Improving the Tractability of SVC-based Robust Optimization

Benoît Loger, Alexandre Dolgui, Fabien Lehuédé, Guillaume Massonnet

Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 55-10 (2022) 719–724

Improving the Tractability of SVC-based Robust Optimization

Benoît Loger * **Alexandre Dolgui** * **Fabien Lehuédé** * **Guillaume Massonnet** *

* *IMT Atlantique, LS2N, 4 rue Alfred Katler 44300 Nantes FRANCE (e-mail: {benoit.loger, alexandre.dolgui, fabien.lehuede, guillaume.massonnet}@imt-atlantique.fr).*

Abstract: Support Vector Clustering (SVC) has been proposed in the literature as a data-driven approach to build uncertainty sets in robust optimization. Unfortunately, the resulting SVC-based uncertainty sets induces a large number of additional variables and constraints in the robust counterpart of mathematical formulations. We propose two methods to approximate the resulting uncertainty sets and overcome these tractability issues. We evaluate these approaches on a production planning problem inspired from an industrial case study. The results obtained are compared with those of the SVC-based uncertainty set and the well known budget-based uncertainty set. We find that the approximated uncertainty set based formulation can be solved much faster than the SVC-based formulation. Still, the obtained solutions are comparable to the SVC-based solutions in term of performance.

Keywords: Data-driven, Robust Optimization, Production planning

1. INTRODUCTION

Robust optimization (RO) has become a popular approach to deal with uncertainties in optimization problems in the last decades and have been applied to a wide range of application fields (Sözüer and Thiele, 2016). The core concept of RO is to restrict the possible realizations of uncertain parameters to a given *uncertainty set* $\mathcal{U}$, then to optimize against the worst case within this set to obtain solutions that are immunized against all scenarios included in $\mathcal{U}$. With the growing complexity of supply chain and production systems, RO appears to be a promising approach to improve the performances and the reliability of industrial systems by reducing the impact of uncertain or unpredictable events. Since the pioneering work of Soyster (1973), a lot of efforts have been dedicated to propose and characterize different types of uncertainty sets in order to obtain tractable robust models (El Ghaoui et al., 1998; Ben-Tal and Nemirovski, 1998; Bertsimas and Sim, 2004).

Among the numerous applications of RO, Bertsimas and Thiele (2006) were the first to apply the model of Bertsimas and Sim (2004) to multi-period inventory management problems where demands are uncertain. On similar topics, José Alem and Morabito (2012) considered a production planning problem with uncertain costs and demands, while Wei et al. (2011) considered uncertain returns and demands. In Aouam and Brahimi (2013), the authors proposed a robust model for an integrated production planning problem. The RO approach have been extended to different source of uncertainty Varas et al. (2014), Thorsen and Yao (2017).

Thanks to the growing amount of data collected in industrial processes, the paradigm of *Data-driven robust optimization (DDRO)* have recently emerged with the objective to find new ways to use historical data as a support to build more sophisticated uncertainty sets. Several mathematical tools have been applied to incorporate a large fraction of the support of random parameters in the uncertainty set (Bertsimas and Brown, 2009; Bertsimas et al., 2018; Zhang et al., 2018). Chassein et al. (2019) have proposed a construction procedure of different class of uncertainty sets from real observations on a robust shortest path problem and compared their respective performance on a real-case application. Other approaches have addressed the exploitation of available data as an unsupervised learning problem, leading to the application of Machine Learning methods (Ning and You, 2019) to construct uncertainty sets. Among the different approaches proposed in the literature, Ning and You (2018) apply principal component analysis and kernel density estimation to construct uncertainty sets for different optimization problem encountered in the chemical industry. Multiple kernel support vector machine is applied in Han et al. (2021) for a multistage inventory management problem.

Among kernel learning approaches, Shang et al. (2017) proposed a support vector clustering (SVC) uncertainty set that efficiently captures correlations and asymmetries in the distribution of uncertain parameters. In addition, this technique produces polyhedral uncertainty sets, which offers a relative computational efficiency. This has been demonstrated by some recent studies such as a multi-product inventory management problem (Qiu et al., 2019), an energy system optimization (Shen et al., 2020) or a resource allocation in a cellular network (Wu et al., 2021). As another perspective on this technique, the recent work of Goerigk and Kurtz (2021) compare deep learning methods with the SVC-based method of Shang et al. (2017). As we shall see later, one drawback of this type of uncertainty

10.1016/j.ifacol.2022.09.492

set comes from the number of variables and constraints introduced in the robust model. Indeed, these are linearly increasing in both the number of uncertain parameters and the size of the data set. In many industrial applications, this results in a large mathematical formulation that quickly become intractable.

In this paper we address this scalability issue by introducing two simple methods to approximate the SVC-based uncertainty set described in Shang et al. (2017). Both approaches rely on a parameter that directly controls the number of additional variables and constraints in the robust model. This allows us to compute solutions that are identical or comparable to those obtained with the SVC-based uncertainty set with a significant reduction in the computational burden.

The remaining of this paper is organized as follows: Section 2 describes the uncertainty set of Shang et al. (2017) and exhibits the reasons of the scalability issue. Section 3 describes in details the approximation methods. Section 4 presents the application of our approximations method on a robust production planning problem, before Section 5 concludes this study.

2. SUPPORT VECTOR CLUSTERING BASED UNCERTAINTY SET

Throughout the theoretical construction of the approximate SVC-based uncertainty sets, we consider that our goal is to solve a robust combinatorial optimization problems of the form

$$\min \ \boldsymbol{c}^\intercal \boldsymbol{x} \tag{1}$$
$$\text{s.t. } \boldsymbol{a}_j^\intercal \boldsymbol{x} \leq b_j \ \ \forall \boldsymbol{a}_j \in \mathcal{U} \tag{2}$$
$$\boldsymbol{x} \in \mathbb{N}^m \tag{3}$$

where uncertain parameters $\boldsymbol{a}_j$ belongs to a given uncertainty set $\mathcal{U}_j$. Constraint (2) is equivalent to

$$\max_{\boldsymbol{a}_j \in \mathcal{U}} \boldsymbol{a}_j^\intercal \boldsymbol{x} \leq b_j \tag{4}$$

, where $\mathcal{U}$ is constructed from a sample of data that consists of a set $\mathcal{D} = \left\{\boldsymbol{u}^{(1)}, \ldots, \boldsymbol{u}^{(N)}\right\}$ of N points in a m-dimensional space. Ideally, $\mathcal{U}$ should provide a good representation of the support of the multi-dimensional random variable that generated $\mathcal{D}$. Various Machine Learning techniques have been developed to achieve this goal. SVC (Ben-Hur et al., 2001) is one of them that relies on a single parameter $\nu \in (0, 1)$ to define the boundaries of a cluster that contains at least $(1 - \nu)$ points of $\mathcal{D}$, while at most ν are considered as outliers (i.e. lie outside of the cluster boundaries). On of the major strengths of SVC is to naturally capture information such as the covariance of random variables or asymmetries in their distributions. In addition when the SVC algorithm is well designed, it does not need any sophisticated tuning of hyper-parameters and only requires to solve a quadratic program (QP).

Recently, Shang et al. (2017) have built data-driven uncertainty sets for robust optimization by applying the SVC approach with a custom piecewise linear kernel function (Weighted General Intersection Kernel, WGIK). The solution of their QP defines a subset $\mathcal{S}$ of outliers called *support vectors* (SV), among which a subset $\mathcal{B} \subseteq \mathcal{S}$, called *boundary support vectors* (BSV), are exactly on the boundary of the cluster. To summarize, those two sets

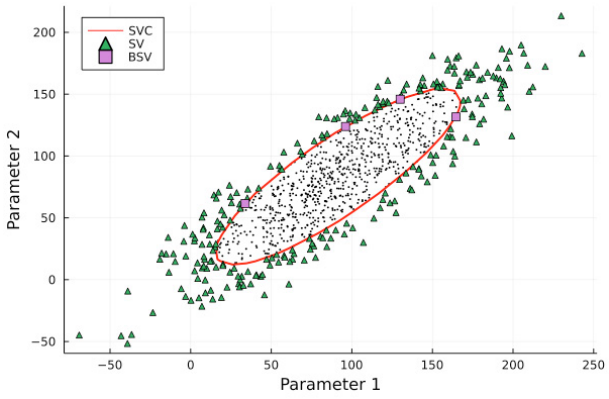

Fig. 1. Representation of $\mathcal{S}$ and $\mathcal{B}$ for a bivariate gaussian distribution when $\nu = 0.15$

are such that $\mathcal{B} \subseteq \mathcal{S} \subseteq \mathcal{D}$ and $|\mathcal{S} \setminus \mathcal{B}| \leq N\nu \leq |\mathcal{S}|$. We refer the reader to Shang et al. (2017) for more details on their method and the theoretical aspects of SVC. Figure 1 represents these sets for 1000 samples following a bivariate gaussian disribution when ν is set to 0.25.

To keep the remainder of this paper concise, we sometimes refer to a data point $\boldsymbol{u}^{(i)}$ by its index i when it is clear from the context. For each point $i \in \mathcal{D}$ the QP computes the vector of weights $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_N]$ where $\alpha_i = 0$ iff $i \in \mathcal{D} \setminus \mathcal{S}$ and $\alpha_i = 1/\nu N$ iff $i \in \mathcal{S} \setminus \mathcal{B}$. Shang et al. (2017) define the SVC based uncertainty set as

$$\mathcal{U}_\nu(\boldsymbol{\alpha}, \mathcal{D}) = \left\{ \boldsymbol{u} \,\middle|\, \sum_{i \in \mathcal{S}} \alpha_i ||\boldsymbol{Q}(\boldsymbol{u} - \boldsymbol{u}^{(i)})||_1 \leq \theta \right\} \tag{5}$$

where $||\cdot||_1$ stands for the ℓ_1-norm and

$$\theta = \min_{i' \in \mathcal{B}} \left(\sum_{i \in \mathcal{S}} \alpha_i ||\boldsymbol{Q}(\boldsymbol{u}^{(i')} - \boldsymbol{u}^{(i)}||_1 \right) \tag{6}$$

Using auxiliary variables $\boldsymbol{v} = [v_1, \ldots, v_N]$, one can then reformulate $\mathcal{U}_\nu(\boldsymbol{\alpha}, \mathcal{D})$ as a polyhedron:

$$\mathcal{U}_\nu(\boldsymbol{\alpha}, \mathcal{D}) = \left\{ \boldsymbol{u} \,\middle|\, \begin{array}{c} \exists \boldsymbol{v}_i, \ i \in \mathcal{S} \ s.t \\ \sum_{i \in \mathcal{S}} \alpha_i \boldsymbol{v}_i^T \boldsymbol{1} \leq \theta \\ -\boldsymbol{v}_i \leq \boldsymbol{Q}(\boldsymbol{u} - \boldsymbol{u}^{(i)}) \leq \boldsymbol{v}_i, \ i \in \mathcal{S} \end{array} \right\} \tag{7}$$

which is bounded and nonempty for $0 < \nu < 1$ (Shang et al., 2017). Based on formulation (7), the left-hand side of the robust linear constraint

$$\max_{\boldsymbol{a} \in \mathcal{U}_\nu(\boldsymbol{\alpha}, \mathcal{D})} \boldsymbol{a}^T \boldsymbol{x} \leq b \tag{8}$$

is thus equivalent to the following LP:

$$\max \ \boldsymbol{a}^T \boldsymbol{x} \tag{9}$$
$$\text{s.t.} \quad \sum_{i \in \mathcal{S}} \alpha_i \boldsymbol{1}^T \boldsymbol{v}_i \leq \theta \tag{10}$$
$$-\boldsymbol{v}_i \leq \boldsymbol{Q}(\boldsymbol{a} - \boldsymbol{u}^{(i)}) \leq \boldsymbol{v}_i \ \forall i \in \mathcal{S} \tag{11}$$

, which is feasible and bounded since $\mathcal{U}_\nu(\boldsymbol{\alpha}, \mathcal{D})$ is nonempty and bounded for $0 < \nu < 1$. By strong duality, the dual of this problem

$$\min \sum_{i\in\mathcal{S}} (\boldsymbol{\mu}_i - \boldsymbol{\lambda}_i)^T \boldsymbol{Q}\boldsymbol{u}^{(i)} + \eta\theta \tag{12}$$

$$\text{s.t.} \sum_{i\in\mathcal{S}} \boldsymbol{Q}(\boldsymbol{\lambda}_i - \boldsymbol{\mu}_i) + \boldsymbol{x} = \boldsymbol{0} \tag{13}$$

$$\boldsymbol{\lambda}_i + \boldsymbol{\mu}_i = \eta\alpha_i \boldsymbol{1} \;\forall i \in \mathcal{S} \tag{14}$$

$$\boldsymbol{\lambda}_i, \boldsymbol{\mu}_i \in \mathbb{R}^n_+ \quad \forall i \in \mathcal{S} \tag{15}$$

$$\eta \geq 0 \tag{16}$$

is also feasible and bounded and their optimal value coincide. Therefore the robust counterpart or constraint (8) is given by

$$\sum_{i\in\mathcal{S}} (\boldsymbol{\mu}_i - \boldsymbol{\lambda}_i)^T \boldsymbol{Q}\boldsymbol{u}^{(i)} + \eta\theta \leq b \tag{17}$$

with the additional constraints (13)-(16).

Formulation (17), (13)-(16) suggests that the size of $\mathcal{S}$ greatly influences the size of the robust counterpart. Consider a robust constraint with m uncertain parameters that belong to the uncertainty set $\mathcal{U}_\nu(\boldsymbol{\alpha}, \mathcal{D})$ defined from support vectors $\mathcal{S}$. The robust counterpart has $2 \cdot |\mathcal{S}| + 1$ new variables (15) and $m \cdot (|\mathcal{S}| + 1)$ new constraints (13), (14) compared to the original one. An increase in the size of the data set $\mathcal{D}$ or of the input parameter ν may thus lead to an intractable formulation, which represents a significant limitation to the application of this technique in practice. This phenomenon is particularly problematic in the context of big data where one would need to take advantage of the information contained in a large volume of data.

3. APPROXIMATIONS OF SVC-BASED UNCERTAINTY SETS

The computational burden described above is a significant drawback for the practitioners interested in robust solutions that use data-driven uncertainty sets based on SVC. We propose a method to alleviate this limitation, with the objective to define an approximate cluster that is highly similar to $\mathcal{U}_\nu(\boldsymbol{\alpha}, \mathcal{D})$ but uses a number of support vectors that is significantly lower than the original one. Our approach consists in computing a vector $\hat{\boldsymbol{\alpha}} = [\hat{\alpha}_1, \ldots, \hat{\alpha}_N]$ of modified weights to derive an approximation of the original uncertainty set. We use a single input parameter $K \in \mathbb{N}^+$ that directly defines the maximum number of strictly positive $\hat{\alpha}_i$ thus limiting the number of support vectors used to define the (approximate) uncertainty set, which is then defined following (5) as the set $\mathcal{U}_\nu(\hat{\boldsymbol{\alpha}}, \mathcal{D})$. This leads to a reduction of the number of variables and constraints in the LP (12)-(16) as well as its robust counterpart and gives the user a parameter to control the tractability of the model.

In what follows, we assume that we start from an initial uncertainty set $\mathcal{U}_\nu(\boldsymbol{\alpha}, \mathcal{D})$ obtained with the methodology of Shang et al. (2017). For simplicity we use the notations

$$\theta^{(i')} = \sum_{i\in\mathcal{S}} \alpha_i ||\boldsymbol{Q}(\boldsymbol{u}^{(i')} - \boldsymbol{u}^{(i)})||_1 \tag{18}$$

$$\hat{\theta}^{(i')} = \sum_{i\in\mathcal{S}} \hat{\alpha}_i ||\boldsymbol{Q}(\boldsymbol{u}^{(i')} - \boldsymbol{u}^{(i)})||_1 \tag{19}$$

to refer to the weighted sum calculated in point $\boldsymbol{u}^{(i')}$ that is used in definition (5) of $\mathcal{U}_\nu(\boldsymbol{\alpha}, \mathcal{D})$ and $\mathcal{U}_\nu(\hat{\boldsymbol{\alpha}}, \mathcal{D})$. We measure the quality of an approximation based on the cumulative absolute deviation of $\hat{\theta}^{(i')}$ from $\theta^{(i')}$ over $\mathcal{D}$.

In the remainder of this section we describe two different approaches to select the set of data points $\widehat{\mathcal{S}} \subset \mathcal{S}$ with strictly positive modified weight. The vector $\hat{\boldsymbol{\alpha}}$ is then computed using a linear program that maximizes the quality of the approximation with respect to $\widehat{\mathcal{S}}$.

3.1 K-medoid based subset selection

From the construction of the SVC, we can assume that relatively close samples that lie strictly outside $\mathcal{U}_\nu(\boldsymbol{\alpha}, \mathcal{D})$ have similar contributions to cluster boundaries. Our first approach selects a subset $\widehat{\mathcal{S}}$ of K data points, each representing the support vectors that are "close" to them. Our procedure relies on the so-called K-medoid clustering algorithm Kaufman and Rousseeuw (1990) to select the subset of points $\widehat{\mathcal{S}}$ that aggregates the information contained in $\mathcal{S}$. We compute the distances between two points $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^m$ using the ℓ_2-norm $||\boldsymbol{u} - \boldsymbol{v}||_2$. We define the set $\widehat{\mathcal{S}}$ of selected support vectors as the union of the K cluster centers collected as the output of the algorithm with the set $\mathcal{B}$, as represented in Figure 2.

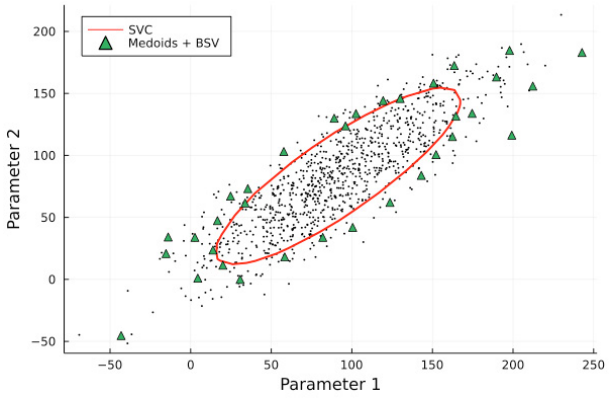

Fig. 2. Subset $\widehat{\mathcal{S}}$ of support vectors selected with the K-medoid approach

3.2 Crown based subset selection

The idea motivating the "crown" method is to only keep the K support vectors that are the closest to the cluster boundaries. We can achieve this selection easily by expressing K as a fraction $\delta = K/N$ of the total number of points and compute a second SVC with parameter $(\nu - \delta)$. Let $\mathcal{S}'$ be the resulting set of support vectors: We simply define $\widehat{\mathcal{S}} = \mathcal{S} \setminus \mathcal{S}'$ as the set of selected support vectors for the approximation. The set $\widehat{\mathcal{S}}$ contains at most K of the original support vectors. Figure 3 illustrates the subset obtained for the example introduced above.

3.3 Updating coefficients $\hat{\boldsymbol{\alpha}}$

The classification of point (i') as being inside or outside of the SVC-based uncertainty set $\mathcal{U}_\nu(\boldsymbol{\alpha}, \mathcal{D})$ is based on the comparison of $\theta^{(i')}$ with a given threshold θ. A reasonable approach to approximate the $\mathcal{U}_\nu(\boldsymbol{\alpha}, \mathcal{D})$ thus consists in defining modified wieghts $\hat{\boldsymbol{\alpha}}$ in such a way that $\hat{\theta}^{(i')}$ is comparable to $\theta^{(i')}$ for any sample point i'. This leads

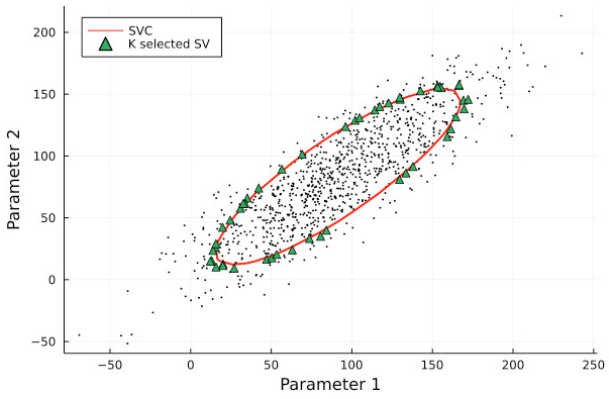

Fig. 3. Subset $\widehat{\mathcal{S}}$ of support vectors selected with the Crown approach when $N = 1000$ and $K = 30$

to the definition of a distance between $\mathcal{U}_\nu(\boldsymbol{\alpha}, \mathcal{D})$ and its approximation $\mathcal{U}_\nu(\hat{\boldsymbol{\alpha}}, \mathcal{D})$ as:

$$\Delta = \sum_{i' \in \mathcal{D}} |\hat{\theta}^{(i')} - \theta^{(i')}| \tag{20}$$

Obtaining the best approximation then comes down to computing weights $\hat{\boldsymbol{\alpha}}$ that minimize Δ, which can be formulated as the following LP:

$$\min \Delta \tag{21}$$

$$\text{s.t.} \sum_{i' \in \mathcal{D}} \sum_{i \in \mathcal{S}} (\alpha_i - \hat{\alpha}_i) ||\boldsymbol{Q}(\boldsymbol{u}^{(i')} - \boldsymbol{u}^{(i)})||_1 \leq \Delta \tag{22}$$

$$\sum_{i' \in \mathcal{D}} \sum_{i \in \mathcal{S}} (\hat{\alpha}_i - \alpha_i) ||\boldsymbol{Q}(\boldsymbol{u}^{(i')} - \boldsymbol{u}^{(i)})||_1 \leq \Delta \tag{23}$$

$$\hat{\alpha}_i \geq 0 \qquad \forall i \in \mathcal{S} \tag{24}$$

$$\Delta \geq 0 \tag{25}$$

This LP enables us to derive the optimal weights $\hat{\boldsymbol{\alpha}}$ and the approximate uncertainty set $\mathcal{U}_\nu(\hat{\boldsymbol{\alpha}}, \mathcal{D})$ is the defined as:

$$\mathcal{U}_\nu(\hat{\boldsymbol{\alpha}}, \mathcal{D}) = \left\{ \mathbf{u} \mid \sum_{i \in \widehat{\mathcal{S}}} \hat{\alpha}_i ||\mathbf{Q}(\mathbf{u} - \mathbf{u}^{(i)})||_1 \leq \theta \right\} \tag{26}$$

where θ is the initial bound of (5).

4. APPLICATION TO AN ASSEMBLY PROBLEM

We consider the production planning problem introduced in Loger et al. (2021) where an assembly line combines different components into a set of final products over a finite planning horizon $\mathcal{T}$. The set of final products and components are respectively denoted by I and J. d_{it} denotes the demand for product $i \in I$ faced by the system in each period $t \in \mathcal{T}$. For each product $i \in I$ and component $j \in J$, let r_{ij} be the number of components j required to produce one unit of product i. We assume that all component are available in a warehouse managed by a third party logistic provider (TPL). The TPL is in charge of delivering components to the assembly line according to its orders as represented on Figure 4. For each picking operation, an operator collects a given quantity of a single type of component, bounded by a maximum batch size of m_j units. Thus several picking operations of the same component $j \in J$ may be scheduled in the same period due to this limitation. We presume that the

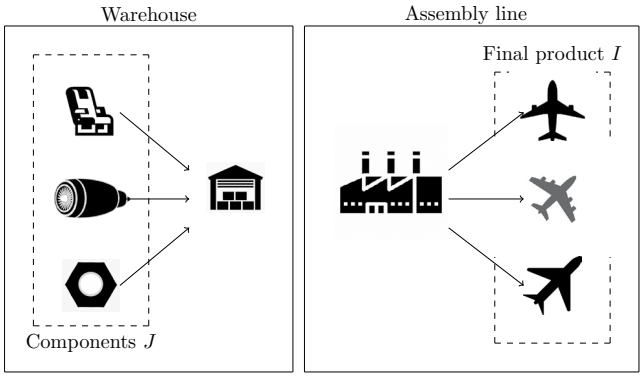

Fig. 4. Diagram of the problem

picking time of a particular batch of component j of size x_j is composed of (1) a fixed time p_j, that corresponds to the travel time between the shipping point and the zone where components j are stored and (2) a picking time per unit denoted τ_j. The overall picking time of a given period $t \in \mathcal{T}$ is bounded by the maximum work capacity of the TPL C_t. Any demand for product i that is not satisfied immediately is backlogged and incurs a backlogging penalty cost b_i in each period until the corresponding product is assembled in a subsequent period. Whenever a component j is available on the assembly line but is not immediately used to manufacture an end product, it disturbs the production process by interfering with people and other goods moving nearby. We model this situation with a per-unit, per-period obstruction cost o_j. The problem consists in planning the quantity of each component delivered to the assembly line by the TPL in each period such that the sum of the obstruction and backlogging costs is minimized. To model this problem, we define variables x_{jt} which represent the quantity of components $j \in J$ that should be delivered to the assembly line during period t. These variables are closely related to variables v_{jt}, representing the number of distinct picking operations of components $j \in J$ performed during period t. Variables s_{jt} represent the number of components $j \in J$ held on the border of the assembly line at the end of period t. Variables u_{it} denote the number of final products $i \in I$ produced during period t and finally, variables P_t and O_t denote the total backlogging penalty and obstruction costs incurred in period t, respectively.

We formulate the problem with the following MIP:

$$\min \sum_{t=1}^{T} O_t + P_t \tag{27}$$

$$\text{s.t. } s_{jt} = s_{j1} + \sum_{k=1}^{t-1} x_{jk} - \sum_{i \in I} r_{ij} u_{ik} \;\; \forall j \in J, \forall t \in \mathcal{T} \setminus 1 \tag{28}$$

$$\sum_{i \in I} u_{it} r_{ij} \leq s_{jt} + x_{jt} \qquad \forall j \in J, \forall t \in \mathcal{T} \tag{29}$$

$$\sum_{k=1}^{t} u_{ik} \leq \sum_{k=1}^{t} d_{ik} \qquad \forall i \in I, \forall t \in \mathcal{T} \tag{30}$$

$$v_{jt} \geq x_{jt}/m_j \qquad \forall j \in J, \forall t \in \mathcal{T} \tag{31}$$

$$\sum_{j \in J} v_{jt} p_j + \tau_j x_{jt} \leq C_t \qquad \forall t \in \mathcal{T} \tag{32}$$

$$O_t \geq \sum_{j \in J} o_j \left(s_{jt} + x_{jt} - \sum_{i \in I} u_{it} r_{ij} \right) \quad \forall t \in \mathcal{T} \tag{33}$$

$$P_t \geq \sum_{i \in I} b_i \left(\sum_{k=1}^{t} d_{ik} - u_{ik} \right) \quad \forall t \in \mathcal{T} \tag{34}$$

$$x_{jt}, s_{jt}, P_t, O_t \in \mathbb{R}_+ \quad \forall j \in J, \forall t \in \mathcal{T} \tag{35}$$

$$u_{it}, v_{jt} \in \mathbb{N}_+ \quad \forall i \in I, \forall j \in J, \forall t \in \mathcal{T} \tag{36}$$

The objective (27) aims at minimizing the total cost incurred over the planning horizon. The inventory balance constraints (28) update the stock levels in each period. Constraints (29) ensure that the quantity of component $j \in J$ available in period t is sufficient to perform the planned assembly operations. Constraints (30) impose that the system never manufactures more units of product $i \in I$ than the expressed demand. The minimum number of picking batches corresponding to the quantity of components j ordered in each period t is defined in constraints (31). Constraints (32) ensure that the total picking time does not exceed the picking capacity of the TPL provider in any period t. Constraints (33) and (34) define the components holding costs and final product backlogging costs incurred in each period t, respectively. Finally, constraints (35) and (36) define the domain of the decision variables.

4.1 Picking time uncertainties

We consider a case where the manufacturer has an incomplete or imprecise knowledge on the picking times. As a consequence, some combinations of his orders may exceed the picking capacity of the TPL provider, forcing the latter to postpone some operations to subsequent periods. We assume that we are given a set of historical setup times $\mathcal{D} = \{\mathbf{p}^{(1)}, \ldots, \mathbf{p}^{(N)}\}$ and we present a comparison of the performance of three different robust models, each based on a particular implementation of $\mathcal{U}(\mathcal{D})$. Namely, we compare the performances of the approximate SVC uncertainty set we propose built with the K-medoid method and the Crown method with both the original SVC of Shang et al. (2017) and a budget based uncertainty set similar to the one derived in Bertsimas and Sim (2004):

$$\mathcal{U}_\Gamma(\mathcal{D}) = \{\boldsymbol{p} = \bar{\boldsymbol{p}} + \hat{\boldsymbol{p}}^\intercal \boldsymbol{z} \,|\mathbf{1}^\intercal \boldsymbol{z} \leq \Gamma, \boldsymbol{z} \in [0,1]^m\} \tag{37}$$

, where $\hat{\boldsymbol{p}} = [\hat{p}_1, \ldots, \hat{p}_m]$ for m the number of components is calculated such that the interval $[\bar{p} - \hat{p}, \bar{p} + \hat{p}]$ contains 95% of the data $\{p^{(1)}, \ldots, p^{(N)}\}$. We assume that the TPL provider organizes the storage of the components in order to optimize the picking operations. Specifically, components are stored in such a way that their accessibility improves with their order frequency. We consider three types of components and separate them based on their storage area. We assume that both the mean and the variability of the setup times decrease with component accessibility. In our instances, we thus we generate setup times using Gamma distributions with different shape parameters for each type of component. As in Loger et al. (2021), we generate the instances in order to reflect a practical application case inspired from the aircraft industry.

4.2 Experimental results

Figure 5 and 6 presents the results obtained for instances with 5 final products over 10 periods and for five different

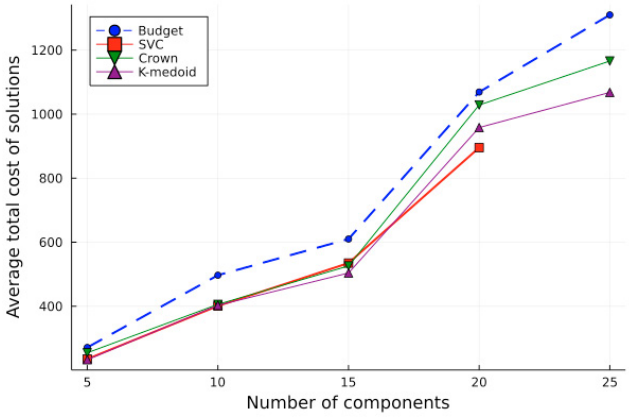

Fig. 5. Average total cost of the best solution when the number of component increases

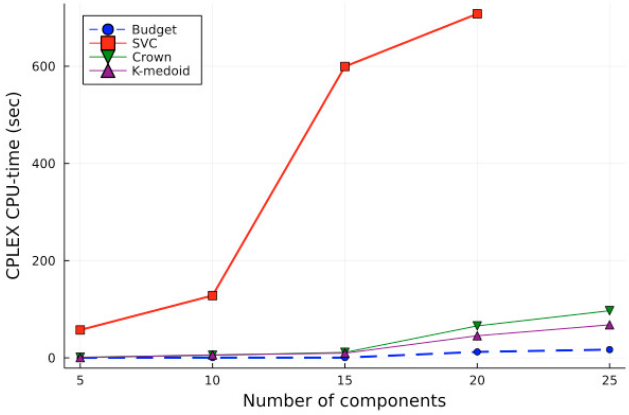

Fig. 6. CPLEX CPU-time (s) of the best solution when the number of component increases

number of components in $\{5, 10, 15, 20, 25\}$. The different uncertainty sets are built with a data-set $\mathcal{D}$ of $N = 400$ data-points and the solutions obtained are evaluated on a larger test data-set of $N' = 10000$ data points following the same distribution. For all uncertainty sets, we consider different values of parameters Γ and ν. The budget based solutions presented in Figure 5 and 6 always correspond to the best overall solutions obtained with this type of uncertainty set. The solutions presented for the SVC-based uncertainty set and the two approximations are always obtained with the same values of ν and the additional parameter is set to $K = 2 * |J|$.

Figure 5 clearly shows that SVC based model can help to reduce costs when in the case of asymmetric distributions. We observe that both the Crown and the K-medoid based approximations efficiently captured asymmetries in the distribution of picking time and lead to better solutions than the classic budget based uncertainty set. For a given value of ν, the solution obtained with the K-medoid based approximation tends to be closer to the solution obtained with $\mathcal{U}_\nu(\boldsymbol{\alpha}, \mathcal{D})$ than those obtained with the Crown based approximation.

Figure 6 shows the solving time of the four different robust counterparts obtained for each instances. We observe for both approximation methods that restricting the number of support vector efficiently reduces the time needed to

solve the robust counterpart. On the instance with 25 components, the gap achieved with the SVC-based model exceeds 99% at the end of the maximum allocated time of 20 minutes. On the same instance, the K-medoid and the crown based approximations are respectively solved to optimality in 68 and 97 seconds. In our test set, the budget-based models always required less computational efforts than the SVC-based model and the two approximations, with a time difference that increases with the number of components.

5. CONCLUSION AND PERSPECTIVES

We propose two approximations of the SVC-based uncertainty set that efficiently reduce the computational effort while maintaining good quality solutions. Our main contribution is to make it possible to apply this type of data-driven approach to larger industrial problems in a reasonable amount time. Both methods lead to comparable computational time but the K-medoid based method produces a better approximation than the Crown based one. Future works could extend the K-medoid based approach by defining a specific metric space to compute the distance between two points in the K-medoid algorithm in order to improve the quality of the approximation. Using an optimization approach to select the best support vectors to define $\mathcal{U}_\nu(\hat{\boldsymbol{\alpha}}, \mathcal{D})$ could also be an interesting perspective for future research work.

REFERENCES

Aouam, T. and Brahimi, N. (2013). Integrated production planning and order acceptance under uncertainty: A robust optimization approach. *European Journal of Operational Research*, 228, 504–515.

Ben-Hur, A., Horn, D., Siegelmann, H., and Vapnik, V. (2001). Support vector clustering. *Journal of Machine Learning Research*, 2, 125–137.

Ben-Tal, A. and Nemirovski, A. (1998). Robust convex optimization. *Mathematics of Operations Research*, 23(4), 769–805.

Bertsimas, D. and Brown, D.B. (2009). Constructing Uncertainty Sets for Robust Linear Optimization. *Operations Research*, 57(6), 1483–1495.

Bertsimas, D., Gupta, V., and Kallus, N. (2018). Data-driven robust optimization. *Mathematical Programming*, 167(2), 235–292.

Bertsimas, D. and Sim, M. (2004). The price of robustness. *Operations Research*, 52, 35–53.

Bertsimas, D. and Thiele, A. (2006). A robust optimization approach to inventory theory. *Operations Research*, 54, 150–168.

Chassein, A., Dokka, T., and Goerigk, M. (2019). Algorithms and uncertainty sets for data-driven robust shortest path problems. *European Journal of Operational Research*, 274(2), 671–686.

El Ghaoui, L., Oustry, F., and Lebret, H. (1998). Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9(1), 33–52.

Goerigk, M. and Kurtz, J. (2021). Data-Driven Robust Optimization using Unsupervised Deep Learning. *arXiv:2011.09769*.

Han, B., Shang, C., and Huang, D. (2021). Multiple kernel learning-aided robust optimization: Learning algorithm, computational tractability, and usage in multi-stage decision-making. *European Journal of Operational Research*, 292(3), 1004–1018.

José Alem, D. and Morabito, R. (2012). Production planning in furniture settings via robust optimization. *Computers & Operations Research*, 39(2), 139–150.

Kaufman, L. and Rousseeuw, P.J. (1990). Partitioning Around Medoids (Program PAM). In *Finding Groups in Data*, 68–125. John Wiley & Sons, Ltd.

Loger, B., Dolgui, A., Lehuédé, F., and Massonnet, G. (2021). A robust data driven approach to supply planning. In *Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems*, 169–178. Springer International Publishing, Cham.

Ning, C. and You, F. (2018). Data-driven decision making under uncertainty integrating robust optimization with principal component analysis and kernel smoothing methods. *Computers & Chemical Engineering*, 112, 190 – 210.

Ning, C. and You, F. (2019). Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming. *Computers & Chemical Engineering*, 125, 434–448.

Qiu, R., Sun, Y., Shu, P., and Sun, M. (2019). Robust multi-product inventory optimization under support vector clustering-based data-driven demand uncertainty set. *Soft Computing*, 24, 6259–6275.

Shang, C., Huang, X., and You, F. (2017). Data-driven robust optimization based on kernel learning. *Computers & Chemical Engineering*, 106, 464–479.

Shen, F., Zhao, L., Du, W., Zhong, W., and Qian, F. (2020). Large-scale industrial energy systems optimization under uncertainty: A data-driven robust optimization approach. *Applied Energy*, 259, 114199.

Soyster, A.L. (1973). Technical note—convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21(5), 1154–1157.

Sözüer, S. and Thiele, A.C. (2016). The State of Robust Optimization. In *Robustness Analysis in Decision Aiding, Optimization, and Analytics*, International Series in Operations Research & Management Science, 89–112. Springer International Publishing, Cham.

Thorsen, A. and Yao, T. (2017). Robust inventory control under demand and lead time uncertainty. *Annals of Operations Research*, 257, 207–236.

Varas, M., Maturana, S., Pascual, R., Vargas, I., and Vera, J. (2014). Scheduling production for a sawmill: A robust optimization approach. *International Journal of Production Economics*, 150, 37–51.

Wei, C., Li, Y., and Cai, X. (2011). Robust optimal policies of production and inventory with uncertain returns and demand. *International Journal of Production Economics*, 134(2), 357–367.

Wu, W., Liu, R., Yang, Q., and Quek, T.Q.S. (2021). Learning-Based Robust Resource allocation for D2D Underlaying Cellular Network. *arXiv:2105.08324*.

Zhang, Y., Jin, X., Feng, Y., and Rong, G. (2018). Data-driven robust optimization under correlated uncertainty: A case study of production scheduling in ethylene plant. *Computers & Chemical Engineering*, 109, 48–67.