Submitted to *Operations Research*
manuscript (Please, provide the manuscript number!)

The Price of Flexibility

Dimitris Bertsimas
Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139, dbertsim@mit.edu

Hoda Bidkhori
Swanson School of Engineering, University of Pittsburgh, Pittsburgh, PA 15261, bidkhori@pitt.edu

Iain Dunning
DeepMind, London, United Kingdom, iaindunning@gmail.com

Process flexibility is a popular operations strategy that has been employed in many industries to help firms respond to uncertainty in product demand. However, additional flexibility comes at a cost that firms must balance against the reduction of risk it can provide. In this paper, we reduce the price of flexibility by taking an optimization approach. Unlike many approaches previously discussed in the literature, we consider systems that may have nonhomogenous parameters and unbalanced capacity and demand. We formulate the problem as a robust adaptive optimization model and propose a computationally tractable method for solving this model using standard integer optimization software. To further reduce the price, we consider Pareto-efficient and relaxed Pareto-efficient robust solutions and show that they correspond to flexibility designs that perform well in both worst-case and average-case demand scenarios. We demonstrate through simulation experiments that our method can find flexible designs that reduce the price of flexibility by approximately 30% or more versus common flexibility designs, with a negligible impact on revenues and thus higher profits. Furthermore, we provide theoretical results, we show that if we have two groups of similar products, it is more profitable to have short chains containing similar products, rather than a long chain. Furthermore, under some assumptions of symmetry, we show that a long chain design cannot be dominated by any other designs with the same number of links

Key words: Process flexibility, Manufacturing systems, Adaptive optimization, Robust optimization

1. Introduction

Process flexibility is an important strategy used by many manufacturing firms to manage uncertain product demand. Jordan and Graves (1995) define it as a firm's ability to "build different types of products in the same manufacturing plant or on the same production line at the same time." Flexibility enables, for example, a firm to utilize surplus capacity that may exist at some plants to produce products that have higher-than-expected demand that would otherwise exceed the capacity of any single plant. Many authors have related tales of both the successes of flexibility and the failures of inflexibility. For example, Biller et al. (2006) report that a failure by Chrysler to keep up with demand, despite having underutilized capacity at some plants, lead to an estimated loss of $240M in pretax profit. Mak and Shen (2009) relay media accounts that the Ford Motor Company made a $485M investment in 2002 to increase flexibility at two of their plants and to add flexibility at most of their engine and transmission plants worldwide; Chou et al. (2010) report that similar initiatives have been undertaken by both GM and Nissan.

The more products each plant is capable of producing, the more flexible the production system becomes. With a "full" flexibility *design* each plant can produce all products, while under a "dedicated" flexibility design each plant produces a single product (as visualized in Figure 1). While a full flexibility design places a firm in the best position to respond to uncertain demand, it is normally unrealistic from both a cost and management perspective for many firms to implement such a design. The natural goal then is to reduce the *price of flexibility* while realizing as many of the benefits of flexibility as possible, given considerations of profitability and feasibility.

Contributions The primary contribution of this paper is a group of methods based on *adaptive robust optimization* that reduce the price of flexibility by optimizing for *profitable* flexibility designs in unbalanced, nonhomogenous manufacturing systems in which (a) the number of plants and products may differ, (b) the plants may have varying capacities, (c) the products may have varying marginal profit by plant, and (d) the costs for flexibility may vary between plants. We demonstrate through simulations that our method can reduce the price by 30% or more versus standard flexibility designs described in the literature, with negligible effects on revenues. A secondary contribution is the theoretical results related to the application of *Pareto efficiency* to adaptive robust optimization. In particular, we demonstrate that Pareto efficient solutions can be found with little additional computational effort and show through computation that considering Pareto efficiency improves the designs produced by our methods. As a third contribution, we provide theoretical proofs to show that certain design guidelines that perform better than a long chain. Unlike what has been discussed in the literature, we prove that under certain conditions short chains perform better than a long chain.

Previous work Designs with a low price of flexibility will typically be those in which each plant produces a small number of products and thereby gains most of the benefit of full flexibility at only a fraction of the cost. Jordan and Graves (1995) represents one of the first comprehensive studies into the effectiveness of partial flexibility designs. In particular, they seek to explore the effect of designs on expected sales and capacity utilization and offer a convincing demonstration that "chaining" is a key property of effective designs. If process flexibility is viewed as a network of links between plants and products, then a long chain design is one in which a "cycle" can be traced between any product and any plant by following these links. (See the center of Figure 1). Jordan and Graves (1995) quantify both analytically and by simulation that such designs are capable of approximating the performance of a full flexibility design with substantially fewer links. However, they do not address the problem of determining a design for any specific firm or setting; in those cases, the costs of adding flexibility at each plant could vary widely. For some pairs of plant and product, adding flexibility may not even be possible.

Since the work of Jordan and Graves (1995) many researchers have attempted to analytically explain the observed effectiveness of long chain and other partial flexibility designs, mainly in *balanced* homogeneous systems (i.e., systems in which the number of plants and products are the same, all plants have equal capacities, and all products face identical demand distributions). For example, in terms of average performance, Chou et al. (2010) provide theoretical performance

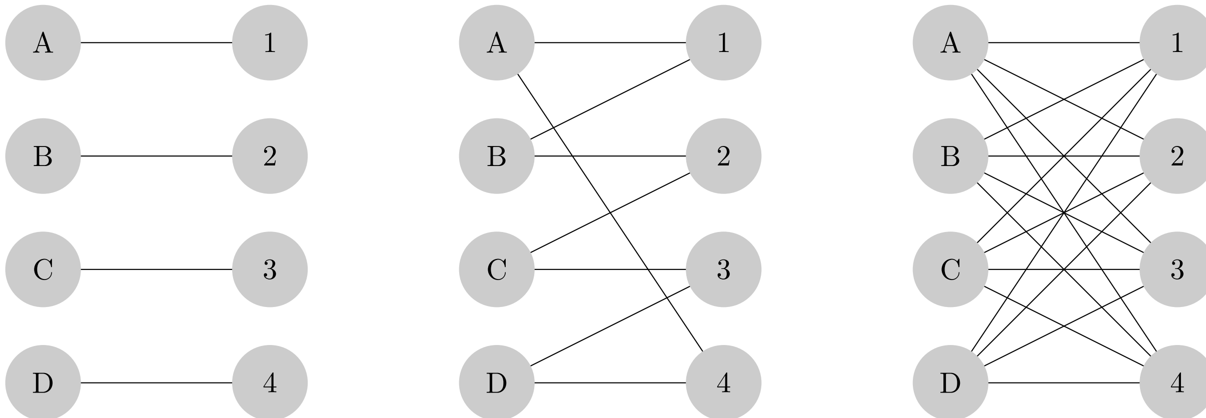

Figure 1 Examples of process flexibility designs. Left: "dedicated", Center: "long chain", Right: "full".

guarantees for long chain designs under some distributions, including two-point, uniform, and normal distributions. Simchi-Levi and Wei (2012) proved that, in a balanced and symmetrical system with exchangeable random demand, the long chain design maximizes the expected maximum sales among all "2-regular" flexibility designs. More recently, Wang and Zhang (2015) established a lower bound on the ratio between the expected sales for a long chain design and a full flexibility design when all products share an identical demand distribution. Worst-case performance of the long chain and other sparse designs for balanced and homogeneous systems has also been discussed by Wei and Simchi-Levi (2015).

Many authors have proposed guidelines to identify effective partial flexibility designs, including Chou et al. (2010), Deng and Shen (2013), Chou et al. (2011), and Chen et al. (2014). These guidelines consist of heuristics, and they assume homogeneity in relevant parameters such as identical demand distributions. To the best of our knowledge, the only paper addressing the problem of finding optimal flexibility designs in unbalanced, nonhomogeneous manufacturing systems is by Mak and Shen (2009). The authors propose a stochastic programming approach, with demand modeled as belonging to a known distribution from which a finite set of scenarios can be drawn. The resulting model is computationally challenging to solve, so in order to find feasible solutions they propose a relaxation that sacrifices solution quality and requires a complex implementation. They demonstrate that, despite the relaxation and in some settings, they are able to produce designs (a) identical to long chain designs, and (b) considerably better than long chain designs in settings where a "generic" long chain design has a high price, primarily due to the use of high and nonhomogeneous costs.

Structure We have structured the paper as follows:

- In Section 2, we model the process flexibility design problem as a robust adaptive optimization problem, where the demand for each product is treated as an uncertain parameter of the model, and the objective is to maximize profits and thereby reduce the price of flexibility. We present a *relative* robust objective variant, and detail how our model can be reformulated as a computationally tractable deterministic optimization problem solvable with off-the-shelf optimization software. We provide guidance on various ways the uncertain demand can be modeled within our optimization model depending on the risk preferences of management and available data. Finally, we demonstrate the extensibility and generality of our model, including how it can incorporate uncertainty in other parameters (e.g., disruptions in capacity) and restrictions on capabilities.
- In Section 3, we introduce the notion of *Pareto efficiency* to adaptive robust optimization, and apply it to manufacturing process flexibility and we provide a method to design pareto efficient flexibility configurations. We develop a theoretical result that demonstrates how we can obtain flexibility designs that perform well in both "worst-case" and "average-case" demand scenarios for little extra computational effort. We present a new notion of *relaxed-Pareto efficiency* that allows us to further explore the trade-offs between worst-case and average-case performance according to the preferences of a decision maker.
- In Section 4, we obtain some theoretical results about the efficiency of a long chain design versus multiple short chains. We show that when there are two groups of products which are similar within groups and different between groups, it is beneficial to have two short chains which connect similar products. Finally, under some assumptions of symmetry, we show that a single long chain design cannot be dominated by any other designs with the same number of links.
- In Section 5, we first provide computational evidence that we can solve instances of practical size quickly on commodity hardware (e.g., instances with 15 plants and 15 products in one hour).

We next provide insights derived from comparing the relative merits of different flexibility designs as viewed through our optimization model, which qualitatively reproduce some of the insights provided by Jordan and Graves (1995) and others. We then quantify to what degree Pareto efficient designs improve over initial designs, and then compare these Pareto efficient designs with designs obtained by a stochastic programming model under a variety of parameter choices. We finally compare the designs produced by our method with alternatives such as long chains and find substantial cost savings with little impact on revenue, and thus higher profits.

2. Process flexibility as adaptive robust optimization

In this section, we present an optimization model that can be used to reduce the price of flexibility. We achieve this by optimizing for a profitable design that is both robust with respect to demand uncertainty and able to incorporate nonhomogenous capacities, costs, and revenues.

We consider systems with n plants, each with a capacity $c_i \geq 0$, and with m products. We model the demand for the products as an uncertain vector $\mathbf{d} = (d_1, \ldots, d_m)$. As the demand is not known with certainty, we treat the problem of flexibility design as a two-stage process: the design itself must be decided *here-and-now*, but the production decision for how much of the demand for each product to satisfy from each plant is a *wait-and-see* decision that is deferred until after the demand is realized.

Multiple techniques have been developed in the optimization literature for modeling uncertainty in optimization problems. Here we have elected to take a *robust optimization* (RO) perspective, wherein we model the uncertain demand vector as being drawn from a known *uncertainty set* $\mathcal{U}$. Instead of assigning a probability distribution over the demand scenarios, as is often done in the stochastic programming literature, we will instead optimize with respect to the *worst-case realization* of the demand over the uncertainty set (for a review of RO, please refer to the survey by Bertsimas et al. (2011)). Our model thus represents an *adaptive robust optimization* (ARO) approach, with a $\max - \min - \max$ objective: we initially select a flexibility design (max), "nature" adversarially selects a demand scenario from the uncertainty set (min), and finally we select a production plan to maximize profits (max) and reduce the price of flexibility.

2.1. Adaptive robust optimization model

Perhaps the most natural objective is to maximize total profit. Let $i \in \{1, \ldots, n\}$ index over the plants, and $j \in \{1, \ldots, m\}$ index over the products. We can enable the flexibility for plant i to produce product j in exchange for a one-time investment cost of B_{ij}. We define T_{ij} to be the production cost for manufacturing one unit of product j from plant i, and define p_j to be the revenue for selling one unit of product j. The ARO formulation for this objective is

$$\max_{\mathbf{x},\mathbf{y}(\mathbf{d})} \min_{\mathbf{d}\in\mathcal{U}} \sum_{i,j} (p_j - T_{ij})\, y_{ij}(\mathbf{d}) - \sum_{i,j} B_{ij} x_{ij} \tag{1}$$

$$\text{subject to} \quad \sum_j y_{ij}(\mathbf{d}) \le c_i \qquad \forall i,\ \forall \mathbf{d}\in\mathcal{U} \tag{1a}$$

$$\sum_i y_{ij}(\mathbf{d}) \le d_j \qquad \forall j,\ \forall \mathbf{d}\in\mathcal{U} \tag{1b}$$

$$0 \le y_{ij}(\mathbf{d}) \le d_j x_{ij} \qquad \forall i,j,\ \forall \mathbf{d}\in\mathcal{U} \tag{1c}$$

$$\mathbf{x} \in \{0,1\}^{n\times m},$$

where x_{ij} are binary decisions for the flexibility design (whether to enable plant i to produce product j) that must be made here-and-now, and $y_{ij}(\mathbf{d})$ are the continuous wait-and-see production decisions (how much of the demand for product j to satisfy from plant i) that can be made after $\mathbf{d}$ is known. Note that we moved the inner maximization problem over the production decisions to outside the minimization over the uncertainty set: in this perspective, we are optimizing here-and-now over a space of possible functions $\mathbf{y}(\mathbf{d})$. These two representations are equivalent, but the representation we have selected here simplifies the discussion of computational details in Section 5. The objective (1) is the marginal profit from selling the products less the costs of the flexibility design. Constraint (1a) restricts the total amount of production at each plant i to be no more than the the plant's capacity. Constraint (1b) restricts the total amount produced and sold for each product j to be no more than the demand for that product. Finally constraint (1c) forces the production of product j from plant i to be zero if plant i is unable to produce product j (that is, if $x_{ij} = 0$).

2.2. Relative profit adaptive robust optimization model

While the absolute profit objective function of (1) is easy to understand, it may not be the best choice when coupled with a RO model of uncertainty for reducing the price of flexibility. Flexible designs excel in a setting where unused capacity at some plants can be used to satisfy higher-than-expected product demand. However, the objective function of (1) does not value such designs particularly highly. As evidence of this, consider fixing the flexibility design and determining the worst-case demand scenario, which leaves only the marginal profit component of the objective function. We can observe that the worst-case scenario will be one in which all the demands are low, as this minimizes total profit and penalizes designs that add flexibility:

EXAMPLE 1. Pessimism of model (1): Consider a problem with $n=5$ plants and $m=5$ products. We will assume that all plants are identical, as are all products. The profit p_j for all products is 1, and already includes production costs (i.e., $T_{ij}=0$). We assume that the dedicated flexibility design is already in place ($B_{ii}=0$), and we are interested in whether it is worth adding extra flexibility at a cost $B_{ij}=5$. The uncertainty set restricts each d_j to the range $[50,150]$ with a restriction that no more than two of the products can deviate from their "mean" demand of 100. We set the capacity of each plant c_i to be equal to 110, slightly more than the mean demand.

If we consider a dedicated design where we add no additional flexibility, our objective function simplifies to

$$\min_{\mathbf{d}\in\mathcal{U}} \sum_{i,j=1}^{5} y_{ij}(\mathbf{d}) = \sum_{k=1}^{5} \min\{c_k, d_k\}. \tag{2}$$

The worst-case scenario is one in which demand is as low as possible, giving a profit of 400. Consider now a "long chain" design: this design requires adding flexibility to each of the 5 plants at a total cost of 25. The worst-case demand is again the case where demand is low, e.g. $\mathbf{d}=[50,50,100,100,100]$, so profit for the long chain design will be 375 (worse than a dedicated design). Indeed, for these parameters any flexibility design with more flexibility than the dedicated design will never have a better worst-case profit that the dedicated design as, in the low demand scenario, no benefit is realized from the flexibility yet we still pay a price for the flexibility. □

We thus propose a modification to the objective function (1) that will make the model less "pessimistic" and produce solutions that we would expect to perform well in both worst-case scenarios and more "average" scenarios as well. In particular, we propose to optimize for a worst-case *relative* measure of profit, where we normalize the profits for our design by *total revenue*:

$$\max_{\mathbf{x},\mathbf{y}(\mathbf{d})} \min_{\mathbf{d}\in\mathcal{U}} \frac{\sum_{i,j} (p_j - T_{ij})\, y_{ij}(\mathbf{d}) - \sum_{i,j} B_{ij} x_{ij}}{\sum_j p_j d_j}. \tag{3}$$

The total revenue can be viewed as an optimistic estimate of what profits could be achieved, and so by normalizing by total revenue the worst-case scenario is no longer the low-demand scenario. This objective function is similar to a "competitive ratio" or "relative regret" objective function. Simchi-Levi and Wei (2015) consider a similar competitive ratio criterion for assessing flexibility designs in manufacturing, while in the field of network revenue management Ball and Queyranne (2009) consider competitive ratios and Perakis and Roels (2010) consider robust regret objective functions. While total revenue is not the only possible measure that profits can be normalized by, it has the benefit of being interpretable, in the same units as profit, and being linear in the demand (with positive implications for computational tractability).

Our ARO model for manufacturing flexibility with a relative profit objective function can thus be expressed as

$$\begin{aligned}
\max_{\mathbf{x},\mathbf{y}(\mathbf{d}),z} \quad & z && (4)\\
\text{subject to} \quad & \sum_{i,j} (p_j - T_{ij})\, y_{ij}(\mathbf{d}) - \sum_{i,j} B_{ij} x_{ij} \geq \sum_j p_j d_j z && \forall \mathbf{d} \in \mathcal{U}\\
& \sum_j y_{ij}(\mathbf{d}) \leq c_i && \forall i,\ \forall \mathbf{d} \in \mathcal{U}\\
& \sum_i y_{ij}(\mathbf{d}) \leq d_j && \forall j,\ \forall \mathbf{d} \in \mathcal{U}\\
& 0 \leq y_{ij}(\mathbf{d}) \leq d_j x_{ij} && \forall i,j,\ \forall \mathbf{d} \in \mathcal{U}\\
& \mathbf{x} \in \{0,1\}^{n\times m},
\end{aligned}$$

where we have converted the ratio objective implied by (18) into a linear uncertain constraint by using an epigraph formulation with auxiliary variable z. In our preliminary experiments, this formulation usually produced better designs than (1), so we will focus on (4) throughout the paper.

EXAMPLE 2. Example 1 with a relative regret objective: We now revisit the previous problem, using model (4). Considering first the dedicated design (no additional flexibility), we note that the worst-case demand scenarios had the form $\mathbf{d} = (50, 50, 100, 100, 100)$ – however, under the new objective function we would have a ratio of 1. If we consider scenarios such as $\mathbf{d} = (150, 150, 100, 100, 100)$ instead, we obtain a ratio of $(2 \times 110 + 3 \times 100)/600 \approx 0.867$ (as the denominator is not limited by capacity), which is the worst-case scenario for this design. The long chain design in the low demand scenario obtains a ratio of $375/400 \approx 0.938$, worse than the dedicated design. However, the worst-case demand scenario for the ratio objective is again the high-demand scenario. By using the power of flexibility, we can satisfy $5 \times 110 = 550$ out of the total demand of 600, achieving a ratio of $(550 - 25)/600 = 0.875$, which is superior to the ratio for the dedicated design. We would thus prefer the long chain design for these parameters. □

We will now discuss how to solve problem (4), before addressing the selection of the uncertainty set $\mathcal{U}$ (Section 2.4) and possible extensions (Section 2.5).

2.3. Solving the ARO model

To this point we have deferred discussing in detail how to obtain a solution to problem (4). The key difficulty that must be addressed is that the optimal wait-and-see decision $\mathbf{y}(\mathbf{d})$ may have a complex relationship with demand $\mathbf{d}$, making the optimization problem challenging to solve. Ben-Tal et al. (2004) have shown that, even for the case where the design $\mathbf{x}$ is fixed and we thus have only continuous decisions, this problem is hard in both the computational complexity sense and in practice. A popular remedy, first applied to the problem of ARO by Ben-Tal et al. (2004), is to drastically simplify the problem by restricting the space of admissible adaptive continuous decisions to those that are affine with respect to the uncertain parameters $\mathbf{d}$. This simplification is referred to as *affine adaptability* or *linear decision rules* in the literature, and has been applied successfully in application areas such as project management (Chen et al. 2007) and supply chains (Ben-Tal et al. 2005). It also been shown from a theoretical point of view (Bertsimas et al. 2010, Bertsimas and Goyal 2012, Bertsimas and Bidkhori 2014) that, in many cases, affine adaptability is optimal or a good approximation of the fully adaptive ideal.

We now apply the concept affine adaptability to our model. Let

$$y_{ij}(\mathbf{d}) = \mathbf{Q}_{ij}^T \mathbf{d} + q_{ij}, \tag{5}$$

where $\mathbf{Q}_{ij} \in \mathbb{R}^m$ and $q_{ij} \in \mathbb{R}$ are the decision variables that we will now optimize over. We can substitute (5) into (4) to obtain the problem

$$\begin{aligned}
\max_{\mathbf{x},\mathbf{Q},\mathbf{q},z} \quad & z && (6)\\
\text{subject to} \quad & \sum_{i,j} (p_j - T_{ij}) \left(\mathbf{Q}_{ij}^T \mathbf{d} + q_{ij}\right) - \sum_{i,j} B_{ij} x_{ij} \geq \sum_j p_j d_j z && \forall \mathbf{d} \in \mathcal{U}\\
& \sum_j \left(\mathbf{Q}_{ij}^T \mathbf{d} + q_{ij}\right) \leq c_i && \forall i,\ \forall \mathbf{d} \in \mathcal{U}\\
& \sum_i \left(\mathbf{Q}_{ij}^T \mathbf{d} + q_{ij}\right) \leq d_j && \forall j,\ \forall \mathbf{d} \in \mathcal{U}\\
& 0 \leq \mathbf{Q}_{ij}^T \mathbf{d} + q_{ij} \leq d_j x_{ij} && \forall i,j,\ \forall \mathbf{d} \in \mathcal{U}\\
& \mathbf{x} \in \{0,1\}^{n \times m},
\end{aligned}$$

which can be viewed as a relaxation of (4), and thus the objective value will be an upper bound on the fully adaptive problem's objective value. In return for this approximation we obtain a robust optimization problem with constraints that are linear with respect to the decision variables for fixed values of $\mathbf{d}$, and are linear with respect to $\mathbf{d}$ for fixed values of the decision variables. As a result, we are able to reformulate (6) to a deterministic mixed-integer optimization with convex constraints for many choices of $\mathcal{U}$. For example, if $\mathcal{U}$ is a polyhedron, then the resulting reformulation is a mixed-integer linear optimization problem, and if $\mathcal{U}$ is ellipsoidal, then the reformulation will be a mixed-integer second-order cone optimization problem. In Section 5.1 we provide computational evidence that (6) can be easily solved on commodity hardware for instances of realistic size.

2.4. Selecting the uncertainty set

One of the key modeling decisions that we have deferred to this point is the selection of the uncertainty set of possible demand scenarios. There are many possible choices, but key considerations include

- whether the resulting reformulation of (6) is tractable,
- what data we have available about the uncertain demand, and
- how robust we seek to be against uncertainty in the demand.

Given the wide variety of choices available we will focus here on polyhedral *budget* uncertainty sets based on the uncertainty set described in Bertsimas and Sim (2004). Other alternatives include sets such as the ellipsoidal set presented in Ben-Tal and Nemirovski (1999) and "data-driven" sets such as those described in Bertsimas et al. (2013).

The first uncertainty set we will describe, the *budget* uncertainty set, is defined by the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$. In particular, for each product j we have a nominal (or average) demand parameter μ_j and a deviation parameter σ_j. Both these parameters can be either estimated from data (for example, a sample mean and standard deviation), or from some other source (such as estimates by a domain expert). The uncertainty set is then the polyhedron

$$\mathcal{U} = \left\{ \mathbf{d} \ \middle| \ d_j = \mu_j + \sigma_j z_j \ \forall j \in \{1, \dots, m\}, \ \|\mathbf{z}\|_1 \le \Gamma, \ \|\mathbf{z}\|_\infty \le 1 \right\}, \tag{7}$$

where Γ is the *budget of uncertainty*. When $\Gamma = 0$ the uncertainty set reduces to the nominal demand scenario, i.e., $\mathcal{U} = \{\boldsymbol{\mu}\}$, and when $\Gamma = m$ the set reduces to the "hypercube" $\mathcal{U} = \{\mathbf{d} | \mu_j - \sigma_j \le d_j \le \mu_j + \sigma_j\}$. We can interpret Γ (when Γ is an integer) as controlling the number of products whose demand can vary from the nominal case. While Bertsimas and Sim (2004) provide some theoretical guidance on how one can select Γ, in practice the best approach is to solve the problem for multiple values of Γ and select amongst the solutions based on some metric of interest, e.g., performance of the optimal designs in simulation. We sketch this process, which is related to the process of cross-validation for parameter tuning in machine learning, in our computational experiments in Section 5.4.

The budget uncertainty set (7) assumes there is no correlation between products. If we have some estimation available of the covariance between products, e.g., a covariance matrix $\mathbf{\Sigma}$, then we can create a *correlated budget* uncertainty set

$$\mathcal{U} = \left\{ \mathbf{d} \;\middle|\; d_j = \mu_j + \sum_k \sigma_{jk} z_k \;\forall j \in \{1, \dots, m\},\; \|\mathbf{z}\|_1 \leq \Gamma,\; \|\mathbf{z}\|_\infty \leq 1 \right\}. \tag{8}$$

As a covariance matrix is positive-definite by construction, we can easily obtain $\mathbf{\Sigma}^{\frac{1}{2}}$ by performing a Cholesky factorization. The use of this uncertainty set does not make the optimization problem significantly more difficult to solve, so if sufficient data or expertise is available to estimate $\mathbf{\Sigma}$, we would expect higher quality and less pessimistic solutions with the correlated budget uncertainty set. We explore the effects of correlation on the relative merits of different possible designs in Section 5.2.

2.5. Extensibility of formulation

Part of the benefit of our ARO approach to flexibility is that incorporating extensions is relatively simple. For example, process flexibility puts the firm in a better position to match available capacity with variable demand, as described in the introduction. However, many firms are also interested in *operational flexibility*, which in the context of this problem could refer to the possibility of allocating extra capacity at each plant after the demand is realized.

To incorporate operational flexibility in our model we will define $s_i(\mathbf{d})$ to be the amount of additional extra capacity we will allocate at plant i as a function of the uncertain demand $\mathbf{d}$. Let the cost of additional capacity be h_i per unit. We can now extend (4) to include this new decision:

$$\begin{aligned}
\max_{\mathbf{x},\mathbf{y}(\mathbf{d}),\mathbf{s}(\mathbf{d}),z} \quad & z && (9)\\
\text{subject to} \quad & \sum_{i,j} (p_j - T_{ij})\, y_{ij}(\mathbf{d}) - \sum_{i,j} B_{ij} x_{ij} - \sum_i h_i s_i(\mathbf{d}) \geq \sum_j p_j d_j z && \forall \mathbf{d} \in \mathcal{U}\\
& \sum_j y_{ij}(\mathbf{d}) \leq c_i + s_i(\mathbf{d}) && \forall i,\ \forall \mathbf{d} \in \mathcal{U}\\
& \sum_i y_{ij}(\mathbf{d}) \leq d_j && \forall j,\ \forall \mathbf{d} \in \mathcal{U}
\end{aligned}$$

$$0 \le y_{ij}(\mathbf{d}) \le d_j x_{ij} \qquad \forall i,j,\ \forall \mathbf{d} \in \mathcal{U}$$

$$\mathbf{x} \in \{0,1\}^{n\times m},$$

where $\mathbf{s}(\mathbf{d})$ appears in the objective function constraint and in the constraint that controls the maximum amount of demand satisfied by plant i. As with (4), we can replace $\mathbf{s}(\mathbf{d})$ with an affine adaptability approximation, for example

$$s_i(\mathbf{d}) = \mathbf{R}_i^T \mathbf{d} + r_i, \tag{10}$$

which we can then substitute into (9) as we did with $\mathbf{y}$.

A partial list of possible extensions that can be incorporated into this model includes:

- *capacity disruption*, where $\mathbf{c}$ is itself uncertain, and the production decisions are a function of both the realized demand and the realized available capacity, e.g., $\mathbf{y}(\mathbf{d},\mathbf{c})$.
- *side constraints*, that allow for capturing more real-life restrictions. For example, linear constraints that could be added that enforce restrictions such as "plant 1 cannot produce both products 1 and 2".

3. Pareto efficiency in manufacturing operations

Our model is aimed at reducing the price of flexibility by finding flexibility designs that maximize relative profitability with respect to the worst-case demand scenarios contained within our uncertainty set. We do not explicitly consider performance under more average-case scenarios when optimizing, but ideally our design would also perform well across a range of realizations. This notion is closely related to the concept of *Pareto efficiency*: we wish to find designs that are not *dominated* by any another design. In this case, that means designs that are optimal in the worst-case, but there does not exist another design that is also optimal in the worst-case, but is better for some (or many) "average" scenarios. This concept was first applied in the context of robust optimization setting by Iancu and Trichakis (2013), who consider the general RO problem

$$z^* = \max_{\mathbf{x}\in X} \min_{\mathbf{d}\in\mathcal{U}} f(\mathbf{d},\mathbf{x}), \tag{11}$$

where z^* is the optimal objective value, and X^* is the set of all optimal solutions (that is, all solutions whose worst-case objective value is equal to z^*). A solution $\mathbf{x} \in X^*$ is said to be a *Pareto robustly optimal solution* if there is no other $\tilde{\mathbf{x}} \in X^*$ such that $f(\mathbf{d}, \tilde{\mathbf{x}}) \geq f(\mathbf{d}, \mathbf{x})$ for all $\mathbf{x} \in \mathcal{U}$ and $f(\tilde{\mathbf{d}}, \tilde{\mathbf{x}}) > f(\tilde{\mathbf{d}}, \mathbf{x})$ for some $\tilde{\mathbf{d}} \in \mathcal{U}$.

In this section, we extend the work of Iancu and Trichakis (2013) to two-stage ARO problems, and provide a tractable method for finding these Pareto robustly optimal solutions. We then define a notion of *relaxed Pareto efficiency* that allows us to systematically trade-off worst-case performance for average case performance. By doing so, we can obtain superior designs which further reduce the price of flexibility.

3.1. Pareto efficiency for two-stage problems

In order to motivate the need to consider Pareto efficiency, consider the following example that demonstrates that giving consideration to demand scenarios other than the worst-case can improve the performance of the final chosen designs:

EXAMPLE 3. Consider a setting with five plants and five products. We set the price $p_j = 1$ for all products, and set $T_{ij} = 0$. Assume that a dedicated design is already in place, and we are interested in whether it is worth adding flexibility ($B_{ij} = 9$). We will use a budget uncertainty set, with $\mu_j = 100$, $\sigma_j = 50$, and $\Gamma = 3$.

If we now solve the ARO model (6), we find that there are multiple optimal solutions. In particular, the dedicated design (i.e., no additional flexibility) and a long chain design both obtain an optimal objective function value of 0.7. To distinguish between the two, we can consider other demand scenarios and the objective function values of the two designs under those scenarios. For example:

- A "mixed" demand scenario $\mathbf{d}' = (50, 150, 75, 125, 100)$: the long chain design value is 0.91, and the dedicated design value is 0.85.
- A "low" demand scenario $\mathbf{d}' = (50, 50, 75, 100, 125)$: long chain design value is approximately 0.89, and the dedicated design value is approximately 0.94.

We now formalize this process by first considering the generic two-stage ARO problem:

$$z^* = \max_{\mathbf{x},\mathbf{y}(\mathbf{d})} \min_{\mathbf{d}\in\mathcal{U}} f(\mathbf{d},\mathbf{x},\mathbf{y}(\mathbf{d})) \tag{12}$$

$$\text{subject to } (\mathbf{x},\mathbf{y}(\mathbf{d})) \in X(\mathbf{d}) \quad \forall \mathbf{d}\in\mathcal{U},$$

which is a generalization of (11), and where z^* is the optimal objective value.

Definition 1. A *Pareto optimal adaptive solution* is an optimal solution $(\mathbf{x},\mathbf{y}(\mathbf{d}))$ to (12) such that there is no other optimal solution $(\tilde{\mathbf{x}},\tilde{\mathbf{y}}(\mathbf{d}))$ such that $f(\mathbf{d},\tilde{\mathbf{x}},\tilde{\mathbf{y}}(\mathbf{d})) \geq f(\mathbf{d},\mathbf{x},\mathbf{y}(\mathbf{d}))$ for all $\mathbf{d}\in\mathcal{U}$ and there does not exist a $\tilde{\mathbf{d}}\in\mathcal{U}$ such that $f(\tilde{\mathbf{d}},\tilde{\mathbf{x}},\tilde{\mathbf{y}}(\tilde{\mathbf{d}})) > f(\tilde{\mathbf{d}},\mathbf{x},\mathbf{y}(\tilde{\mathbf{d}}))$.

In the above example both designs are Pareto optimal adaptive solutions, as neither dominates the other across all demand scenarios.

We use affine adaptability to create a tractable relaxation of the "fully" adaptive original ARO problem. We can restate (12) by substituting in an affine policy (5) to obtain:

$$z^* = \max_{\mathbf{x},\mathbf{Q},\mathbf{q}} \min_{\mathbf{d}\in\mathcal{U}} f(\mathbf{d},\mathbf{x},\mathbf{Q},\mathbf{q}) \tag{13}$$

$$\text{subject to } (\mathbf{x},\mathbf{Q},\mathbf{q}) \in X(\mathbf{d}) \quad \forall \mathbf{d}\in\mathcal{U},$$

and thus obtain a new definition for this specific case:

Definition 2. A *Pareto optimal affine adaptive solution* is an optimal solution $(\mathbf{x},\mathbf{Q},\mathbf{q})$ to (13) such that there is no other optimal solution $(\tilde{\mathbf{x}},\tilde{\mathbf{Q}},\tilde{\mathbf{q}})$ such that $f(\mathbf{d},\tilde{\mathbf{x}},\tilde{\mathbf{Q}},\tilde{\mathbf{q}}) \geq f(\mathbf{d},\mathbf{x},\mathbf{Q},\mathbf{q})$ for all $\mathbf{d}\in\mathcal{U}$ and there does not exist a $\tilde{\mathbf{d}}\in\mathcal{U}$ such that $f(\tilde{\mathbf{d}},\tilde{\mathbf{x}},\tilde{\mathbf{Q}},\tilde{\mathbf{q}}) > f(\tilde{\mathbf{d}},\mathbf{x},\mathbf{Q},\mathbf{q})$.

The affine policy substitution reduces the problem to nearly the same definition as provided by Iancu and Trichakis (2013), with the exception that in this model both the constraints and the objective function contain uncertain parameters.

We can now describe a general method to find a Pareto optimal affine adaptive solution. We first solve the ARO problem (6) to obtain the worst-case objective function value z^*. We then solve a second optimization problem:

$$\max_{\mathbf{x},\mathbf{Q},\mathbf{q}} \sum_{i,j} (p_j - T_{ij}) \left(\mathbf{Q}_{ij}^T \hat{\mathbf{d}} + q_{ij} \right) - \sum_{i,j} B_{ij} x_{ij} \tag{14}$$

$$\begin{aligned}
\text{subject to} \quad & \sum_{i,j} (p_j - T_{ij}) \left(\mathbf{Q}_{ij}^T \mathbf{d} + q_{ij} \right) - \sum_{i,j} B_{ij} x_{ij} \geq \sum_j p_j d_j z^* && \forall \mathbf{d} \in \mathcal{U} \\
& \sum_j \left(\mathbf{Q}_{ij}^T \mathbf{d} + q_{ij} \right) \leq c_i && \forall i, \ \forall \mathbf{d} \in \mathcal{U} \\
& \sum_i \left(\mathbf{Q}_{ij}^T \mathbf{d} + q_{ij} \right) \leq d_j && \forall j, \ \forall \mathbf{d} \in \mathcal{U} \\
& 0 \leq \mathbf{Q}_{ij}^T \mathbf{d} + q_{ij} \leq d_j x_{ij} && \forall i,j, \ \forall \mathbf{d} \in \mathcal{U} \\
& \mathbf{x} \in \{0,1\}^{n \times m},
\end{aligned}$$

where $\hat{\mathbf{d}}$ is a feasible demand scenario drawn from $\mathcal{U}$. The first constraint ensures that a solution to this problem is no worse in the worst-case than the solution to (6), and the objective function ensures that we are not dominated by another design for the selected demand scenario. We claim that a solution to this second problem will be a Pareto optimal affine adaptive solution.

To understand why solving (14) is sufficient, consider the optimization problem:

$$\begin{aligned}
\max_{\mathbf{x},\mathbf{Q},\mathbf{q}} \quad & \sum_{k=1}^{m} \left[\sum_{i,j} (p_j - T_{ij}) \left(\mathbf{Q}_{ij}^T \mathbf{d}_k + q_{ij} \right) - \sum_{i,j} B_{ij} x_{ij} \right] && (15) \\
\text{subject to} \quad & \sum_{i,j} (p_j - T_{ij}) \left(\mathbf{Q}_{ij}^T \mathbf{d} + q_{ij} \right) - \sum_{i,j} B_{ij} x_{ij} \geq \sum_j p_j d_j z^* && \forall \mathbf{d} \in \mathcal{U} \\
& \sum_j \left(\mathbf{Q}_{ij}^T \mathbf{d} + q_{ij} \right) \leq c_i && \forall i, \ \forall \mathbf{d} \in \mathcal{U} \\
& \sum_i \left(\mathbf{Q}_{ij}^T \mathbf{d} + q_{ij} \right) \leq d_j && \forall j, \ \forall \mathbf{d} \in \mathcal{U} \\
& 0 \leq \mathbf{Q}_{ij}^T \mathbf{d} + q_{ij} \leq d_j x_{ij} && \forall i,j, \ \forall \mathbf{d} \in \mathcal{U} \\
& \mathbf{x} \in \{0,1\}^{n \times m},
\end{aligned}$$

where $\mathbf{d}_1, \ldots, \mathbf{d}_{m+1}$ are affinely independent demand scenarios in $\mathcal{U} \subset \mathbb{R}^m$.

THEOREM 1. *Optimal solutions to Problem* (15) *are Pareto optimal affine adaptive solutions.*

Proof of Theorem 1 First, we note that the first constraint of (15) is sufficient to constrain us to only robustly optimal affine solutions, as we can be no worse than z^* under any scenario. We now proceed to prove it is Pareto optimal by contradiction. Let $(\mathbf{x}, \mathbf{Q}, \mathbf{q})$ be an optimal solution, but not a Pareto optimal one. Therefore, there exists an optimal solution $(\tilde{\mathbf{x}}, \tilde{\mathbf{Q}}, \tilde{\mathbf{q}})$ that has an objective function value greater than or equal to the objective function value for $(\mathbf{x}, \mathbf{Q}, \mathbf{q})$ for all $\mathbf{d} \in \mathcal{U}$, and is strictly greater for at least one scenario. Equivalently, the inequality

$$\sum_{i,j} (p_j - T_{ij}) \left(\mathbf{Q}_{ij}^T \mathbf{d} + q_{ij}\right) - \sum_{i,j} B_{ij} x_{ij} \leq \sum_{i,j} (p_j - T_{ij}) \left(\tilde{\mathbf{Q}}_{ij}^T \mathbf{d} + \tilde{q}_{ij}\right) - \sum_{i,j} B_{ij} \tilde{x}_{ij} \tag{16}$$

holds for all $\mathbf{d} \in \mathcal{U}$. However, as we are optimizing for (15), we know that the following $m+1$ equalities must hold for $\mathbf{d}_1, \ldots, \mathbf{d}_{m+1}$ for an optimal solution:

$$\begin{aligned}
\sum_{i,j} (p_j - T_{ij}) \left(\mathbf{Q}_{ij}^T \mathbf{d}_1 + q_{ij}\right) - \sum_{i,j} B_{ij} x_{ij} &= \sum_{i,j} (p_j - T_{ij}) \left(\tilde{\mathbf{Q}}_{ij}^T \mathbf{d}_1 + \tilde{q}_{ij}\right) - \sum_{i,j} B_{ij} \tilde{x}_{ij} \\
\sum_{i,j} (p_j - T_{ij}) \left(\mathbf{Q}_{ij}^T \mathbf{d}_2 + q_{ij}\right) - \sum_{i,j} B_{ij} x_{ij} &= \sum_{i,j} (p_j - T_{ij}) \left(\tilde{\mathbf{Q}}_{ij}^T \mathbf{d}_2 + \tilde{q}_{ij}\right) - \sum_{i,j} B_{ij} \tilde{x}_{ij} \\
\vdots \quad &= \quad \vdots \\
\sum_{i,j} (p_j - T_{ij}) \left(\mathbf{Q}_{ij}^T \mathbf{d}_{m+1} + q_{ij}\right) - \sum_{i,j} B_{ij} x_{ij} &= \sum_{i,j} (p_j - T_{ij}) \left(\tilde{\mathbf{Q}}_{ij}^T \mathbf{d}_{m+1} + \tilde{q}_{ij}\right) - \sum_{i,j} B_{ij} \tilde{x}_{ij}
\end{aligned}$$

Recall that any $\mathbf{d} \in \mathcal{U} \subset \mathbb{R}^m$ can be generated by $m+1$ affinely independent vectors, for example $\mathbf{d}_1, \ldots, \mathbf{d}_{m+1}$. Since the two solutions $(\mathbf{x}, \mathbf{Q}, \mathbf{q})$ and $(\tilde{\mathbf{x}}, \tilde{\mathbf{Q}}, \tilde{\mathbf{q}})$ take the same value for $m+1$ affinely independent points in $\mathcal{U}$, and the objective function is an affine function of $\mathbf{d}$, they must take the same value for all other points in $\mathcal{U}$ as linear combination of $\mathbf{d}_1, \cdots, \mathbf{d}_{m+1}$. Therefore, the inequality (16) cannot be strict for any $\mathbf{d} \in \mathcal{U}$ and we obtain a contradiction. □

Using Theorem 1 we can then obtain the following result, which justifies our method for finding Pareto solutions:

COROLLARY 1. *Optimal solutions to Problem* (14) *are Pareto optimal affine adaptive solutions.*

Proof of Corollary 1 Any interior point $\hat{\mathbf{d}} \in \mathcal{U}$ can be written as a linear combination of $m+1$ affinely independent demand scenarios in $\mathcal{U}$. Therefore, the optimization problem (14) reduces to (15). □

Solving (14) is no more difficult than solving the initial optimization problem, and may in fact be easier as an optimal solution to (6) is a feasible solution to (14). In Section 5.3, we quantify the degree to which the simulation performance of Pareto optimal designs improves over initial designs, and how often these improvements occur.

3.2. Relaxed Pareto efficiency

The applicability of the above approach rests on there being multiple designs with the same worst-case performance to select among. In practice, this may not occur due to idiosyncrasies of a given instance.

To address this problem we suggest a simple, pragmatic extension to our method for finding Pareto optimal designs. In particular, instead of searching for a non-dominated solution amongst all robust optimal solutions, we will instead search amongst all "α-robust" optimal solutions, where $\alpha \in (0,1]$ is the degree to which we are willing to relax the requirement that we are no worse in the worst-case. Formally, we wish to find solutions to the problem:

$$\begin{aligned}
\max_{\mathbf{x},\mathbf{Q},\mathbf{q}} & \sum_{i,j} (p_j - T_{ij}) \left(\mathbf{Q}_{ij}^T \hat{\mathbf{d}} + q_{ij} \right) - \sum_{i,j} B_{ij} x_{ij} && (17) \\
\text{subject to} & \sum_{i,j} (p_j - T_{ij}) \left(\mathbf{Q}_{ij}^T \mathbf{d} + q_{ij} \right) - \sum_{i,j} B_{ij} x_{ij} \geq \sum_j p_j d_j \alpha z^* && \forall \mathbf{d} \in \mathcal{U} \\
& \sum_j \left(\mathbf{Q}_{ij}^T \mathbf{d} + q_{ij} \right) \leq c_i && \forall i, \ \forall \mathbf{d} \in \mathcal{U} \\
& \sum_i \left(\mathbf{Q}_{ij}^T \mathbf{d} + q_{ij} \right) \leq d_j && \forall j, \ \forall \mathbf{d} \in \mathcal{U} \\
& 0 \leq \mathbf{Q}_{ij}^T \mathbf{d} + q_{ij} \leq d_j x_{ij} && \forall i,j, \ \forall \mathbf{d} \in \mathcal{U} \\
& \mathbf{x} \in \{0,1\}^{n \times m},
\end{aligned}$$

where $\hat{\mathbf{d}}$ is a given demand scenario in the interior of $\mathcal{U}$, and z^* is the optimal objective value to (6). We will refer to the designs produced by this optimization as *relaxed Pareto optimal designs*. We present computational evidence that this relaxation can lead to higher-quality designs than non-relaxed Pareto-efficient designs in Section 5.3, thereby reducing the price of flexibility even further.

4. Insights from Theoretical Results

In this section we show two main results. We show that when we have two separate groups of similar products, it is efficient to use a couple of short chains, each connecting similar plants and products, as it is less expensive for a plant to manufacture similar products. Additionally, we show that, under some symmetry assumptions, a long chain design cannot be dominated by any other design with the same number of links in all scenarios.

4.1. Chaining Similar Products

In this section, we show that if we have two groups of similar products that are very different from one another, it is better to make two short chains that connect similar products. In this model, there is a cost for a plant to produce multiple similar products, but this cost is less if the plant produces multiple significantly different products. We prove the following lemma:

LEMMA 1. *Suppose we have $n+m$ plants and $n+m$ products. All plants have the same capacity c. The demand uncertainty $\mathcal{U}$ can be anything. Suppose $p_j = p$, $T_{ij} = T$. The fixed cost of adding flexible links between first n plants and products is B_1. The fixed cost of adding flexible links between second m plants and products is B_2. The cost of adding flexible links between these two groups is B_3. When $2B_3 > B_1 + B_2 + 2(p-T)c$, two short chains connecting the first n and second m plants and products perform better than a long chain.*

Proof of Lemma 1 We look at two designs: $[\mathcal{C}_n, \mathcal{C}_m]$ and $\mathcal{C}_{n+m}$, and we consider the relative profit function. It is enough to show that the relative profit for $\mathcal{C}_{n+m}$ is less than the relative profit for $[\mathcal{C}_n, \mathcal{C}_m]$.

Let $\mathbf{x}$ be a $0,1$ vector associated with $[\mathcal{C}_n, \mathcal{C}_m]$. Suppose the optimal relative profit for $[\mathcal{C}_n, \mathcal{C}_m]$ occurs in vector $\bar{\mathbf{d}} = (\bar{\mathbf{d}}_1, \bar{\mathbf{d}}_2)$, where $\bar{\mathbf{d}}_1$ is a vector containing first n coordinates of $\bar{\mathbf{d}}$, and $\bar{\mathbf{d}}_2$ is a vector containing the next m coordinates. The relative profit for $[\mathcal{C}_n, \mathcal{C}_m]$ can be expressed as follows:

$$\max_{\mathbf{y}(\bar{\mathbf{d}})} \frac{\sum_{i,j} (p-T)\, y_{ij}(\bar{\mathbf{d}}) - \sum_{i,j} B_{ij} x_{ij}}{\sum_j p \bar{d}_j}. \tag{18}$$

Clearly for $[\mathcal{C}_n, \mathcal{C}_m]$, the $\max \sum_{i,j} y_{ij}(\bar{\mathbf{d}}) = vc(\mathcal{C}_n, \bar{\mathbf{d}}_1) + vc(\mathcal{C}_m, \bar{\mathbf{d}}_2)$, where $vc(\mathcal{C}_n, \bar{\mathbf{d}}_1)$ and $vc(\mathcal{C}_m, \bar{\mathbf{d}}_2)$ are the costs of minimum vertex covers for $\mathcal{C}_n$ (with capacities c facing a demand vector $\bar{\mathbf{d}}_1$) and $\mathcal{C}_m$ (with capacities c facing a demand vector $\bar{\mathbf{d}}_2$), respectively. If we add the nth plant and the

$(n+m)$th plant to these vertex covers, we obtain a vertex cover for $\mathcal{C}_{n+m}$. Therefore, we have $vc(\mathcal{C}_{n+m},\bar{\mathbf{d}}) \le vc(\mathcal{C}_n,\bar{\mathbf{d}}_1) + vc(\mathcal{C}_m,\bar{\mathbf{d}}_2) + 2c$.

So,

$$\frac{(p-T)\,(vc(\mathcal{C}_n,\bar{\mathbf{d}}_1)+vc(\mathcal{C}_m,\bar{\mathbf{d}}_2)) - 2nB_1 - 2mB_2}{\sum_j p\bar{d}_j} > \frac{(p-T)\,(vc(\mathcal{C}_{n+m},\bar{\mathbf{d}})) - (2n-1)B_1 - (2m-1)B_2 - 2B_3}{\sum_j p\bar{d}_j}$$

The left side of the above equation is the relative profit of $[\mathcal{C}_n,\mathcal{C}_m]$ and the right side of the equation is at least the relative profit of $\mathcal{C}_{n+m}$. The above inequality holds because $2B_3 > B_1 + B_2 + 2(p-T)c$. This completes the proof.

□

4.2. Domination of a Long Chain

In this section, we show that in designs with symmetric parameters, a long chain cannot be dominated by other designs with $2n$ links in all demand scenarios.

Lemma 2. *Suppose we have n plants and products. The following is the demand uncertainty set:*

$$\mathcal{U}_n = \left\{ \mathbf{d} \;|d_j = c + \hat{d}_j z_j \;\; \forall j \in \{1,\cdots,n\},\;\; \|\mathbf{z}\|_1 \le n,\;\; \|\mathbf{z}\|_\infty \le 1 \right\}. \tag{19}$$

We have all capacities equal to c, with $\hat{d}_j \ge \gamma$ where $c \ge \gamma > \frac{c}{2}$ and $n \ge 4$. Suppose $p_j = p$, $T_{ij} = T$, and $B_{ij} = B$. Then among all designs with $2n$ links, a long chain cannot be dominated by another design that is not a long chain in all demand scenarios.

Proof of Lemma 2 For a given design, the relative profit for a demand $\mathbf{d}$ can be evaluated as follows: suppose $\mathbf{x}$ is a $0,1$ vector associated with that particular design. Then the relative profit can be calculated as:

$$\max_{\mathbf{y}(\mathbf{d})} \min_{\mathbf{d}\in\mathcal{U}} \frac{\sum_{i,j} (p-T)\, y_{ij}(\mathbf{d}) - \sum_{i,j} B x_{ij}}{\sum_j p d_j}. \tag{20}$$

$$0 \le y_{ij}(\mathbf{d}) \le d_j x_{ij}. \tag{21}$$

We prove by contradiction that any design with $2n$ links that is not a long chain can be dominated by a long chain in at least one demand scenario.

(i) Suppose in design G_0 one product has degree one. Without loss of generality, we assume the product n has degree one.

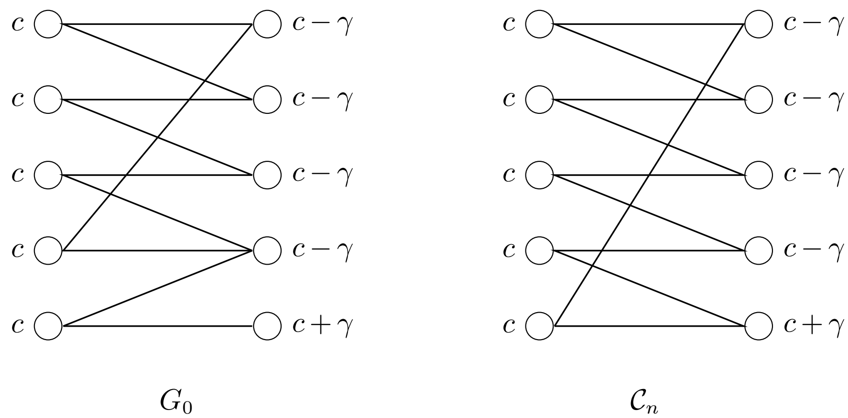

Figure (a)

We claim that a long chain design $\mathcal{C}_n$ in which the degrees of all its products are 2 will dominate G_0 on the demand scenario $(c-\gamma,\ldots,c-\gamma,c+\gamma)$. This is due to the fact that in $\mathcal{C}_n$, we can fulfill this demand vector but we cannot fulfill it in G_0.

(ii) From part (i), we now only consider the designs in which the degrees of all the products are 2. Consider the design G_1 that satisfies this condition and has at least one plant with degree 1. Without loss of generality, suppose the first plant has degree 1 and this plant is connected to product 1. Because the number of arcs is $2n$, there must be a plant s that has a degree more than $r>2$, and is connected to $r\geq 3$ products. Consider the demand vector $(c-\gamma,c+\gamma,c+\gamma,\cdots,c+\gamma)$:

There exists a plant s with degree more than 2 connected to $r\geq 3$ products. The degree of each product is 2, so these r products are at most connected to $r+1$ plants. Consider the optimal capacity allocation for this demand and design. It is not hard to see that, for at least one of the r products (excluding the first product - say the product r_1), at most half of the capacity of plant s is allocated to the product r_1 and the demand of this product r_1 is not fully satisfied. This is due to the fact that all of these r products with demands $c+\gamma>\dfrac{3c}{2}$ are only connected to the plant s and one other plant with capacity c. We then remove the link between product r_1 and plant s, and add a link between product r_1 and plant 1. Plant 1 is only connected to products 1 and r_1, and the demand of product 1 is less than $\frac{c}{2}$. Therefore, in the obtained design, plant 1 can allocate more than $\frac{c}{2}$ capacity to product r_1. Therefore, the total demand that we can satisfy increases.

It is easy to see that in the above procedure, we always decrease the number of plants with degree 1, with the degrees of the products remaining the same. So, any design with at least one degree 1 plant and degree 2 products is dominated by a design in which the degrees of all plants and products are 2 in the demand scenario $(c-\gamma, c+\gamma, c+\gamma, \cdots, c+\gamma)$. It is easy to see that the maximum capacity allocation of all designs with degree 2 plants and degree 2 products are at most the maximum capacity allocation of a long chain for a particular demand $(c-\gamma, c+\gamma, c+\gamma, \cdots, c+\gamma)$. The maximum capacity allocation is equal to the cost of the minimum vertex cover, which for all of these designs is at most nc. The cost of the minimum vertex cover for a long chain is nc for this demand scenario as well. Therefore a long chain design $\mathcal{C}_n$ will have a larger relative profit and will dominate G_1 for the demand $(c-\gamma, c+\gamma, c+\gamma, \cdots, c+\gamma)$.

(iii) The only designs we need to consider are combinations of chains in which the degrees of the plants and products are 2. Consider two short chains in which the first short chain has r plants, and the demand vector $(d_1, \ldots, d_r, 0, \ldots, 0)$. The long chain will dominate the short chains in this demand scenario where $d_i > c$ for $1 \le i \le r$. This happens because in two short chain designs the first r product demands are satisfied by the first r plant capacities, and in a long chain the first r product demands can be satisfied by $r+1$ plant capacities. Therefore a long chain will have a larger relative profit.

□

5. Insights from computations

In this section, we detail the performance of the designs produced by model (6). Before doing so, we first establish that model (6) is *tractable*, which we define as allowing us to quickly solve instances of practical size on commodity hardware (Section 5.1). In particular, we show that we can obtain near-provably-optimal solutions for instances with 15 plants and 15 products in about an hour on a laptop computer.

We derive insight into the solutions produced by our model by considering the case of homogeneous parameters. We find the results match our intuition and observations made previously in the literature (Section 5.2).

We compare the initial designs produced by model (6) with their Pareto-efficient counterparts, and show that Pareto-efficient designs usually improve over the initial design for little extra effort (Section 5.3). Finally we investigate how our designs perform when the true distribution of the demand is substantially different from the distribution we had estimated (Section 5.4). We compare different Pareto efficient designs for various values of Γ in (7), as well as with designs produced by a stochastic programming model similar to that presented in Mak and Shen (2009). We find that, despite optimizing for a worst-case objective, we improve slightly over the stochastic programming designs when there is no difference between the estimated and true distribution and improve substantially when there is. Finally, we demonstrate in Section 5.5 that our overall approach produces solutions that substantially reduce the price of flexibility compared to standard designs like the long chain design, while maintaining the same level of revenue.

Throughout this section we will primarily use two summary statistics to evaluate simulation performance: the arithmetic mean of the measured quantity, and the *Conditional Value at Risk* (Bertsimas and Brown 2009) at the 10% level ("$CVaR_{10\%}$"), which we take as the arithmetic mean of our metric over the worst 10% of scenarios. The use of $CVaR$ allows us to capture worst-case performance without being unduly affected by the most extreme cases.

5.1. Tractability of adaptive formulation

As described in Section 2, model (6) is an ARO problem with binary here-and-now variables and continuous wait-and-see variables. As we are using a polyhedral "budget" uncertainty set and affine adaptability, we can reformulate this problem to obtain a robust mixed-integer linear optimization problem that can be solved with any general-purpose integer optimization solver. Understanding the tractability of our approach requires a brief description of the *branch-and-bound* method these solvers employ. In branch-and-bound, a continuous relaxation of the problem is solved, giving an upper bound z_{bound} on the problem (as we are maximizing). This bound is tightened progressively by creating subproblems by branching on fractional variables. The best integer solution found so far is stored (with value $z_{integer}$ providing a lower bound), and the usual termination criterion, other than a time limit, is to stop this branching process when the *bound gap*, defined as

$$\frac{z_{bound} - z_{integer}}{z_{integer}}, \tag{22}$$

falls below a tolerance τ. The rate at which the two bounds change depends on the structure of the problem. For example, a good integer solution may be found quickly but it may take a very long time to prove optimality of the solution. In Figure 2, we visualize the progress of these bounds for an instance of the flexibility model: we see that a "good" integer solution is found within 1 second, and a near-optimal solution is found within 6 seconds, but the solver takes approximately 8 seconds to prove that it is within 1% of optimal by successive tightening of the upper bound. We also note that the initial "good" solution was actually within 3% of optimal, although given the bound known at the time we could only guarantee that it was within 13% of optimal.

While proving a design is optimal may be desirable after a model is finalized, of more immediate interest is the time to find "good" solutions. We thus consider the *true gap*

$$\frac{z^*_{integer} - z_{integer}}{z_{integer}}, \tag{23}$$

where $z^*_{integer}$ is the objective function value of a design that is within 1% of the optimal solution. As our goal is to find such a solution, we would like to compare the time taken to find it on average with the time it takes us to confirm it. In Figure 3, we plot the percentage of instances achieving a true gap of less than 1%, 2%, 5% as a function of time for instances of size $n = m = 7$, and find that 75% of instances have a true gap from optimal of less than 5% within 5 seconds, but only approximately 60% of instances have a "proved" gap less than 1% after 40 seconds. Based on this insight, in Figure 4 we show that we can find solutions, for instances of size $n = m = 15$, with a proved gap of less than 5% in approximately 2 hours, and find solutions with a true gap within 2% in approximately 1 hour.

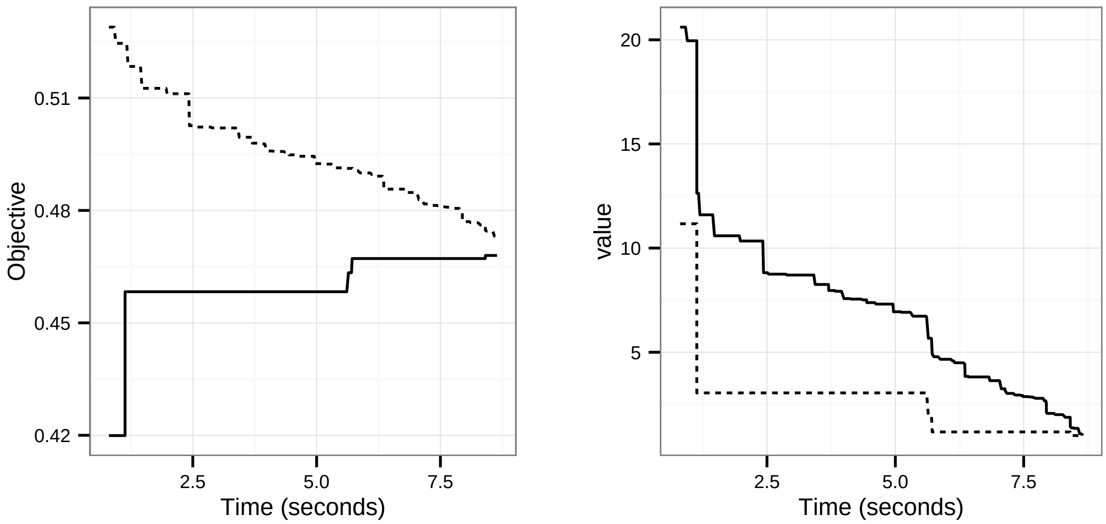

Figure 2 Visualization of the progression of a mixed-integer optimization solver for a random instance. Left: the objective value of the solver's best integer solution so far (solid) and the bound (dashed). Right: the bound gap reported by the solver (solid) and the bound gap between between the best integer solution so far and the bound when solver stops (dashed).

All experiments were conducted on a laptop computer with an Intel Core i7 2.6GHz CPU. All models were implemented using the JuMP (Lubin and Dunning 2015) and JuMPeR (Dunning 2015) packages in the Julia programming language, with Gurobi 6.0.4 used as the solver (Gurobi Optimization 2015).

5.2. Impact of parameter selections

We considered two flexibility designs—a "dedicated" design $\mathcal{C}_D$ and a 'long chain" design $\mathcal{C}_L$—under these restricted, homogeneous settings commonly discussed in the literature and analyzed how the model values different designs under different parameter selections.

$$\mathcal{C}_D = \{(1,1),(2,2),\ldots,(n,n)\}.$$

$$\mathcal{C}_L = \{(1,1),(1,2),(2,2),(2,3),\ldots,(n,n),(n,1)\}.$$

Flexibility cost versus Γ We first consider the relationship between the cost of adding flexibility, B, with the uncertainty set parameter Γ. To do so, we fix $\bar{p}=1$, $\bar{\mu}=100$, and $\bar{\sigma}=50$.

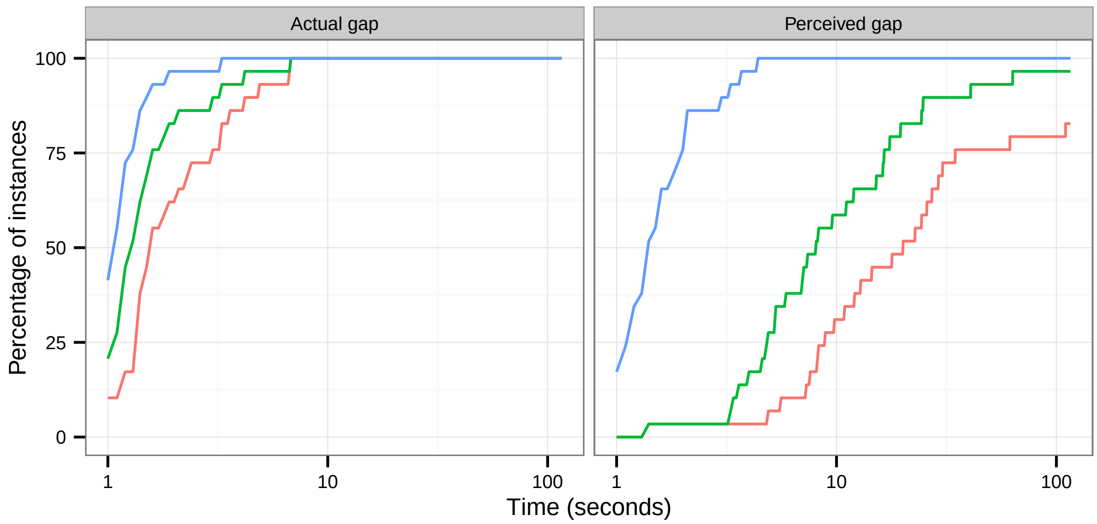

Figure 3 Percentage of a set of 50 random instances of size $n = m = 7$ that have a gap lower than 5%, 2%, or 1% (lines, top to bottom) at a given time. Left, the "actual gap": the gap between the best integer solution so far and the final integer solution. Right, the "perceived gap": the gap between the best integer solution so far and the bound at that point in time.

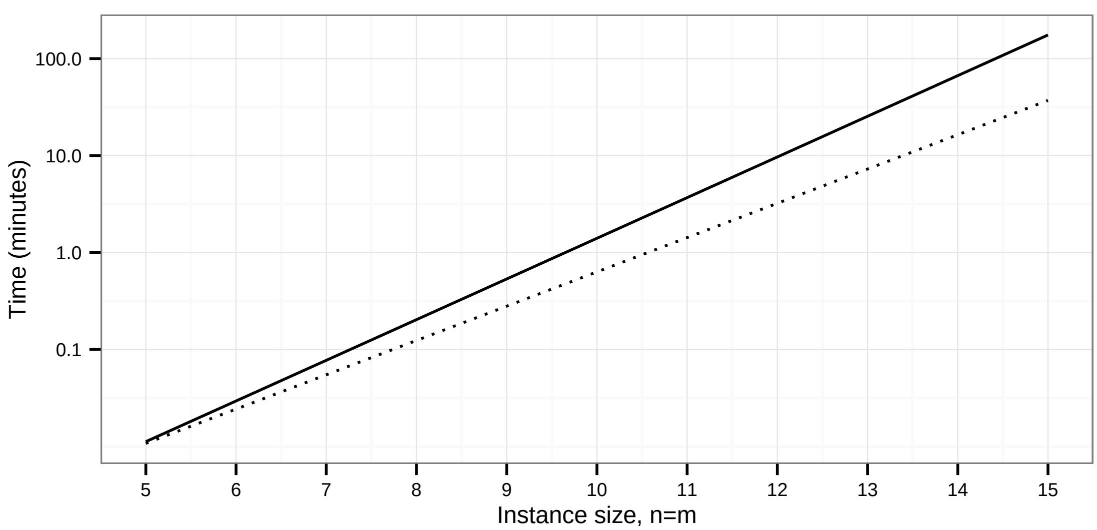

Figure 4 Time for 90% of instances to reach a proved bound gap of 5% (solid line), and time for 90% of instances to find a solution within 2% of the eventual integer solution (dashed line).

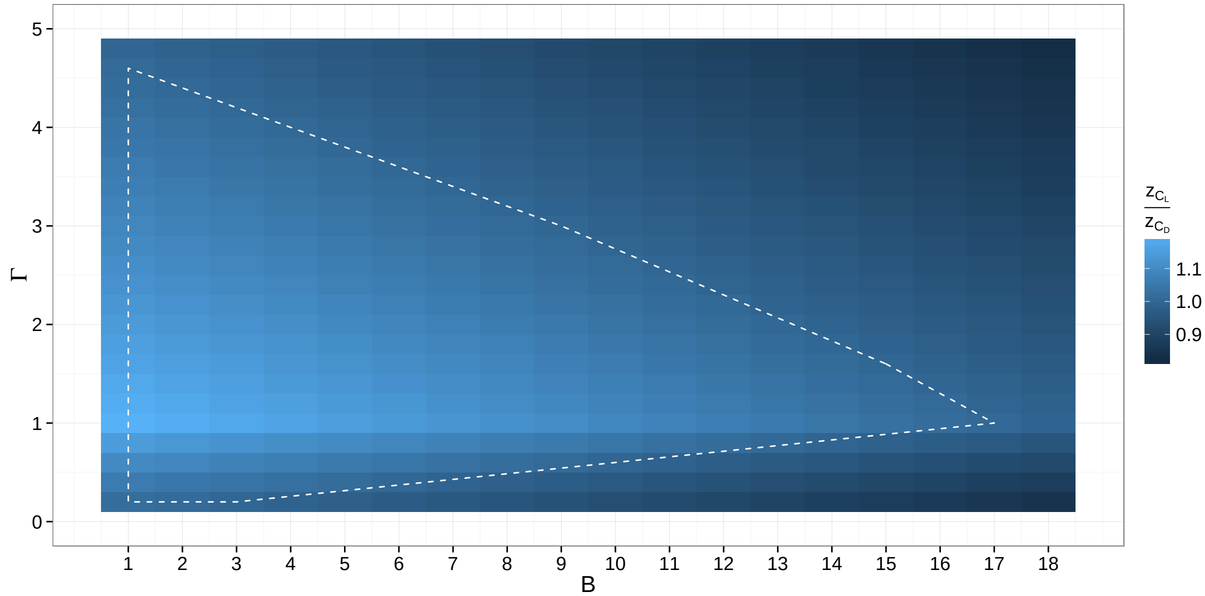

Figure 5 Ratio of the objective function value for a "long chain" design versus a "dedicated" design for varying flexibility costs B and uncertainty set budget Γ. The dotted line marks the parameter range in which the long chain design is better than the dedicated design.

We then vary $B \in \{1, \ldots, 18\}$ and $\Gamma \in \{0, 0.2, \ldots, n\}$, and record the objective function values for the long chain design and dedicated designs. In Figure 5, we plot the ratio of the objective values ($z_{\mathcal{C}_D}/z_{\mathcal{C}_L}$) with respect to the two varying parameters B and Γ. We first note that the maximum flexibility cost for which the long chain is better than the dedicated is approximately 17. The maximum revenue that could be obtained per product in this setting is 50 ($\bar{p}\bar{\sigma}$), which matches the intuition that the cost of additional flexibility should be less than the extra profits we could possibly hope to capture by having the flexibility. As Γ varies between 0 and 1, the range of B for which the long chain solution is better increases, before decreasing again as Γ ranges from 1 to 5. This again matches intuition; if the variability is high, then there is little more to gain from the long chain as we will either exhaust our capacity easily, or we will be able to satisfy each product from one plant, leaving less room for the long chain design to be beneficial.

Flexibility cost versus correlation We also consider how correlation between products affects the design choice. Suppose the mean and standard deviation of demand for each product

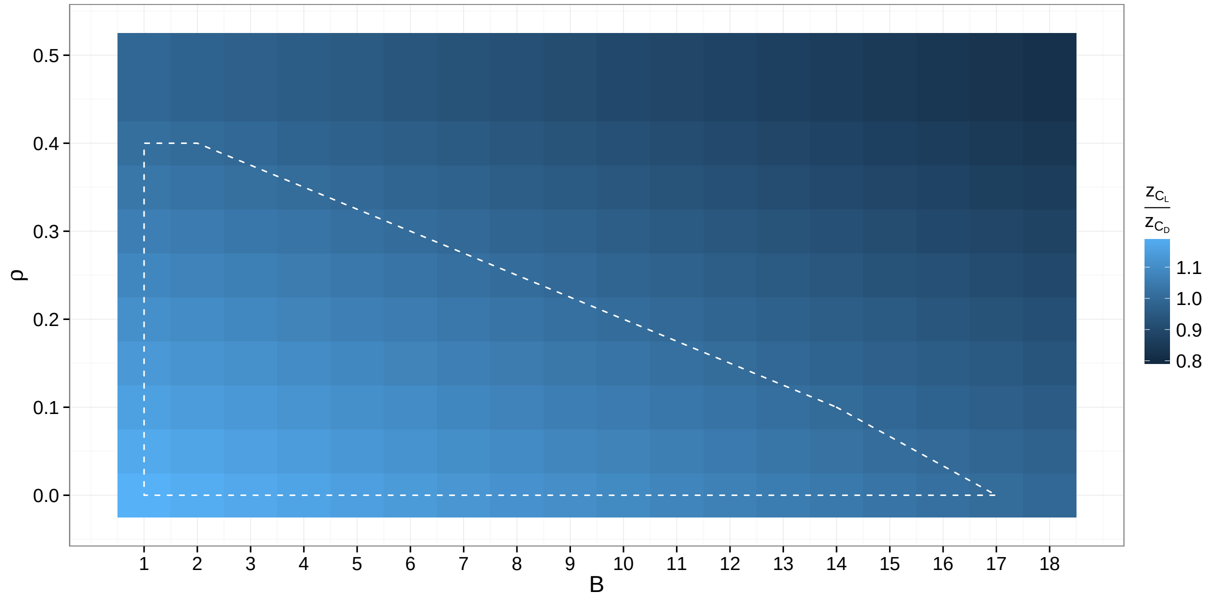

Figure 6 Ratio of the objective function value for a "long chain" design versus a "dedicated" design for varying flexibility costs B and pair-wise demand correlation ρ. The dotted line marks the parameter range in which the long chain design is better than the dedicated design.

is again $\bar{\mu}$ and $\bar{\sigma}$, respectively, but the demand for each product has a pairwise correlation of ρ with the other products. The covariance matrix Σ thus has diagonal elements $\Sigma_{i,i} = \sigma^2$, with off-diagonals of $\rho\sigma^2$. We again fix $\bar{p} = 1$, $\bar{\mu} = 100$, and $\bar{\sigma} = 50$. We then vary $B \in \{1, \ldots, 18\}$ and $\rho \in \{0, 0.05, \ldots, 0.5\}$, with Γ fixed to 1.

In Figure 6, we plot the same quantity as before. In the case where $\rho = 0$ (no correlation) we return to the same setting as above. However, for any fixed B we note that as ρ increases, the long chain design becomes less attractive. For values of ρ above approximately 0.4, we always prefer the dedicated design over the long chain design. This matches intuition, as at least in the limiting case where $\rho = 1$, we would expect to gain no benefit from the long chain design as there would be no variability in demand between products. The overall benefits of flexibility are reduced as product demand correlation increases.

5.3. Utility of Pareto efficient designs

We presented the notion of Pareto efficient designs in Section 3. In brief, we are interested in finding designs that are not only robust with respect to the objective function, but also able to perform well in non-worst-case demand scenarios. To find these designs we first solve model (6), and then seek to find a (possibly) different design with the same worst-case objective value but with non-dominated performance for another scenario (like the nominal or mean scenario). We also presented a relaxation of the worst-case criterion, in which the "relaxed Pareto" design only needs to achieve $\alpha \in (0,1]$ of the worst-case objective value of the initial design. In other words, we exchange worst-case performance for possible improvements in the alternative scenario. In this section, we demonstrate how often we obtain a Pareto or relaxed-Pareto efficient design that improves over the initial (possibly Pareto) design, and to what degree these alternative designs improve over the initial designs.

We considered 100 random instances drawn from the following family of instances: $n = m = 5$, $B_{ij} \sim Uniform(50,350)$, $B_{ii} = 0$, $p \sim Uniform(10,20)$, $T_{ij} = 0$, $\mu_j \sim Uniform(150,250)$, $\sigma_j = \frac{1}{2}\mu_j$, and $c_i = \mu_i$. For each instance we generated 100 demand scenarios, where the demand d_j for product j is drawn from $\max\{Normal(\mu_j, \sigma_j), 0\}$. For each instance and demand scenario a deterministic problem was solved to find the best-possible profit that could be obtained if we had clairvoyant knowledge of the demand. This naturally represents an upper bound on the profit that can be obtained by any design for a given demand scenario.

We considered two values of α: the relaxed-Pareto parameter 0.8, and the standard Pareto parameter 1, i.e., not relaxed at all. The scenario used to find a Pareto efficient design was the mean demand scenario $d_j = \mu_j$. We used a budget uncertainty set (7) with either $\Gamma = 1$ or $\Gamma = 3$. For each of the instances and demand scenarios we thus have six realized profits: one for each of two Pareto efficient designs (for $\alpha = 0.8$ or 1) and the initial (possibly non-Pareto) design (solution to (6)), and for the two values of Γ. These profits were all normalized by the best possible profit obtained from solving the deterministic problem described above.

		Mean			$CVaR_{10\%}$		
Γ	α	Worse	Same	Better	Worse	Same	Better
1	0.8	10	17	73	14	17	69
1	1.0	4	45	51	10	45	45
3	0.8	30	18	52	32	18	50
3	1.0	32	34	34	28	34	38

Table 1 Results for evaluation of Pareto efficient design (Section 5.3). For each of the 100 random instances considered, the Pareto ($\alpha = 1$) and relaxed-Pareto ($\alpha = 0.8$) efficient designs are either worse, the same, or better than the initial designs in simulation under a mean or CVaR metric of performance. The relaxed-Pareto efficient design is different from the initial design more frequently, and is better than the initial design in over half of the instances by either metric.

The results are depicted in Figure 7. The mean and $CVaR$ of the normalized profits across the demand scenarios for each instance are sorted from lowest to highest independently for each design. The left figure, showing the mean performance, reveals that there is little difference for the relaxed-Pareto and Pareto efficient designs in the $\Gamma = 3$ case, but that both have marginally better profitability than the initial design. For $\Gamma = 1$ we see more substantial differentiation between the designs, with the relaxed-Pareto design performing slightly better than the Pareto design. The figures for the $CVaR$ metric are similar in appearance, although the absolute magnitudes of the differences are more substantial than for the mean. For both $\Gamma = 1$ and $\Gamma = 3$ the relaxed-Pareto design outperforms both alternatives.

Figure 7 summarizes the overall performance of the designs, but does not capture on a per-instance basis how Pareto efficient designs compare to the initial designs. To address this, we constructed Table 1 which describes, for each of the 100 instances, whether the mean and CVaR of the Pareto efficient designs were the same, worse, or better. We note that the Pareto efficient designs are at least as good as the initial designs in approximately 70% to 95% of instances, with Pareto efficient designs being strictly better in approximately 35% to 75% of instances. The relaxed-Pareto efficient designs are, as expected, different from the initial designs for more instances, and in general they are better than the initial designs more often than the Pareto efficient designs.

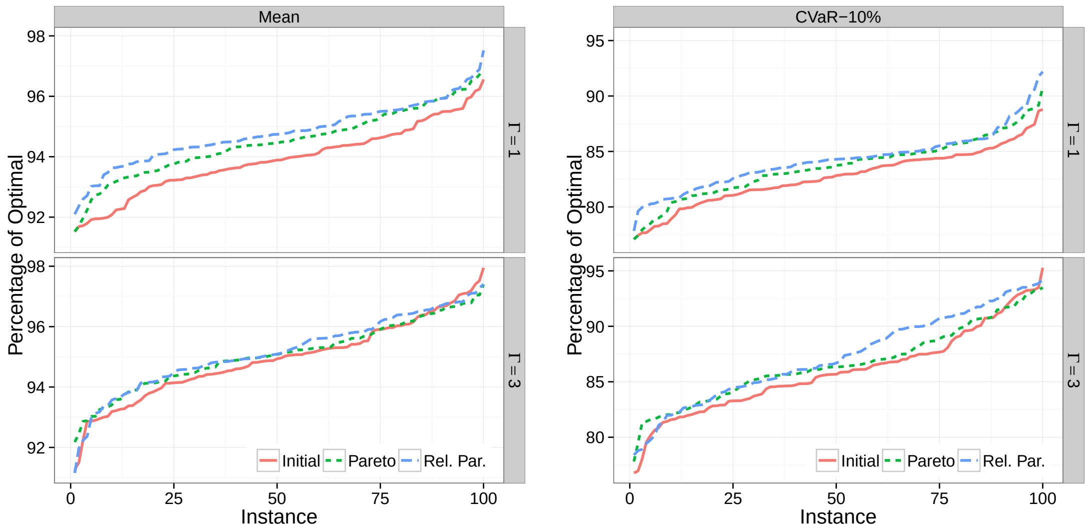

Figure 7 Results for evaluation of Pareto efficient designs (Section 5.3). Each plot displays the mean (left) or $CVaR_{10\%}$ (right) of simulated normalized profits of the relaxed-Pareto efficient design (purple), Pareto efficient design (red), and initial design (teal) for each random instance, ordered from worst to best independently for each design type. The top figures when the uncertainty set parameter Γ is set to 1, and bottom are for $\Gamma = 3$.

5.4. Quality of designs

To quantify the degree to which our designs performed in a realistic setting with both uncertainty in the data and variability in the realized demand, we considered families of random instances where the "true" demand distribution differs from perceived demand. We compare our designs with the designs obtained from the stochastic programming approach in Mak and Shen (2009). In particular, we implemented a sample-average approximation (SAA) version of the model presented in Mak and Shen (2009). The method describes a relaxation of that approach, so the results we present should be considered an upper bound on the quality of their solutions that are obtained at the cost of tractability.

We considered 100 instances drawn from the same family of instances described in Section 5.3. To investigate the effect of uncertainty in the data we considered three levels of perturbation of the distribution from which demand scenarios are drawn for simulation, $\eta \in \{0.0, 0.5, 0.9\}$. Given η, the mean demand $\hat{\mu}_j$ for product j is drawn from $Uniform\left((1-\eta)\,\mu_j, (1+\eta)\,\mu_j\right)$, where μ_j is the mean that is provided to the optimization model. We then generated 100 random demand scenarios for each instance and level of perturbation: that is, demand d_j is drawn from $\max\{Normal(\hat{\mu}_j, \sigma_j), 0\}$.

For each instance we calculated the relaxed-Pareto efficient design ($\alpha = 0.8$) with $\Gamma \in \{0,1,2,3,4,5\}$, where $\Gamma = 0$ corresponds to the deterministic problem and $\Gamma = 5$ is equivalent to a "box" uncertainty set. We also calculated the SAA design with an expectation objective. As in Section 5.3, we considered the mean and $CVaR$ of profit normalized by the clairvoyant-optimal profit.

We first investigated the effect of Γ on design quality, and the results are displayed in Figure 8. To summarize, we see that by tuning Γ for an "intermediate" level of variability we can obtain good average and worst-case behaviour. In particular, we see that both $\Gamma = 0$ and $\Gamma = 5$ perform poorly in comparison to $\Gamma = 1$ and $\Gamma = 3$ at all perturbation levels by mean performance. However, the $\Gamma = 3$ design distinguishes itself from the $\Gamma = 1$ solution for $CVaR$ performance even without disruption, and performs better than all alternatives.

We then compared the $\Gamma = 3$ designs with the SAA designs in Figure 9. Under the no-perturbation setting ($\eta = 0$) there is little difference in mean performance between the two, and similarly for $CVaR$. However, as perturbation increases we see that our designs perform better more often than the SAA designs, with the largest differences occurring at the highest level of disruption.

5.5. Price of flexibility

We established that model (6) is tractable (Section 5.1) and produces intuitive designs (Section 5.2), that Pareto efficiency adds value (Section 5.3), and that the designs are robust against uncertainty (Section 5.4). We now seek to quantify how much our method reduces the price of flexibility.

We will compare the design selected by the relaxed-Pareto model (with parameter $\alpha = 0.8$) with three alternatives: no additional flexibility ("dedicated"), the classic "long chain" design, and a "three chain" design in which each plant produces three products, and vice versa. We evaluate them on 100 instances drawn from the same family of instances described in Section 5.3, and simulate random demand scenarios as before.

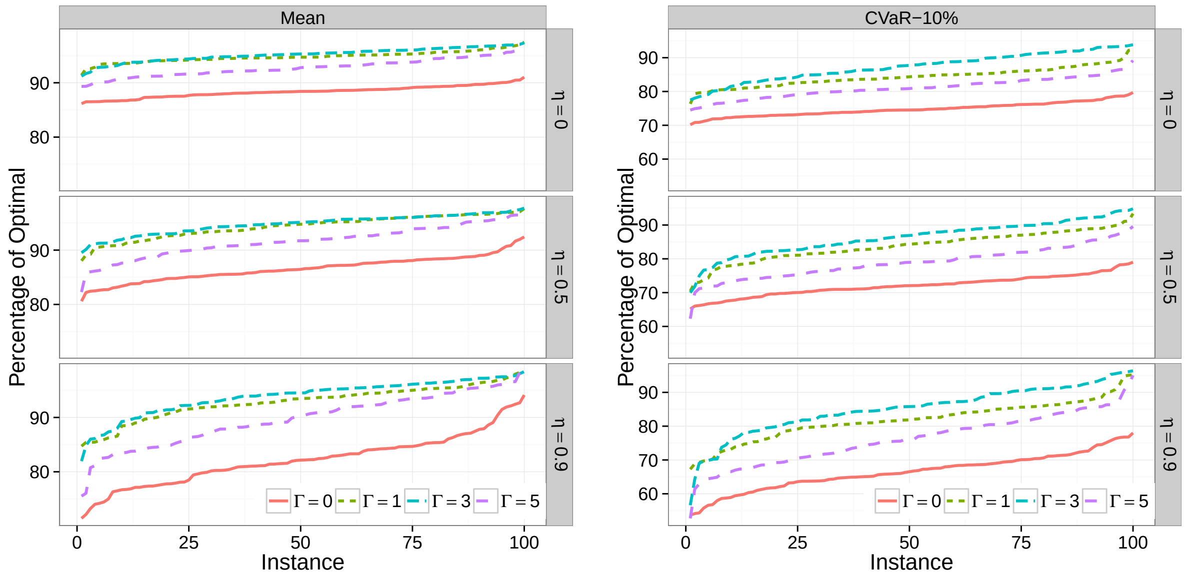

Figure 8 Results for evaluation of effect of Γ for different levels of perturbation η (Section 5.4). Each plot displays the mean (left) or $CVaR_{10\%}$ (right) of simulated normalized profits for each random instance, ordered from worst to best independently for each design type. Each line represents the designs corresponding to different values of Γ, with the perturbation increasing from top to bottom.

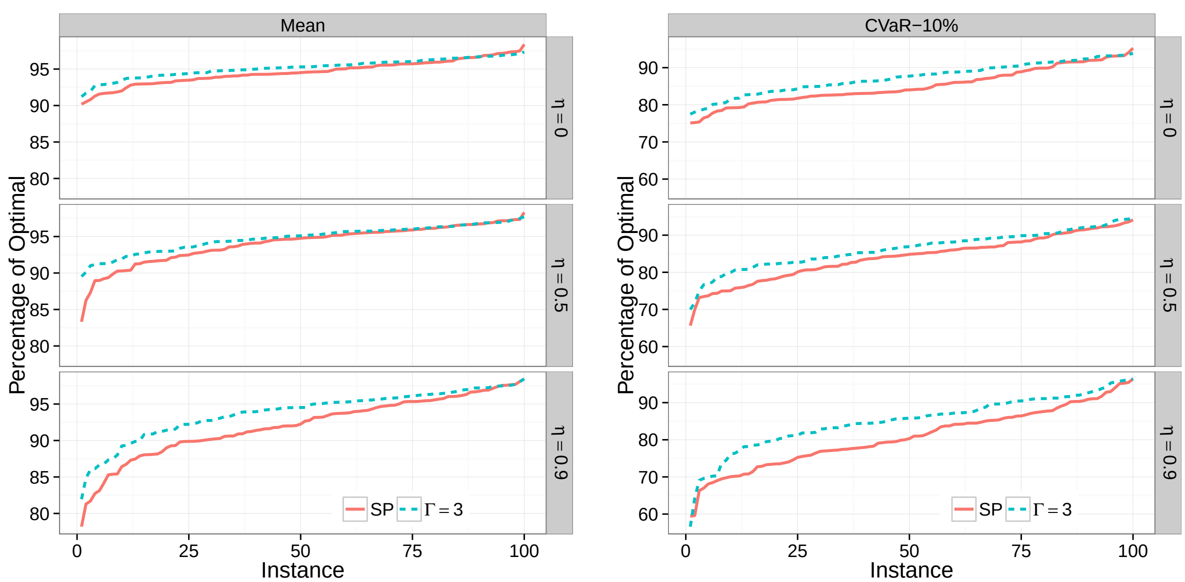

Figure 9 Results for the comparison of our designs with SAA designs for different levels of perturbation η (Section 5.4). Each plot displays the mean (left) or $CVaR_{10\%}$ (right) of simulated normalized profits for each random instance, ordered from worst to best independently for each design type. The top line represents the designs corresponding to either our $\Gamma = 3$, or the SAA stochastic programming design.

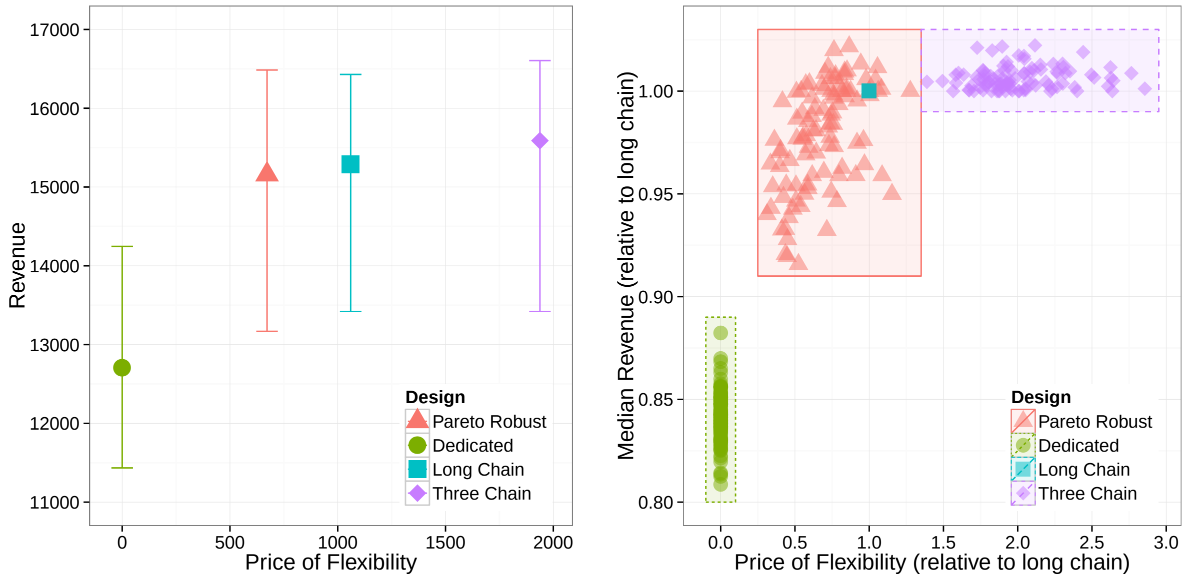

Figure 10 Left: comparison of the "price of flexibility" for four different flexibility designs for a representative random instance. The price is the cost of adding flexibility over the "dedicated" design, and revenue is the distribution (first quartile, median, and third quartile) of revenues over 100 demand scenarios. Right: median revenues and prices for all 100 random instances, normalized by the long chain design's median revenue and price.

We are interested in two key metrics for each design: the "price of flexibility" (equal to the cost of the flexibility added $\sum_{ij} B_{ij}x_{ij}$), and total revenue. In the left plot of Figure 10 we show these metrics for one of the 100 instances. Note that the price is fixed for each design, but revenue has a distribution. We can see that in this particular instance, the distribution of revenues for the robust design and the two chain designs have a roughly similar distribution. However, the additional costs of the long chain design and the three chain design add only slightly improve revenues obtained, with the robust design being the most profitable.

In the right plot of Figure 10 we summarize the median revenue and price of flexibility for all 100 instances, normalized by the corresponding revenue and price for the long chain design. We see that the dedicated design achieves approximately 85% of the revenue of the long chain design, and is the cheapest option (by definition). The robust design achieves approximately 98% of the revenue on average of the long chain design, but at only 70% of the cost. The "three chain" design is roughly double the cost of the long chain design, but improves revenues by little more than 1% point. This demonstrates that our method can effectively reduce the price of flexibility without impacting revenues, and therefore drive higher profitability.

6. Concluding remarks

In this paper, we have demonstrated that we can reduce the price of flexibility by formulating the process flexibility design problem as an adaptive robust optimization problem. In particular, in simulation experiments our designs often obtained more than 90% of the maximum clairvoyant possible profit. They proved to be robust against considerable uncertainty in the distribution of demand, in terms of both average profit and worst-case profit. Most importantly, we showed that our designs reduce the price of flexibility by 30% versus long chain designs. This was achieved with negligible impact on revenues, and thus profitability was substantially higher for our designs.

Our approach improves over more ad-hoc approaches and heuristics as it explicitly characterizes the trade-off between flexibility and profits (or, equivalently, costs). The use of Pareto-efficient and relaxed Pareto-efficient solutions was a key part of our strategy for producing solutions that perform well in both "worst-case" and "average-case" scenarios.

Our models represent a general tool for obtaining profitable manufacturing process flexibility designs. As we utilize established techniques such as mixed-integer optimization and robust optimization, our model can be extended to handle a variety of firm-specific constraints and additions. It thus represents a general way to reduce the price of flexibility and improve profitability in settings encountered in practice by manufacturing firms.

Acknowledgments

References

Ball, Michael O., Maurice Queyranne. 2009. Toward robust revenue management: Competitive analysis of online booking. *Operations Research* **57**(4) 950–963.

Ben-Tal, Aharon, Boaz Golany, Arkadi Nemirovski, Jean-Philippe Vial. 2005. Retailer-supplier flexible commitments contracts: a robust optimization approach. *Manufacturing & Service Operations Management* **7**(3) 248–271.

Ben-Tal, Aharon, Alexander Goryashko, Elana Guslitzer, Arkadi Nemirovski. 2004. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* **99**(2) 351–376.

Ben-Tal, Aharon, Arkadi Nemirovski. 1999. Robust solutions of uncertain linear programs. *Operations Research Letters* **25**(1) 1–13.

Bertsimas, Dimitris, Hoda Bidkhori. 2014. On the performance of affine policies for two-stage adaptive optimization: a geometric perspective. *Mathematical Programming* 1–18.

Bertsimas, Dimitris, David Brown. 2009. Constructing uncertainty sets for robust linear optimization. *Operations Research* **57**(6) 1483–1495.

Bertsimas, Dimitris, David Brown, Constantine Caramanis. 2011. Theory and applications of robust optimization. *SIAM Review* **53**(3) 464–501.

Bertsimas, Dimitris, Vineet Goyal. 2012. On the power and limitations of affine policies in two-stage adaptive optimization. *Mathematical Programming* **134**(2) 491–531.

Bertsimas, Dimitris, Vishal Gupta, Nathan Kallus. 2013. Data-driven robust optimization. *arXiv preprint arXiv:1401.0212* .

Bertsimas, Dimitris, Dan Iancu, Pablo Parrilo. 2010. Optimality of affine policies in multistage robust optimization. *Mathematics of Operations Research* **35**(2) 363–394.

Bertsimas, Dimitris, Melvyn Sim. 2004. The price of robustness. *Operations Research* **52**(1) 35–53.

Biller, Stephan, Ana Muriel, Yongmei Zhang. 2006. Impact of price postponement on capacity and flexibility investment decisions. *Production and Operations Management* **15**(2) 198–214.

Chen, Xi, Jiawei Zhang, Yuan Zhou. 2014. Optimal sparse designs for process flexibility via probabilistic expanders. *Available at SSRN 2400768* .

Chen, Xin, Melvyn Sim, Peng Sun. 2007. A robust optimization perspective on stochastic programming. *Operations Research* **55**(6) 1058–1071.

Chou, Mabel C., Geoffrey A. Chua, Chung-Piaw Teo, Huan Zheng. 2010. Design for process flexibility: Efficiency of the long chain and sparse structure. *Operations Research* **58**(1) 43–58.

Chou, Mabel C., Geoffrey A. Chua, Chung-Piaw Teo, Huan Zheng. 2011. Process flexibility revisited: the graph expander and its applications. *Operations Research* **59**(5) 1090–1105.

Deng, Tianhu, Zuo-Jun Max Shen. 2013. Process flexibility design in unbalanced networks. *Manufacturing & Service Operations Management* **15**(1) 24–32.

Dunning, Iain. 2015. JuMPeR: JuMP extension for robust optimization. URL `https://github.com/IainNZ/JuMPeR.jl`.

Gurobi Optimization, Inc. 2015. Gurobi optimizer reference manual. URL `http://www.gurobi.com`.

Iancu, Dan A, Nikolaos Trichakis. 2013. Pareto efficiency in robust optimization. *Management Science* **60**(1) 130–147.

Jordan, William C, Stephen C Graves. 1995. Principles on the benefits of manufacturing process flexibility. *Management Science* **41**(4) 577–594.

Lubin, Miles, Iain Dunning. 2015. Computing in operations research using julia. *INFORMS Journal on Computing* **27**(2) 238–248.

Mak, Ho-Yin, Zuo-Jun Max Shen. 2009. Stochastic programming approach to process flexibility design. *Flexible Services and Manufacturing Journal* **21**(3-4) 75–91.

Perakis, Georgia, Guillaume Roels. 2010. Robust controls for network revenue management. *Manufacturing & Service Operations Management* **12**(1) 56–76.

Simchi-Levi, David, Yehua Wei. 2012. Understanding the performance of the long chain and sparse designs in process flexibility. *Operations Research* **60**(5) 1125–1141.

Simchi-Levi, David, Yehua Wei. 2015. Worst-case analysis of process flexibility designs. *Operations Research* **63**(1) 166–185.

Wang, Xuan, Jiawei Zhang. 2015. Process flexibility: A distribution-free bound on the performance of k-chain. *Operations Research* **63**(3) 555–571.

Wei, Yehua, David Simchi-Levi. 2015. Worst-case analysis of process flexibility designs. *Operations Research* **63**(1) 166–185.