a negative effect on average and total discounted rewards, an increase in reliability by 0.2 percentage points (from 99.3% to 99.5%) has the potential to increase average and total discounted rewards by 6% and 7%, respectively (Table 9). The present value of this average operating profit increase in perpetuity is at least c.1.4 m€ (10% p.a. discount rate, 730 yearly cycles). Therefore, investments aimed at increasing process reliability, e.g., in newer fermentation equipment, are economically feasible up to this amount. The effect of further probability increases, however, showed opposite effects between the two measures. Their managerial interpretation should, therefore, be subject to further validation.

This work's answers to RQ1 and RQ2 extend academia's and practice's understanding of fed-batch fermentation control and harvesting decisions under simultaneous consideration of downstream purification and maintenance operations. The presented Markov decision process provides a dynamic framework for the stochastic optimization of biopharmaceutical operations. Furthermore, it corroborates existing research on the optimal harvesting decisions under fermentation uncertainties (Martagan et al., 2016; Schmidt, 1996), extends the literature on the timing of resin exchange activities (Liu et al., 2014), and aids practical decision making on an operational- and financial-level.

6.1.3. Sensitivity to varying resin exchange-related parameters

This sub-section answers RQ3 by discussing the optimal policy's sensitivity to process parameter variations relating to the chromatography resin economics. To this end, the 1:1 model analyzed above was run with different parameter sets and the effects were studied.

Because of its high economic impact, the timing of resin exchange, i.e., how long this costly activity is at least postponed, is particularly interesting for academia and practice. Decisions regarding the timing are mainly driven by the trade-off between the expected higher payoff of future harvests and the incurring of high material costs. This subsection, therefore, analyses the sensitivity of the timing of the earliest exchange activity to changes in related parameters, such as resin exchange costs and minimum viable resin capacities. The "earliness" of resin exchanges in a policy is operationalized as the least number of batches purified before any resin exchange. Visually, this is the upper-leftmost occurrence of the downstream sub-problem action 3 in the optimal policy as pictured in Table 7. In this specific case, that occurrence is in the state (p_{18}, 80%), meaning four batches have been purified at 100% and one batch each at 95, 90, and 85%. It can be found by traversing the optimal policy column-wise, starting from the upper left state (e, 100%). The optimal average reward was analyzed as a second important Key Performance Indicator (KPI).

> Observation 3: Higher resin material cost negatively influenced both monitored KPIs; higher minimum resin capacities led to higher average rewards.

Resin material cost had a negative impact on both the optimal average rewards and earliness of the resin exchanges. The rest of this thesis considers resin material costs, c_{resin}, equivalent to 0.75 times the value of a full TPA batch, i.e.,

[5] The average reward maximizing policy under deterministic decay is provided in Appendix 17

Table 8: Sensitivity of the optimal average reward and total expected discounted reward functions to changes in the resin decay probabilities in the DSP

Probability scenarios				$\mathcal{V}^*$	$\mathcal{V}^*(e, 100\%)$
	No decay	by one step	by two steps		
Base case	5%	90%	5%	1.00	1.00
Deterministic case	0%	100%	0%	0.97	1.00

Table 9: Sensitivity of the optimal average reward and total expected discounted reward functions to changes in the probability of a successful transition in the USP

Probability scenarios	$\mathcal{V}^*$[6]	$\mathcal{V}^*(e, 100\%)$[7]
Base case$p(s_t^U, s_{t+1-}U, a_{t+1}^U) = 99.3\%$	1.00	1.00
90%	0.40	0.40
99%	0.90	0.89
99.5%	1.06	1.07
99.8%	1.01	1.19
99.9%	0.76	1.23

96,480\$. For this analysis, c_{resin} is varied in the range [192,960\$, 189,0220\$ ..0\$], where 192,960\$ is equivalent to one and a half times the value of a full TPA batch. Figure 5 visualizes the negative correlation of more expensive resin exchanges and earliness of the exchange. For increasing resin costs, it becomes optimal to postpone the earliest exchange further. If the resin material were free, it can be optimal to exchange the chromatography resin as early as after four batches, i.e., when 100% of the produced TPA could be purified for the last time.

On the other extreme end of the range, for material costs of at least one and a half times the value of a full TPA batch, resin exchange is postponed until nine batches have been purified and the performance of the resin has decayed to 75%. Furthermore, the negative correlation of higher resin costs and the maximum achievable average operating profit is apparent in Figure 5. This relationship should not surprise. As the cost of replacing spent resin increases, resin exchanges are postponed further and further. This puts negative pressure on the average operating profit from three directions: (1) when resin exchange is carried out, higher one-time costs are incurred, (2) because the exchange is postponed, more batches are harvested at lower relative purification yields, and (3) because more batches are harvested before the resin is exchanged, operating costs from ongoing fermentation are incurred for longer. While the minimum allowed resin capacity had a strong negative impact on the earliness of the resin exchanges, its effect on the optimal average reward was modestly positive. Values for $\text{capacity}_{\text{MaxBatch}}$ were varied in the interval [100%,98% ..0%]. However, the maximum number of batches purified before the resin must be exchanged, MaxBatch, was not altered. Therefore, lower values of $\text{capacity}_{\text{MaxBatch}}$ imply a steeper decay profile after the fourth batch. Increasing minimum resin capacities showed a strong negative correlation with the earliness of resin exchange activities, as visualized in Figure 6. For higher minimum allowed purification yield and, therefore, a flatter resin decay pattern, more batches are purified before the earliest optimal resin exchange. For $\text{capacity}_{\text{MaxBatch}} \geq 90\%$, a premature exchange is no longer optimal. In these cases, the resin is always fully depleted and only exchanged in its final unusable state. In contrast to the effect which increasing resin costs had on the optimal average reward, higher minimum resin capacities led to slightly higher average operating profits (Figure 6). This positive correlation could be driven by the comparatively higher number of batches which can be purified at close to full yield due to the flatter slope of the resin's performance decline.

In providing an answer to RQ3, this contribution extends the existing literature on the scheduling of chromatography resin maintenance activities by a consideration of its system-level operational implications. Additional theoretical contributions are made by considering stochastic decay and different resin costs. The analyses in this work provide an understanding of the sensitivity of optimal maintenance timing and average profitability to these process parameter changes. The discussed sensitivities also aid managerial decision making under the uncertainties of biopharmaceutical production. The next section discusses the 2:1 case of two parallel production reactors being harvested into one chromatography step.

6.2. Optimal control of two parallel USPs and chromatography

In this section, the scenario of two parallel production reactors being harvested into one chromatography step in a 2:1 ratio is studied to answer RQ4. After describing the

[6] Average reward maximizing value function (normalized)

[7] Total expected discounted reward value function in the initial system state (normalized)

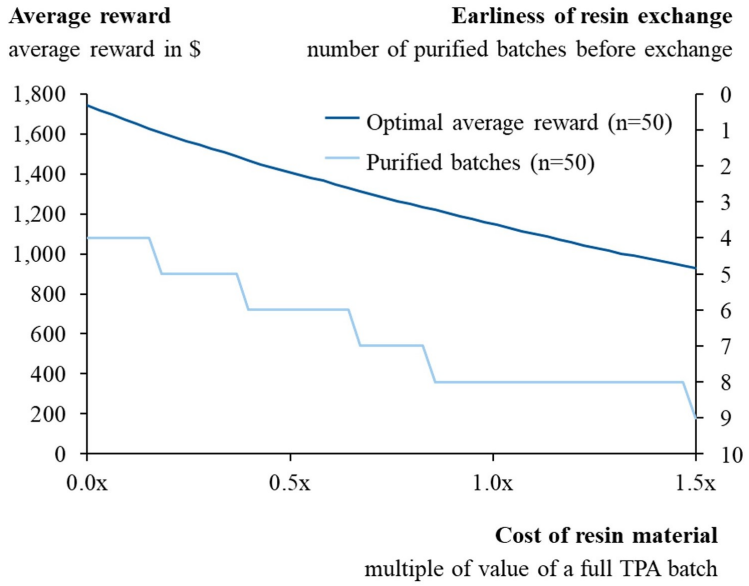

Figure 5: Sensitivity of the optimal average reward and the lateness of the first scheduled exchange to changes of the chromatography resin's material cost

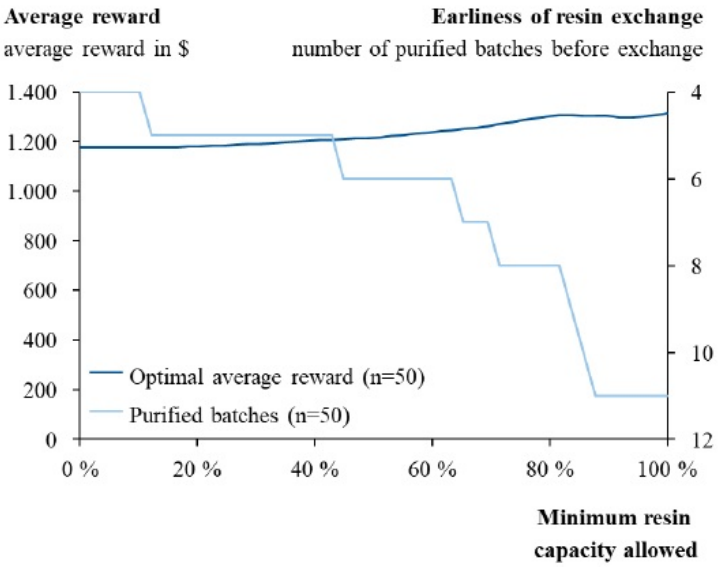

Figure 6: Sensitivity of the optimal average reward and the lateness of the first scheduled exchange to changes of the minimum allowed chromatography resin capacity

necessary adjustments to the model (sub-section 6.2.1), the optimal control policy is discussed (sub-section 6.2.2) and its sensitivity to process parameter changes analyzed (sub-section 6.2.3).

6.2.1. Model adjustment for the 2:1 case

This sub-section introduces the adjustments to the proposed model which were necessary to study parallel production reactors. Both production reactors are inoculated by a single seed-train (see Appendix 2). For this to be feasible, the following equation must hold

$$N_{PR} = \frac{46 \times 1}{\max(T)} \geq 2 \tag{16}$$

Therein, N_{PR} is the number of parallel production reactors which can be inoculated by a single seed-train, 1+1+8+36 = 46 is the number of 12-hour cycles in a production reactor from an empty reactor to the culture's first growth state until the latest possible time of harvesting the batch (see optimal policy in Table 7), 1 is the number of seed-trains feeding into

the production reactors, and $max(T)$ is the maximum seed-train occupancy time (Jagschies et al., 2018, p. 654). For $max(T) \leq 23$, i.e., at most 23 12-hour cycles or 11.5 days, it is feasible for a single seed-train to inoculate two production reactors. Because seed-train dynamics are not explicitly discussed in this contribution, this is assumed to be given.

Due to the exponential growth of the state space from adding a second reactor, the original model had to be scaled down to make computation feasible. This process is described in the following paragraph before the analysis of RQ4 commences.

Both parallel production reactors and the first chromatography step must be represented to study RQ4. However, the addition of a second, identical production reactor to the model drastically increased the state space from $|\mathscr{S}| = (1 + 1 + 8 + 36 + 1) \times 12 = 564$ states to $|\mathscr{S}| = (1 + 1 + 8 + 36 + 1)^2 \times 12 = 26,508$

The resulting transition matrix had c.12.6 billion entries, which made the calculation of optimal policies infeasible on available equipment. Therefore, to reduce the state space of the MDP and escape the curse of dimensionality, the number of decision epochs was scaled down linearly by a factor of three. Each decision epoch, therefore, represented 36 hours, compared to 12 hours in the basic model. This reduced the growth phase from eight epochs to three, and the production phase from 36 to twelve epochs. Additional repercussions on some process dynamics include: the culture decline phase at the end of the production phase was reduced to two transitions (a sixth of the now 12 production states) with probabilities 84% and 34%, respectively, and G_p only included the last growth state. The size of the resulting state space was $|S| = 3888$ and a transition matrix with c.272 million entries. The reduced scale made optimization of the 2:1 model feasible while holding the model robust (see robustness check for longer decision periods in the single reactor case above). Furthermore, the probability of a successful state transition had to be adjusted for the reduced number of cycles per week. Following Schmidt (1996) argumentation of a weekly contamination probability of 5% to 10% and the number of cycles per week scaled-down to from 14 to c.4.67, the probability of a successful cycle transition, $p(s^{U_l}, s'^{U_l}, a^{U_l})$, was assumed to be 97.8%.

6.2.2. Optimal control of the parallel reactors and the DSP

This sub-section discusses the optimal policy obtained during this variation of the TPA case study.

The optimal policy for the two production reactors if the chromatography resin is yet unused is presented in Table 10. Following the visualization scheme introduced in Section 6.1, the rows and columns represent the states of the two production reactors, respectively, while the remaining performance of the chromatography resin is held constant. The ordered triplets within each section prescribe the optimal action to take in the first reactor (represented by the rows), the second reactor (represented by the columns), and in the chromatography column. Average reward maximizing policies for all other resin capacities are provided in Appendix 19 - Appendix 23. Policy-related insights remain robust under total expected discounted reward maximization, although slight individual differences in the optimal policies exist.

We can't robustly conclude the economic feasibility of producing TPA in two parallel production reactors under the assumed process parameters. The average reward maximizing policy for the control of two parallel production reactors and one chromatography column is $V^* = -24,122$\$ per decision epoch. However, under the total expected discounted reward criterion, the optimal value function in the initial state is $V^*(e, e, 100\%) = 827,959$\$, and, therefore, strongly positive.

The decision zones of both production reactors are fairly symmetric, independent of the remaining resin performance. We can deduct this visually. For example, regard the state pair (p_{10}, p_{12}) in Table 10, i.e., reactor one is in its tenth production cycle and reactor two is in its twelfth. The optimal action is (3,6,2), i.e., to add production medium in the first reactor and harvest the second reactor. Conversely, in the state (p_{12}, p_{10}), the optimal actions are mirrored: (6,3,2). This symmetry holds for all USP reactor-state pairs except those in which both reactors are in the same state within the last third of their production phase, i.e., $(s^{U_1}, s^{U_1}) \in \{(p_j, p_j) : 9 \leq j \leq 12\}$. In these states, both reactors have a TPA concentration of at least 24.4 mg/L (batch value of at least 93,556\$).

This symmetry should not surprise the attentive reader. Contrary to a serial set-up, such as the one modeled by Schmidt (1996) where intra-reactor transfers between sequentially operating bioreactors are possible, interdependencies between parallel reactors only come into play in harvesting decisions. As described earlier (Section 5.2), both reactors can't be harvested simultaneously. Additionally, because purification only takes one decision epoch (by assumption, as discussed in Section 5.4), purification of one reactor's medium never blocks the other reactor.

> Observation 4: The decision zones of two parallel production reactors exhibit strong symmetry due to limited interdependency between the reactors them.

For the optimal average financial reward, both reactors are operated in a "staggered" fashion. Following the first harvested batch, one reactor is delayed from the other by one cycle. Regard the optimal path of a system with two parallel bioreactors and one DSP, starting from its initial state $(e, e, 100\%)$, in Table 11. Because simultaneous harvesting is not feasible, reactor two is harvested prematurely by one production cycle (in p9) and a new batch started, followed by harvesting reactor one in the next decision epoch. This leads to reactor two being one cycle further in its process than reactor one because the chromatography step only takes one decision epoch. This allows harvesting future batches of both reactors in the optimal tenth production cycle. For all subsequent batches until the tenth, it is optimal to harvest in their tenth production cycle, i.e., p_{10}, or on the 13th production phase day. These yield revenues of 105,251\$, 105,251\$,

Table 10: Optimal policy for the operation of two parallel bioreactors into a single chromatography column, before the first and second purified batch

	e	r g_1 g_2	p_{10} p_1 ... p_8	p_9 p_{10} p_{11} p_{12}	u
e	551	521	531	562	561
r g_1 g_2	251	221	231	262	261
p_{10} p_1 ... p_8	351	321	331	362 Continue production in reactor one, harvest reactor two	361
p_9 p_{10} p_{11} p_{12}	652	622	632	162	662
u	651	621	631	662	661

105,251\$, 99,988\$, 95,726\$, 89,463\$, 84,201\$, 78,938\$, 73,676\$, respectively. The eleventh batch is then purified in its eleventh production cycle, generating 76,015 \$ of operating revenue. Afterward, the resin is exchanged for the first time.

Furthermore, in states where the resin is already depleted to its minimum capacity, one batch is repeatedly dumped while the other is continued. Regard the sequence $(r, g_1, 65\%)$, $(g_1, g_2, 65\%)$, $(r, g_3, 65\%)$, $(g_1, p_1, 65\%)$ in Table 11. Because the resin is at its minimum capacity, operating costs are spent on producing only one batch (in reactor two) while the other (in reactor one) is repeatedly dumped (realizing its salvage value) and restarted in the first growth cycle.

6.2.3. Sensitivity to process parameter changes

Finally, this sub-section analyzes the optimal policy's sensitivity to process parameter changes.

> Observation 5: Both optimal average rewards and earliness of resin exchange were strongly sensitive to changes in the cost of resin material and the minimum allowed resin capacity.

Decreases in the cost of resin material, on the one hand, led to decreasingly increasing optimal average rewards. Below resin costs equivalent to half of the value of a full TPA batch, they led to almost linearly earlier resin exchanges. Firstly, cheaper resins are necessary for the profitable production and purification of TPA under the proposed process parameters. The average operating profit was negative for resin material costs of 0.75 times the value of a full TPA batch. While the single production reactor setup was profitable for all resin costs, the system of two parallel reactors only breaks even for reductions at least past 51,193\$ (c.0.24x) (Figure 7). However, reductions past 31,503\$(c.0.3x), show only small marginal profit improvements. Marginal gains per additional 4,000\$ cost reductions diminish to under 10%. Therefore, even if resin material were free, only a maximum of 937\$ of average operating profit per 36 hour-cycle would be earned. Additionally, the cheaper the resin material is, the earlier is its earliest exchange optimal. Figure 7 visualizes this correlation. Below the break-even point, the number of batches before the earliest resin exchange decreases almost linearly in resin cost decreases. On the other hand, both optimal average rewards and earliness of resin exchanges were decreasing in minimum allowed resin capacities. This contrasts with the single production reactor setup.

While delaying the earliest optimal resin exchange does not surprise, the decrease of the maximum average operating profit seems counterintuitive. Increasing the minimum allowed resin capacity flattens the decay profile and a higher number of batches can be harvested at relatively higher resin yields. Therefore, resin exchange becomes necessary later and later. Intuitively, harvesting more batches at almost full resin yield should also increase average rewards. This is not what we observe.

Per-cycle average fermentation operating costs surge in adding a second, parallel production reactor. It seems logical to conclude that the comparatively higher revenues from harvested batches can't fully compensate for this downward pressure on average profits. With the flatter decay curve, average revenues from harvested batches increase (for same protein concentrations at the time of harvesting). This leads to postponing the earliest resin exchange. Because more batches are purified before the resin is exchanged, fermentation operating costs are incurred longer – an effect which the higher average revenues observably don't fully counter.

Nevertheless, if minimum allowed resin performances are reduced below c.16%, thereby reducing the number of batches before the earliest resin exchange, operating the parallel production reactor setup turns profitable. At best, however, an average operating profit of only c.294 \$ is achieved. In answering RQ4, this section discussed the model adjustments necessary to study the simultaneous control of two parallel production reactors, discussed the optimal policies, and analyzed their sensitivity to process parameter changes.

Table 11: Optimal path of a 2:1 system, given no upsets

Path of system state left to right, starting at $(e, e, 100\%)$				
$(e, e, 100\%)$,	$(r, r, 100\%)$,	$(g_1, g_1, 100\%)$,	...	$(p_9, p_9, 100\%)$,
$(p_{10}, r, 100\%)$,	$(r, g_1, 100\%)$,	...	$(g_2, g_3, 100\%)$,	$(g_3, p_1, 100\%)$,
...	$(p_9, p_{10}, 100\%)$,	$(p_{10}, r, 100\%)$,	$(r, g_1, 100\%)$,	...
$(g_2, g_3, 100\%)$,	$(g_3, p_1, 100\%)$,	$(p_9, p_{10}, 100\%)$,	$(p_{10}, r, 95\%)$,	$(r, g_1, 95\%)$,
...	$(p_9, p_{10}, 95\%)$,	$(p_{10}, r, 90\%)$,	$(r, g_1, 85\%)$,	...
$(p_9, p_{10}, 85\%)$,	$(p_{10}, r, 80\%)$,	$(r, g_1, 75\%)$,	...	$(p_9, p_{10}, 75\%)$,
$(p_{10}, r, 70\%)$,	$(r, g_1, 65\%)$,	$(g_1, g_2, 65\%)$,	$(r, g_3, 65\%)$,	$(g_1, p_1, 65\%)$,
$(r, p2, 65\%)$,	$(g_1, p_3, 65\%)$,	...	$(p_6, p_{11}, 65\%)$,	(p_7, r, Δ),
$(p_8, g_1, 100\%)$,	$(p_9, g_2, 100\%)$,	$(r, g_3, 100\%)$,	...	

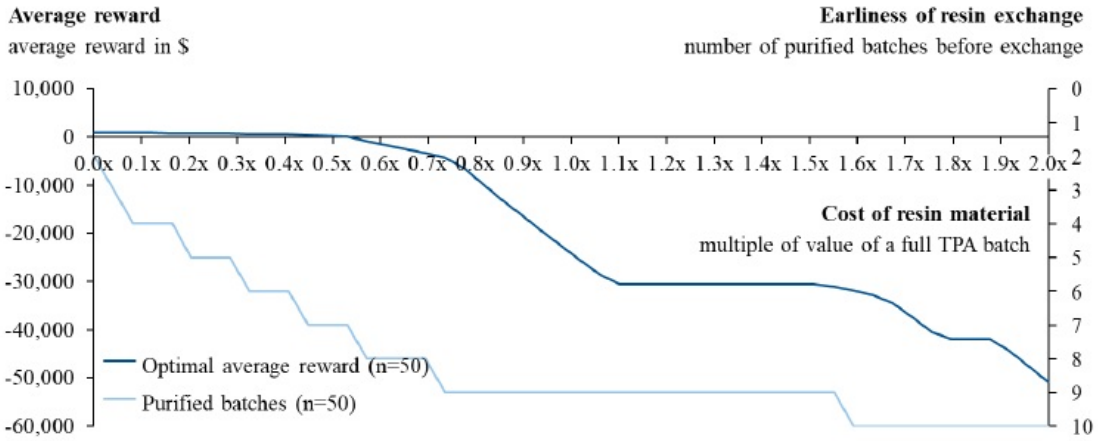

Figure 7: Sensitivity of the optimal average reward to changes of the chromatography resin's material cost

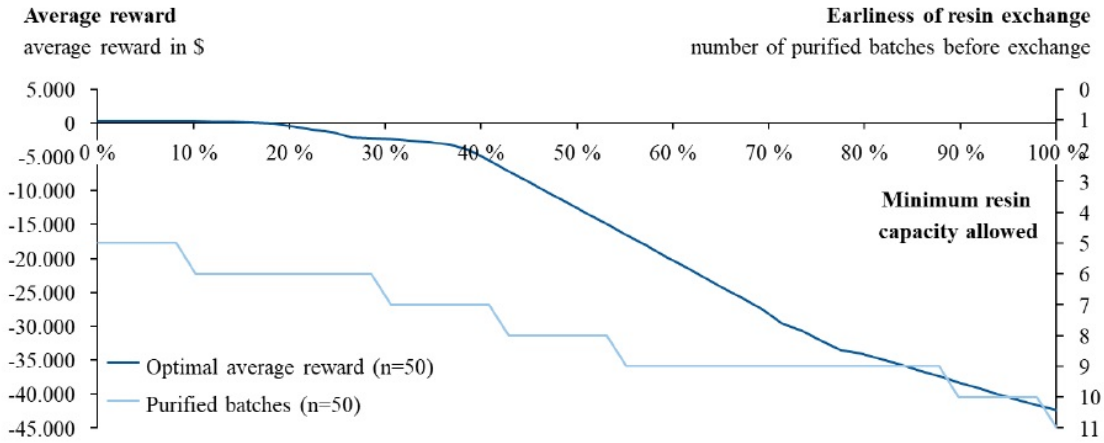

Figure 8: Sensitivity of the optimal average reward to changes of the minimum allowed chromatography resin capacity

In conclusion, the production and fermentation of TPA may not be economically feasible under the proposed setup, however, we were able to generate insights related to the staggered operation of the reactors and necessary parameter changes to make the setup profitable. The next chapter concludes this thesis and provides directions for future research endeavors on the topic.

7. Conclusions & Outlook

Producing biopharmaceutical products is a complex endeavor due to the high stochasticity of the living organisms used in fermentation and uncertainties about purification yields. Due to the high economic and curative value of biopharmaceutical APIs, optimal decision making during

their manufacture is especially important. Currently, operations research methods are under-appreciated in practice, even though, the academic discussion on the topic has been growing since the advent of biopharmaceutical production processes.

Academia has produced contributions on three different abstraction levels: single bioreactor control, capacity and operations planning, and the system-level control of integrated production processes. Literature on the bioreactor-level has dealt mostly with the maximization of protein concentrations constrained by dynamic models of cell-level kinetics. The capacity planning-level considers scheduling interdependencies between up- and downstream operations and the notion of uncertain or decaying chromatography resin yield. Furthermore, it introduced maintenance activities related to the exchange of spent chromatography resin. System-level literature, on the other hand, has modeled the dynamic control of entire production systems. However, it lacks work which considers integrated fermentation and purification control under relevant process uncertainties.

This work, therefore, was motivated by the paucity of literature examining the system-level control of interdependent up- and downstream processes under stochastic fermentation and resin decay. Thus far, no paper has simultaneously considered the upstream fermentation and downstream resin exchange control-problems. In the former sub-problem, decisions about growing the culture, its conversion into the protein-producing phase, and when to harvest the protein from one or multiple parallel production reactors must be made under the risk of costly batch failure due to contamination. After each purified batch, the capacity of the used chromatography resin to bind the protein of future harvests deteriorates stochastically. In the resin exchange sub-problem, therefore, decisions must be made under the trade-off between potentially reduced yields from future harvests and costly maintenance activities to restore the resin to its full performance.

This thesis aimed at answering the following four research questions. What is the optimal, simultaneous control policy for the USP and DSP? When is chromatography resin exchanged under stochastic decay? When under different minimum allowed resin capacities and different resin costs? How does the consideration of two parallel production reactors change the optimal policy?

To this end, an infinite horizon MDP with a discrete state space was proposed. The model allows for a single production reactor in the USP coupled with a single DSP, or, to depict common production realities, parallel USP production reactors coupled with a single DSP. It models the evolution of the production reactors over states, representing the cell culture lifecycle (growth, production, decline). Every decision epoch, the system controller decides whether to do nothing, prepare the empty bioreactor, add growth medium to facilitate culture growth, add production medium to start or continue protein production, or harvest the produced protein. When harvested, the production medium must be accepted into the chromatography column because intermediate storage is not assumed. At some point, the spent chromatography resin is exchanged for fresh resin to restore full chromatography performance.

Furthermore, the case study of TPA fermentation and purification was analyzed and optimal policies interpreted. Both production setups with one production reactor and two parallel reactors were implemented and the sensitivity of critical process parameters was analyzed. The case study results were already summarized in their respective sections in Chapter 6. Here they are synthesized to answer the research questions of this work.

In answering RQ1, we thusly postulate three takeaways: Firstly, the state of the downstream process influences optimal upstream decision making and vice versa. This extends the existing research on system level control in which only one process part was considered in detail while the other was strongly abstracted or omitted. Secondly, batches should optimally be harvested before the culture enters decline. This corroborates prior findings (Martagan et al., 2016) and aids operative decision making. Lastly, a business case for the simultaneous stochastic optimization of up- and downstream decision making exists compared to a rudimentary policy focusing only on the output of the upstream process.

For RQ2, we conclude that premature resin exchanges become optimal under stochastic decay. The timings are even earlier than existing findings under deterministic decay (Liu et al., 2014).

The findings support the hypotheses related to RQ3. Thusly, resin-related parameters affect the profitability and resin exchange timing as expected, aiding managerial decision making and capacity planners.

By considering two parallel production reactors in studying RQ4, we gained the following four insights: For one, the simultaneous operation of two parallel reactors was only interdependent at the point of harvesting, therefore exhibiting strong symmetry in their respective decision zones. Secondly, the financial results cause doubt over this production setup's financial viability. Because the parallelization of production reactors is a common practice, managers must critically examine their production lines for profitability. This corroborates findings on increasing costs of goods in the case of parallel USPs (Liu et al., 2016). Thirdly, parallel production reactors should be operated in a staggered fashion to compensate for the time it takes to harvest one batch. Lastly, contrasting the findings under RQ3, the results fail to support the hypothesis of how flatter resin decay affects financial results of parallel operation.

Nevertheless, generalizability of the presented research has some limitations and warrants future research on the topic.

The presented model incorporates a linear protein accumulation pattern, i.e., the amount of the API of interest in the production medium is linearly dependent on the number of passed production phase cycles. Martagan et al. (2016), however, argue for stochastic accumulation of protein during fermentation based on probability distributions from fermentation experiments. Future research, therefore, could extend

the presented discrete state space formulation into a continuous state space, allowing the modeling of stochastic accumulation based on experimentally derived probability distributions.

The model presented in this contribution makes some significant assumptions about the purification process. This contribution abstracts the purification decision space to whether to accept a harvest for chromatography. In reality, purification decisions are more intricate. Martagan et al. (2018), for example, model the practically relevant decision about which chromatography lanes to pool. Future research on the simultaneous control of fermentation and purification operations should include this level of complexity to better aid practical decision making.

This contribution studied the TPA production case study, which Schmidt (1996) proposed. While the argument for choosing this numerical exercise has been made, corroboration of the presented academic and managerial insights requires the study of additional cases. Future research could, for example, model the fermentation of an MAB, a class of biopharmaceuticals of high commercial prominence.

Furthermore, due to the lack of high-grade literature on the topic of stochastic chromatography performance decay, model assumptions regarding the decay probabilities should be validated in future experimental research.

Lastly, future research on the topic should draw a more concrete picture of the fermentation process by including additional process variables in the state description, such as nutrient concentration, pressure, pH, and oxygen, instead of merely state approximations. Furthermore, future models could consider more concrete control actions such as the substrate feed-rate to even better aid operations.

Ultimately, academia's progress in modeling bioprocesses continues to be hindered by our limited understanding of living organisms' complexities (Jagschies et al., 2018, p. 98). However, this contribution aids practice and academia in increasing their understanding of controlling production systems considering the presented complexities and argues for the simultaneous optimization of up- and downstream operations.

References

Abu-Absi, N. R., Kenty, B. M., Cuellar, M. E., Borys, M. C., Sakhamuri, S., Strachan, D. J., Hausladen, M. C., and Li, Z. J. Real time monitoring of multiple parameters in mammalian cell culture bioreactors using an in–line Raman spectroscopy probe. *Biotechnology and Bioengineering*, 108 (5):1215–1221, 2011. ISSN 00063592.

Allmendinger, R., Simaria, A. S., Turner, R., and Farid, S. S. Closed–loop optimization of chromatography column sizing strategies in biopharmaceutical manufacture. *Journal of Chemical Technology & Biotechnology*, 89 (10):1481–1490, 2014.

Ashoori, A., Moshiri, B., Khaki-Sedigh, A., and Bakhtiari, M. R. Optimal control of a nonlinear fed-batch fermentation process using model predictive approach. *Journal of Process Control*, 19(7):1162–1173, 2009.

Banga, J. R., Alonso, A. A., and Singh, R. P. Stochastic dynamic optimization of batch and semicontinuous bioprocesses. *Biotechnology Progress*, 13(3): 326–335, 1997.

Banga, J. R., Balsa-Canto, E., Moles, C. G., and Alonso, A. A. Dynamic optimization of bioprocesses: efficient and robust numerical strategies. *Journal of Biotechnology*, 117(4):407–419, 2005.

Bellman, R. Dynamic programming. *Princeton University Press*, 89:92, 1957a.

Bellman, R. A Markovian decision process. *Journal of Mathematics and Mechanics*, pages 679–684, 1957b.

Boucherie, R. J. and van Dijk, N. M., editors. *Markov decision processes in practice*, volume 248 of *International series in operations research & management science*. Berlin, Springer, Cham, Switzerland, 1st edition edition, 2017. ISBN 978-3-319-47764-0.

Chades, I., Chapron, G., Cros, M.-J., Garcia, F., and Sabbadin, R. MDPtoolbox: a multi-platform toolbox to solve stochastic dynamic programming problems. *Ecography*, 37(9):916–920 (v. 1.6), 2014. ISSN 0906-7590. doi: 10.1111/ecog.00888.

Chaudhuri, B. and Modak, J. M. Optimization of fed-batch bioreactor using neural network model. *Bioprocess Engineering*, 19(1):71–79, 1998. ISSN 1615-7605.

Datar, R. V., Cartwright, T., and Rosen, C.-G. Process economics of animal cell and bacterial fermentations: a case study analysis of tissue plasminogen activator. *Bio/Technology*, 11(3):349–357, 1993. doi: 10.1038/nbt0393-349.

Delvigne, F., Lejeune, A., Destain, J., and Thonart, P. Stochastic models to study the impact of mixing on a fed-batch culture of Saccharomycescerevisiae. *Biotechnology Progress*, 22(1):259–269, 2006.

Farid, S. S. Process economics of industrial monoclonal antibody manufacture. *Journal of Chromatography B*, 848(1):8–18, 2007.

Farid, S. S., Washbrook, J., and Titchener-Hooker, N. J. Decision–support tool for assessing biomanufacturing strategies under uncertainty: stainless steel versus disposable equipment for clinical trial material preparation. *Biotechnology Progress*, 21(2):486–497, 2005.

Freitas, H., Olivo, J., and Andrade, C. Optimization of bioethanol in silico production process in a fed-batch bioreactor using non-linear model predictive control and evolutionary computation techniques. *Energies*, 10 (11):1–23, 2017.

Gatica, G., Papageorgiou, L. G., and Shah, N. An aggregation approach for capacity planning under uncertainty for the pharmaceutical industry. *Foundations of Computer-Aided Process Operations*, 4:245–248, 2003.

Gupta, A., Jain, R., and Glynn, P. W. An empirical algorithm for relative value iteration for average-cost mdps. In *2015 54th IEEE Conference on Decision and Control*, pages 5079–5084. IEEE, 2015. ISBN 1479978868. doi: 10.1109/CDC.2015.7403014. URL https://doi.org/10.1109/CDC.2015.7403014.

Howard, R. A. *Dynamic programming and Markov processes*. Cambridge, MIT Press, Cambridge, MA, 1960.

Jagschies, G., Lindskog, E., Lacki, K., and Galliher, P. M. *Biopharmaceutical processing: Development, design, and implementation of manufacturing processes*. Amsterdam, Elsevier, Saint Louis, UK, 2018. ISBN 9780128125526. URL http://ebookcentral.proquest.com/lib/munchentech/detail.action?docID=5154046.

Jiang, C., Liu, J., Rubacha, M., and Shukla, A. A. A mechanistic study of Protein A chromatography resin lifetime. *Journal of Chromatography A*, 1216(31):5849–5855, 2009. ISSN 0021-9673.

Johnston, S. C. The economic case for new stroke thrombolytics. *Stroke*, 41 (10 Suppl):S59–S62, 2010.

King, A. J. and Wallace, S. W. Modeling with stochastic programming. In Mikosch, T., Resnick, S. I., and Robinson, S. M., editors, *Springer Series in Operations Research and Financial Engineering*. Berlin, Springer Science & Business Media, New York, NY, 2012.

Lakhdar, K. and Papageorgiou, L. G. An iterative mixed integer optimisation approach for medium term planning of biopharmaceutical manufacture under uncertainty. *Chemical Engineering Research and Design*, 3(86):259–267, 2008. ISSN 0263-8762.

Lakhdar, K., Zhou, Y., Savery, J., Titchener-Hooker, N. J., and Papageorgiou, L. G. Medium term planning of biopharmaceutical manufacture using mathematical programming. *Biotechnology Progress*, 21(5):1478–1489, 2005.

Lakhdar, K., Farid, S. S., Titchener-Hooker, N. J., and Papageorgiou, L. G. Medium term planning of biopharmaceutical manufacture with uncertain fermentation titers. *Biotechnology Progress*, 22(6):1630–1636, 2006.

Lee, J., Lee, S. Y., Park, S., and Middelberg, A. P. J. Control of fed-batch fermentations. *Biotechnology Advances*, 17(1):29–48, 1999.

Liu, C., Gong, Z., Shen, B., and Feng, E. Modelling and optimal control for a fed-batch fermentation process. *Applied Mathematical Modelling*, 37(3): 695–706, 2013.

Liu, S., Yahia, A., and Papageorgiou, L. G. Optimal production and maintenance planning of biopharmaceutical manufacturing under performance decay. *Industrial & Engineering Chemistry Research*, 53(44):17075–17091, 2014.

Liu, S., Farid, S. S., and Papageorgiou, L. G. Integrated optimization of upstream and downstream processing in biopharmaceutical manufacturing under uncertainty: a chance constrained programming approach. *Industrial & Engineering Chemistry Research*, 55(16):4599–4612, 2016.

Lourenço, N. D., Lopes, J. A., Almeida, C. F., Sarraguça, M. C., and Pinheiro, H. M. Bioreactor monitoring with spectroscopy and chemometrics: a review. *Analytical and Bioanalytical Chemistry*, 404(4):1211–1237, 2012. ISSN 1618-2650. doi: 10.1007/s00216-012-6073-9.

Markov, A. A. The theory of algorithms. *Trudy Matematicheskogo Instituta Imeni VA Steklova*, 42:3–375, 1954.

Martagan, T., Krishnamurthy, A., and Maravelias, C. T. Optimal condition-based harvesting policies for biomanufacturing operations with failure risks. *IIE Transactions*, 48(5):440–461, 2016.

Martagan, T., Krishnamurthy, A., Leland, P. A., and Maravelias, C. T. Performance guarantees and optimal purification decisions for engineered proteins. *Operations Research*, 66(1):18–41, 2018.

Martagan, T., Krishnamurthy, A., and Leland, P. A. Managing trade-offs in protein manufacturing: how much to waste? *Manufacturing & Service Operations Management*, Accepted/In press. ISSN 1523-4614.

Nweke, M. C., Rathore, A. S., and Bracewell, D. G. Lifetime and aging of chromatography resins during biopharmaceutical manufacture. *Trends in Biotechnology*, 36(10):992–995, 2018. ISSN 0167-7799.

Otto, R., Santagostino, A., and Schrader, U. Rapid growth in biopharma: challenges and opportunities, 2014. URL https://www.mckinsey.com/industries/pharmaceuticals-and-medical-products/our-insights/rapid-growth-in-biopharma.

Pandian, B. J. and Noel, M. M. Control of a bioreactor using a new partially supervised reinforcement learning algorithm. *Journal of Process Control*, 69:16–29, 2018.

Peroni, C. V., Kaisare, N. S., and Lee, J. H. Optimal control of a fed-batch bioreactor using simulation-based approximate dynamic programming. *IEEE Transactions on Control Systems Technology*, 13(5):786–790, 2005.

Ponte, X., Barrigón, J. M., Maurer, M., Mattanovich, D., Valero, F., and Montesinos-Seguí, J. L. Towards optimal substrate feeding for heterologous protein production in Pichia pastoris (Komagataella spp) fed-batch processes under P AOX1 control: a modeling aided approach. *Journal of Chemical Technology & Biotechnology*, 93(11):3208–3218, 2018.

Powell, W. B. Clearing the jungle of stochastic optimization. In *Bridging data and decisions*, INFORMS TutORials in Operations Research, pages 109–137. INFORMS, 2014.

Puterman, M. L. *Markov decision processes: discrete stochastic dynamic programming*. Hoboken, John Wiley & Sons, 2014.

Rader, R. A. What is a biopharmaceutical? (bio)technology-based definitions. *BioExecutive International*, 1:60–65, 2005.

Rani, K. Y. and Rao, V. S. R. Control of fermenters – a review. *Bioprocess Engineering*, 21(1):77–88, 1999. ISSN 1615-7605.

Rocha, M., Mendes, R., Rocha, O., Rocha, I., and Ferreira, E. C. Optimization

of fed-batch fermentation processes with bio-inspired algorithms. *Expert Systems with Applications*, 41(5):2186–2195, 2014.

Saucedo, V. M. and Karim, M. N. Experimental optimization of a real time fed-batch fermentation process using Markov decision process. *Biotechnology and Bioengineering*, 55(2):317–327, 1997. ISSN 00063592. doi: 10.1002/(SICI)1097-0290(19970720)55:2{\textless}317::AID-BIT9{\textgreater}3.0.CO;2-L.

Saucedo, V. M. and Karim, M. N. Real time optimal feeding in a fermentor using a Markov decision algorithm. *Computers & Chemical Engineering*, 22(4-5):547–558, 1998.

Schmidt, R. L. A stochastic optimization model to improve production planning and R&D resource allocation in biopharmaceutical production processes. *Management Science*, 42(4):603–617, 1996.

Siganporia, C. C., Ghosh, S., Daszkowski, T., Papageorgiou, L. G., and Farid, S. S. Capacity planning for batch and perfusion bioprocesses across multiple biopharmaceutical facilities. *Biotechnology Progress*, 30(3):594–606, 2014.

Skolpap, W., Scharer, J. M., Douglas, P. L., and Moo-Young, M. Fed-batch optimization of alpha-amylase and protease-producing Bacillus subtilis using Markov chain methods. *Biotechnology and Bioengineering*, 86(6): 706–717, 2004. ISSN 00063592.

Skolpap, W., Nuchprayoon, S., Scharer, J. M., Grisdanurak, N., Douglas, P. L., and Moo-Young, M. Fed-batch optimization of alpha-amylase and protease-producing bacillus subtilis using genetic algorithm and particle swarm optimization. *Chemical Engineering Science*, 63(16):4090–4099, 2008.

The MathWorks Inc. Matlab and statistics toolbox r2018b, 2018. URL https://www.mathworks.com/products/matlab.html.

Tijms, H. C. *A first course in stochastic models*. Hoboken, John Wiley & Sons, Chichester, UK, 2003. ISBN 9780470013632. doi: 10.1002/047001363X.

Walsh, G. Biopharmaceutical benchmarks 2014. *Nature Biotechnology*, 32 (10):992–1000, 2014.

Walsh, G. Biopharmaceutical benchmarks 2018. *Nature Biotechnology*, 36 (12):1137–1145, 2018.

Wechselberger, P., Sagmeister, P., and Herwig, C. Real-time estimation of biomass and specific growth rate in physiologically variable recombinant fed-batch processes. *Bioprocess Engineering*, 36(9):1205–1218, 2013. ISSN 1615-7605. doi: 10.1007/s00449-012-0848-4.

Xing, Z., Bishop, N., Leister, K., and Li, Z. J. Modeling kinetics of a large-scale fed-batch CHO cell culture by Markov chain Monte Carlo method. *Biotechnology Progress*, 26(1):208–219, 2010.