Operations Research Letters 44 (2016) 6–11

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

K-adaptability in two-stage distributionally robust binary programming[☆]

Grani A. Hanasusanto [a,*], Daniel Kuhn [a], Wolfram Wiesemann [b]

[a] *Risk Analytics and Optimization Chair, École Polytechnique Fédérale de Lausanne, Switzerland*
[b] *Imperial College Business School, Imperial College London, United Kingdom*

ARTICLE INFO

Article history:
Received 24 April 2015
Received in revised form
26 October 2015
Accepted 26 October 2015
Available online 10 November 2015

Keywords:
Distributionally robust optimization
Integer programming
Two-stage problems

ABSTRACT

We propose to approximate two-stage distributionally robust programs with binary recourse decisions by their associated K-adaptability problems, which pre-select K candidate second-stage policies here-and-now and implement the best of these policies once the uncertain parameters have been observed. We analyze the approximation quality and the computational complexity of the K-adaptability problem, and we derive explicit mixed-integer linear programming reformulations. We also provide efficient procedures for bounding the probabilities with which each of the K second-stage policies is selected.
© 2015 Elsevier B.V. All rights reserved.

1. Introduction

We study two-stage distributionally robust programs of the form

$$\begin{aligned} \text{minimize} \quad & \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{-OCE}_U \left[\tilde{\boldsymbol{\xi}}^\top \boldsymbol{C}\,\boldsymbol{x} + \min_{\boldsymbol{y}\in\mathcal{Y}} \left\{ \tilde{\boldsymbol{\xi}}^\top \boldsymbol{Q}\,\boldsymbol{y} : \boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y} \le \boldsymbol{H}\tilde{\boldsymbol{\xi}} \right\} \right] \\ \text{subject to} \quad & \boldsymbol{x} \in \mathcal{X}, \end{aligned} \qquad (\mathcal{DP})$$

where

$$\mathbb{P}\text{-OCE}_U\left[\phi(\tilde{\boldsymbol{\xi}})\right] = \inf_{\theta\in\mathbb{R}} \theta + \mathbb{E}_{\mathbb{P}}\left[U(\phi(\tilde{\boldsymbol{\xi}}) - \theta)\right]$$

denotes the *optimized certainty equivalent* of a disutility function U under the probability distribution $\mathbb{P}$. Here, $\mathcal{X} \subseteq \mathbb{R}^N$ and $\mathcal{Y} \subseteq \{0,1\}^M$ are bounded mixed-integer linear sets, while $\boldsymbol{C} \in \mathbb{R}^{Q\times N}$, $\boldsymbol{Q} \in \mathbb{R}^{Q\times M}$, $\boldsymbol{T} \in \mathbb{R}^{L\times N}$, $\boldsymbol{W} \in \mathbb{R}^{L\times M}$ and $\boldsymbol{H} \in \mathbb{R}^{L\times Q}$. The first-stage or *here-and-now* decisions $\boldsymbol{x}$ are selected prior to the observation of the uncertain parameters $\tilde{\boldsymbol{\xi}} \in \mathbb{R}^Q$, and the second-stage or *wait-and-see* decisions $\boldsymbol{y}$ are chosen after $\tilde{\boldsymbol{\xi}}$ has been revealed. We assume that all components of $\boldsymbol{y}$ are binary, while $\boldsymbol{x}$ may have continuous as well as binary components. Problem ($\mathcal{DP}$) minimizes the worst-case optimized certainty equivalent over all distributions $\mathbb{P}$ from an ambiguity set $\mathcal{P}$. Problems of the type ($\mathcal{DP}$) have many applications, for example in facility location, vehicle routing, unit commitment, layout planning, project scheduling, portfolio selection and game theory.

Despite the broad applicability of problem $\mathcal{DP}$, its numerical solution is extremely challenging. Instead of solving ($\mathcal{DP}$) exactly, we propose to solve its associated K-adaptability problem

$$\begin{aligned} \text{minimize} \quad & \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{-OCE}_U \left[\tilde{\boldsymbol{\xi}}^\top \boldsymbol{C}\,\boldsymbol{x} + \min_{k\in\mathcal{K}} \left\{ \tilde{\boldsymbol{\xi}}^\top \boldsymbol{Q}\,\boldsymbol{y}^k : \boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}^k \le \boldsymbol{H}\tilde{\boldsymbol{\xi}} \right\} \right] \\ \text{subject to} \quad & \boldsymbol{x} \in \mathcal{X}, \quad \boldsymbol{y}^k \in \mathcal{Y}, \quad k \in \mathcal{K}, \end{aligned} \qquad (\mathcal{DP}_K)$$

where $\mathcal{K} = \{1, \ldots, K\}$. Thus, we pre-select exactly K of the $|\mathcal{Y}| \lesssim 2^M$ possible second-stage decisions in the first stage. After the realization $\boldsymbol{\xi}$ of the random parameters $\tilde{\boldsymbol{\xi}}$ has been observed, we implement the best of these K pre-selected candidate decisions. More precisely, among all candidate decisions that are feasible for $\boldsymbol{\xi}$, we implement the one that achieves the lowest second-stage cost in scenario $\boldsymbol{\xi}$. If all K candidate decisions are infeasible for a given $\boldsymbol{\xi}$, then the innermost minimum in ($\mathcal{DP}_K$) is interpreted as an infimum that evaluates to $+\infty$. By construction, the K-adaptability problem may provide a strict upper bound on the optimal value of problem ($\mathcal{DP}$) unless $K = |\mathcal{Y}|$, but the hope is that the corresponding optimality gap is small already for some $K \ll |\mathcal{Y}|$. The K-adaptability problem has been studied in [3,7] in the context of robust two-stage integer programming, where the objective is to minimize the worst-case cost over all uncertainty

[☆] An extended version of the paper with all the proofs can be accessed at http://www.optimization-online.org/DB_HTML/2015/04/4878.html.

* Corresponding author.
E-mail address: grani.hanasusanto@epfl.ch (G.A. Hanasusanto).

http://dx.doi.org/10.1016/j.orl.2015.10.006
0167-6377/© 2015 Elsevier B.V. All rights reserved.

realizations $\boldsymbol{\xi} \in \Xi$. We refer to [7] for a review of the literature on robust two-stage integer programming.

The rest of the paper develops as follows. We provide a detailed formulation of the K-adaptability problem $(\mathcal{DP}_K)$ in Section 2. Section 3 studies the approximation quality as well as an explicit MILP reformulation of the distributionally robust K-adaptability problem with objective uncertainty, that is, for instances of $(\mathcal{DP}_K)$ where $\tilde{\boldsymbol{\xi}}$ impacts only the objective function. We close with numerical results in Section 4. The Electronic Companion contains a treatment of generic instances of $(\mathcal{DP}_K)$ where both the objective function and the constraints are affected by uncertainty, as well as most of the proofs.

Notation. Variables with tilde signs represent random objects. We denote by $\mathbf{e}$ the vector of ones and by $\mathbf{e}_i$ the vector whose i-th entry is 1 while all other entries are 0. The indicator function $\mathbb{I}[\mathcal{E}]$ of a logical expression $\mathcal{E}$ is defined through $\mathbb{I}[\mathcal{E}] = 1$ if $\mathcal{E}$ is true; $= 0$ otherwise. To avoid tedious case distinctions, we define the minimum (maximum) of an empty set as $+\infty$ $(-\infty)$.

2. Problem formulation

Problem $(\mathcal{DP})$ accounts for both ambiguity aversion (through the ambiguity set) and risk aversion (through the optimized certainty equivalent). We discuss both of these components in turn.

Ambiguity set. By construction, the ambiguity set $\mathcal{P}$ contains all probability distributions that share certain known properties of the unknown true distribution $\mathbb{P}^0$ of $\tilde{\boldsymbol{\xi}}$. Hedging against the worst probability distribution within the ambiguity set $\mathcal{P}$ reflects an aversion against distributional ambiguity, which enjoys strong justification from decision theory [4,5].

We henceforth assume that the ambiguity set $\mathcal{P}$ is of the form

$$\mathcal{P} = \left\{ \mathbb{P} \in \mathcal{M}_+(\mathbb{R}^Q) : \mathbb{P}\left[\tilde{\boldsymbol{\xi}} \in \Xi\right] = 1, \quad \mathbb{E}_{\mathbb{P}}\left[\boldsymbol{g}(\tilde{\boldsymbol{\xi}})\right] \le \boldsymbol{c} \right\}, \tag{1}$$

where $\mathcal{M}_+(\mathbb{R}^Q)$ denotes the cone of nonnegative Borel measures supported on $\mathbb{R}^Q$. The support set Ξ is defined as the smallest set that is known to satisfy $\tilde{\boldsymbol{\xi}} \in \Xi$ w.p. 1, and it constitutes a nonempty bounded polytope of the form $\Xi = \left\{\boldsymbol{\xi} \in \mathbb{R}^Q : \boldsymbol{A}\boldsymbol{\xi} \le \boldsymbol{b}\right\}$ with $\boldsymbol{A} \in \mathbb{R}^{R\times Q}$ and $\boldsymbol{b} \in \mathbb{R}^R$. We assume that $\boldsymbol{c} \in \mathbb{R}^S$ and that $\boldsymbol{g} : \mathbb{R}^Q \to \mathbb{R}^S$ has convex piecewise linear component functions of the form

$$g_s(\boldsymbol{\xi}) = \max_{t\in\mathcal{T}} \boldsymbol{g}_{st}^\top \boldsymbol{\xi} \quad \forall s \in \mathcal{S} = \{1, \ldots, S\}.$$

Without loss of generality, the index t of the linear pieces ranges over the same index set $\mathcal{T} = \{1, \ldots, T\}$ for each component g_s. To ensure the applicability of strong semi-infinite duality results, we finally assume that the ambiguity set $\mathcal{P}$ contains a Slater point in the sense that there is a distribution $\mathbb{P} \in \mathcal{P}$ such that $\mathbb{E}_{\mathbb{P}}[g_s(\tilde{\boldsymbol{\xi}})] < \boldsymbol{c}$ for all $s \in \mathcal{S}$ for which $g_s(\boldsymbol{\xi})$ is nonlinear.

Ambiguity sets of the form (1) are flexible enough to encode bounds on the expected value and the mean absolute deviation of the unknown true distribution $\mathbb{P}^0$, and they can also be used to approximate nonlinear dispersion measures such as the variance or the standard deviation [11].

Optimized certainty equivalent. Intuitively, the optimized certainty equivalent $\mathbb{P}\text{-OCE}_U[\phi(\tilde{\boldsymbol{\xi}})]$ represents the expected present value of an optimized payment schedule that splits an uncertainty-affected future liability $\phi(\tilde{\boldsymbol{\xi}})$ into a fixed installment θ that is paid today and a remainder $\phi(\tilde{\boldsymbol{\xi}}) - \theta$ that is paid after the realization of $\tilde{\boldsymbol{\xi}}$ has been observed [2].

The decision maker's attitude towards risk is controlled by the choice of the disutility function U. We consider increasing, convex and piecewise affine disutility functions of the form

$$U(x) = \max_{i\in\mathcal{I}} \{s_i x + t_i\}, \quad \text{where } \mathcal{I} = \{1, \ldots, I\} \text{ and}$$
$$\boldsymbol{s} \ge \mathbf{0},\ \boldsymbol{s} \ne \mathbf{0}. \tag{2}$$

The choice $I = 1$, $\boldsymbol{s} = 1$ and $\boldsymbol{t} = 0$ corresponds to the worst-case expected value, whereas $I = 2$, $\boldsymbol{s} = ((1-\beta)^{-1}, 0)^\top$ and $\boldsymbol{t} = (0, 0)^\top$ recovers the worst-case conditional value-at-risk at level β, that is, the worst-case expected value of the $(1-\beta)\cdot 100\%$ largest outcomes.

3. The K-adaptability problem with objective uncertainty

In this section, we assume that the random parameters $\tilde{\boldsymbol{\xi}}$ only enter the objective function of the two-stage robust binary program $(\mathcal{DP})$:

$$\begin{array}{ll} \text{minimize} & \displaystyle\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{-OCE}_U\Big[\tilde{\boldsymbol{\xi}}^\top \boldsymbol{C}\boldsymbol{x} \\ & \displaystyle\quad + \min_{\boldsymbol{y}\in\mathcal{Y}} \left\{\tilde{\boldsymbol{\xi}}^\top \boldsymbol{Q}\boldsymbol{y} : \boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y} \le \boldsymbol{h}\right\}\Big] \\ \text{subject to} & \boldsymbol{x} \in \mathcal{X}, \end{array} \tag{$\mathcal{DPO}$}$$

where $\boldsymbol{h} \in \mathbb{R}^L$. Problem $(\mathcal{DPO})$ arises naturally in a number of application domains, such as traveling salesman and vehicle routing problems with uncertain travel times, network expansion problems with uncertain costs, facility location problems with uncertain customer demands and layout planning problems with uncertain production quantities. A discussion of generic instances of $(\mathcal{DP})$, where both the objective function and the constraints are affected by uncertainty, is relegated to the Electronic Companion.

In the following, we study the K-adaptability problem associated with $(\mathcal{DPO})$:

$$\begin{array}{ll} \text{minimize} & \displaystyle\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{-OCE}_U\Big[\tilde{\boldsymbol{\xi}}^\top \boldsymbol{C}\boldsymbol{x} \\ & \displaystyle\quad + \min_{k\in\mathcal{K}} \left\{\tilde{\boldsymbol{\xi}}^\top \boldsymbol{Q}\boldsymbol{y}^k : \boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}^k \le \boldsymbol{h}\right\}\Big] \\ \text{subject to} & \boldsymbol{x} \in \mathcal{X}, \quad \boldsymbol{y}^k \in \mathcal{Y},\ k \in \mathcal{K}. \end{array} \tag{3}$$

We note that the K-adaptability problem (3) is equivalent to the problem

$$\begin{array}{ll} \text{minimize} & \displaystyle\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{-OCE}_U\Big[\tilde{\boldsymbol{\xi}}^\top \boldsymbol{C}\boldsymbol{x} + \min_{k\in\mathcal{K}} \tilde{\boldsymbol{\xi}}^\top \boldsymbol{Q}\boldsymbol{y}^k\Big] \\ \text{subject to} & \boldsymbol{x} \in \mathcal{X}, \quad \boldsymbol{y}^k \in \mathcal{Y},\ k \in \mathcal{K} \\ & \boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}^k \le \boldsymbol{h} \quad \forall k \in \mathcal{K}, \end{array} \tag{$\mathcal{DPO}_K$}$$

where we have shifted the second-stage constraints to the first stage, see [7, Observation 1].

We now show that the problem $(\mathcal{DPO}_K)$ has an equivalent reformulation as a MILP. To this end, we interpret the objective function in $(\mathcal{DPO}_K)$ as a moment problem. The dual of this moment problem constitutes a semi-infinite program which we can simplify using a standard LP dualization.

Theorem 1. *For the ambiguity set $\mathcal{P}$ defined in (1) and the disutility function U defined in (2), problem $(\mathcal{DPO}_K)$ is equivalent to the following MILP.*

$$\begin{array}{ll} \textit{minimize} & \alpha + \boldsymbol{c}^\top \boldsymbol{\beta} + \theta \\ \textit{subject to} & \boldsymbol{x} \in \mathcal{X}, \quad \boldsymbol{y}^k \in \mathcal{Y}, \quad k \in \mathcal{K},\ \alpha \in \mathbb{R},\ \boldsymbol{\beta} \in \mathbb{R}^S_+ \\ & \boldsymbol{\gamma}^i \in \mathbb{R}^R_+, \quad \boldsymbol{\delta}^i \in \mathbb{R}^K_+, \quad \boldsymbol{\Lambda}^i \in \mathbb{R}^{S\times T}_+,\ i \in \mathcal{I} \\ & \boldsymbol{z}^{i,k} \in \mathbb{R}^M_+, \quad i \in \mathcal{I} \textit{ and } k \in \mathcal{K}, \quad \theta \in \mathbb{R} \\ & \boldsymbol{T}\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}^k \le \boldsymbol{h} \quad \forall k \in \mathcal{K} \\ & \left.\begin{array}{l} \boldsymbol{b}^\top \boldsymbol{\gamma}^i + t_i \le \alpha + s_i\theta, \quad \mathbf{e}^\top \boldsymbol{\delta}^i = 1, \quad \boldsymbol{\Lambda}^i \mathbf{e} = \boldsymbol{\beta} \\ \displaystyle\boldsymbol{A}^\top \boldsymbol{\gamma}^i + \sum_{s\in\mathcal{S}} \sum_{t\in\mathcal{T}} \Lambda^i_{st} \boldsymbol{g}_{st} = s_i \boldsymbol{C}\boldsymbol{x} + \sum_{k\in\mathcal{K}} s_i \boldsymbol{Q}\boldsymbol{z}^{i,k} \\ \left.\begin{array}{l} \boldsymbol{z}^{i,k} \le \boldsymbol{y}^k, \quad \boldsymbol{z}^{i,k} \le \delta^i_k \mathbf{e} \\ \boldsymbol{z}^{i,k} \ge (\delta^i_k - 1)\mathbf{e} + \boldsymbol{y}^k \end{array}\right\} \quad \forall k \in \mathcal{K} \end{array}\right\} \quad \forall i \in \mathcal{I}. \end{array} \tag{4}$$

By construction, the optimal value of problem ($\mathcal{DPO}_K$) constitutes an upper bound on the optimal value of problem ($\mathcal{DPO}$) as we restrict our flexibility in the second stage. For classical two-stage robust binary programs, it has been shown that the approximation provided by the K-adaptability formulation is tight whenever K exceeds the affine dimension of either the uncertainty set or the second-stage feasible region, see [7, Theorem 1]. We now show that this favorable approximation behavior generalizes to distributionally robust two-stage binary programs.

Theorem 2. *Problem ($\mathcal{DPO}_K$) has the same optimal value as problem ($\mathcal{DPO}$) if we choose $K \geq I \cdot \min\{\dim \mathcal{Y}, \operatorname{rk} \boldsymbol{Q}\} + I$ policies.*

We note that $\dim \mathcal{Y} \leq M$ and $\operatorname{rk} \boldsymbol{Q} \leq Q$ by construction. Without loss of generality, we can further assume that $\operatorname{rk} \boldsymbol{Q} \leq \dim \Xi + 1$. Indeed, if this is not the case, then there is a matrix $\boldsymbol{Q}'$ such that $\operatorname{rk} \boldsymbol{Q}' \leq \dim \Xi + 1$ and the optimal value and the optimal solutions to ($\mathcal{DPO}_K$) do not change if we replace $\boldsymbol{Q}$ with $\boldsymbol{Q}'$, see also [7, Remark 3].

3.1. Persistence

We now study the contribution of each second-stage policy $\boldsymbol{y}^k$ to the objective value in problem ($\mathcal{DPO}_K$). This analysis can provide important insights for practical decision-making. Amongst others, it can help to determine how much adaptability is needed (*i.e.*, it can inform the choice of K), and it can elucidate the relative 'criticality' of each second-stage policy.

We develop two alternative measures for the importance of a second-stage policy. We first determine the probability with which a particular policy $\boldsymbol{y}^k$ is chosen under a worst-case distribution in problem ($\mathcal{DPO}_K$). This approach is reminiscent of the persistency analysis in [8], and it is justified by the assumption that the decision maker optimizes in view of the *worst* distribution from within $\mathcal{P}$. Afterwards, we determine the minimum and maximum probability that a policy $\boldsymbol{y}^k$ is chosen if the unknown true probability distribution $\mathbb{P}^0$ can be *any* distribution within the ambiguity set $\mathcal{P}$. These probability bounds characterize the ambiguity of the persistence inherited by the ambiguity of $\mathbb{P}^0$ that governs $\tilde{\boldsymbol{\xi}}$.

We now study the probability that a particular second-stage policy $\boldsymbol{y}^k$ is chosen under a worst-case distribution for ($\mathcal{DPO}_K$).

Theorem 3. *For a feasible decision $(\boldsymbol{x}, \{\boldsymbol{y}^k\}_k)$ in problem ($\mathcal{DPO}_K$) with the ambiguity set (1) and the disutility function (2), and let $(\boldsymbol{\phi}, \{\boldsymbol{\chi}^i\}_i, \boldsymbol{\psi}, \{\boldsymbol{\omega}^i\}_i)$ be an optimal solution to the LP*

$$
\begin{array}{ll}
\text{maximize} & \displaystyle\sum_{i\in\mathcal{I}} t_i\phi_i + s_i(\boldsymbol{C}\boldsymbol{x})^\top \boldsymbol{\chi}^i + \psi_i \\
\text{subject to} & \left.\begin{array}{l}
\phi_i \in \mathbb{R}_+, \quad \boldsymbol{\chi}^i \in \mathbb{R}^Q, \quad \psi_i \in \mathbb{R}, \quad \boldsymbol{\omega}^i \in \mathbb{R}^S \\
\boldsymbol{A}\boldsymbol{\chi}^i \leq \boldsymbol{b}\phi_i \\
\psi_i \leq s_i(\boldsymbol{Q}\boldsymbol{y}^k)^\top \boldsymbol{\chi}^i \quad \forall k \in \mathcal{K} \\
\boldsymbol{g}_{st}^\top \boldsymbol{\chi}^i \leq \omega_s^i \quad \forall s \in \mathcal{S},\ \forall t \in \mathcal{T}
\end{array}\right\} \quad \forall i \in \mathcal{I} \\
& \boldsymbol{e}^\top \boldsymbol{\phi} = 1, \quad \boldsymbol{s}^\top \boldsymbol{\phi} = 1, \quad \displaystyle\sum_{i\in\mathcal{I}} \boldsymbol{\omega}^i \leq \boldsymbol{c}.
\end{array} \tag{5}
$$

Then, a worst-case distribution $\mathbb{P}^\star$ for $(\boldsymbol{x}, \{\boldsymbol{y}^k\}_k)$ in problem ($\mathcal{DPO}_K$) is defined through

$$
\mathbb{P}^\star\left[\tilde{\boldsymbol{\xi}} = \frac{\boldsymbol{\chi}^i}{\phi_i}\right] = \phi_i, \quad i \in \mathcal{I} : \phi_i > 0,
$$

and the probability with which a policy $\boldsymbol{y}^k$, $k \in \mathcal{K}$, is chosen under $\mathbb{P}^\star$ is

$$
\begin{aligned}
&\mathbb{P}^\star\left[\tilde{\boldsymbol{\xi}}^\top \boldsymbol{Q}\,\boldsymbol{y}^k \leq \tilde{\boldsymbol{\xi}}^\top \boldsymbol{Q}\,\boldsymbol{y}^{k'} \quad \forall k' \in \mathcal{K}\right] \\
&\quad = \sum_{i\in\mathcal{I}} \phi_i \cdot \mathbb{I}\left[(\boldsymbol{\chi}^i)^\top \boldsymbol{Q}\,\boldsymbol{y}^k \leq (\boldsymbol{\chi}^i)^\top \boldsymbol{Q}\,\boldsymbol{y}^{k'} \quad \forall k' \in \mathcal{K}\right].
\end{aligned}
$$

We now evaluate the maximum probability with which a second-stage policy $\boldsymbol{y}^k$ is chosen under any distribution $\mathbb{P} \in \mathcal{P}$.

Proposition 1. *Let $(\boldsymbol{x}, \{\boldsymbol{y}^k\}_k)$ be a feasible decision in problem ($\mathcal{DPO}_K$) with the ambiguity set (1) and the disutility function (2). The maximum probability with which policy $\boldsymbol{y}^k$, $k \in \mathcal{K}$, is chosen under any probability distribution $\mathbb{P} \in \mathcal{P}$ is given by the optimal value of the LP*

$$
\begin{array}{ll}
\text{minimize} & \alpha + \boldsymbol{c}^\top \boldsymbol{\beta} \\
\text{subject to} & \alpha \in \mathbb{R}, \quad \boldsymbol{\beta} \in \mathbb{R}^S_+, \quad \boldsymbol{\gamma} \in \mathbb{R}^R_+, \quad \boldsymbol{\delta} \in \mathbb{R}^K_+, \quad \boldsymbol{\kappa} \in \mathbb{R}^R_+ \\
& \boldsymbol{\Lambda} \in \mathbb{R}^{S\times T}_+, \quad \boldsymbol{\Phi} \in \mathbb{R}^{S\times T}_+ \\
& \alpha \geq \boldsymbol{b}^\top \boldsymbol{\gamma} + 1, \quad \alpha \geq \boldsymbol{b}^\top \boldsymbol{\kappa}, \quad \boldsymbol{\Lambda}\boldsymbol{e} = \boldsymbol{\beta}, \quad \boldsymbol{\Phi}\boldsymbol{e} = \boldsymbol{\beta} \\
& \boldsymbol{A}^\top \boldsymbol{\gamma} + \displaystyle\sum_{s\in\mathcal{S}}\sum_{t\in\mathcal{T}} \Lambda_{st}\boldsymbol{g}_{st} = \sum_{k'\in\mathcal{K}} \boldsymbol{Q}(\boldsymbol{y}^{k'} - \boldsymbol{y}^k)\delta_{k'} \\
& \boldsymbol{A}^\top \boldsymbol{\kappa} + \displaystyle\sum_{s\in\mathcal{S}}\sum_{t\in\mathcal{T}} \Phi_{st}\boldsymbol{g}_{st} = \boldsymbol{0}.
\end{array}
$$

Proof. The maximum probability with which policy $\boldsymbol{y}^k$ is chosen under any probability distribution $\mathbb{P} \in \mathcal{P}$ is given by the optimal value of the following moment problem.

$$
\begin{array}{ll}
\text{maximize} & \displaystyle\int_\Xi \mathbb{I}\left[\boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k \leq \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^{k'} \quad \forall k' \in \mathcal{K}\right] \mathrm{d}\mu(\boldsymbol{\xi}) \\
\text{subject to} & \mu \in \mathcal{M}_+(\mathbb{R}^Q) \\
& \displaystyle\int_\Xi \mathrm{d}\mu(\boldsymbol{\xi}) = 1 \\
& \displaystyle\int_\Xi \boldsymbol{g}(\boldsymbol{\xi})\,\mathrm{d}\mu(\boldsymbol{\xi}) \leq \boldsymbol{c}.
\end{array}
$$

Strong duality is guaranteed by Proposition 3.4 in [10], which is applicable since the ambiguity set $\mathcal{P}$ contains a Slater point. Thus, the dual problem

$$
\begin{array}{ll}
\text{minimize} & \alpha + \boldsymbol{c}^\top \boldsymbol{\beta} \\
\text{subject to} & \alpha \in \mathbb{R}, \quad \boldsymbol{\beta} \in \mathbb{R}^S_+ \\
& \alpha + \boldsymbol{g}(\boldsymbol{\xi})^\top \boldsymbol{\beta} \geq \mathbb{I}\left[\boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k \leq \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^{k'} \quad \forall k' \in \mathcal{K}\right] \quad \forall \boldsymbol{\xi} \in \Xi
\end{array}
$$

attains the same optimal value. We can replace the semi-infinite constraint in this problem with the two constraints

$$
\begin{aligned}
&\alpha + \boldsymbol{g}(\boldsymbol{\xi})^\top \boldsymbol{\beta} \geq 1 \quad \forall \boldsymbol{\xi} \in \Xi : \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k \leq \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^{k'} \quad \forall k' \in \mathcal{K} \\
&\alpha + \boldsymbol{g}(\boldsymbol{\xi})^\top \boldsymbol{\beta} \geq 0 \quad \forall \boldsymbol{\xi} \in \Xi.
\end{aligned} \tag{6}
$$

In the following, we focus on the reformulation of the first constraint; the second constraint can be dealt with analogously. The first constraint is satisfied if and only if the optimal value of

$$
\begin{array}{ll}
\text{minimize} & \alpha + \boldsymbol{\beta}^\top \boldsymbol{g}(\boldsymbol{\xi}) \\
\text{subject to} & \boldsymbol{\xi} \in \mathbb{R}^Q \\
& \boldsymbol{A}\boldsymbol{\xi} \leq \boldsymbol{b} \\
& \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k \leq \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^{k'} \quad \forall k' \in \mathcal{K}
\end{array}
$$

is greater than or equal to 1. By employing an epigraph reformulation, we can replace $\boldsymbol{g}$ with its definition to obtain the following equivalent problem:

$$
\begin{array}{ll}
\text{minimize} & \alpha + \boldsymbol{\beta}^\top \boldsymbol{\eta} \\
\text{subject to} & \boldsymbol{\xi} \in \mathbb{R}^Q, \quad \boldsymbol{\eta} \in \mathbb{R}^S \\
& \boldsymbol{A}\boldsymbol{\xi} \leq \boldsymbol{b} \\
& \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k \leq \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^{k'} \quad \forall k' \in \mathcal{K} \\
& \eta_s \geq \boldsymbol{g}_{st}^\top \boldsymbol{\xi} \forall s \in \mathcal{S}, \quad \forall t \in \mathcal{T}.
\end{array}
$$

This reformulation exploits the fact that $\boldsymbol{\beta} \geq \mathbf{0}$. The dual problem

$$\begin{array}{ll} \text{maximize} & \alpha - \boldsymbol{b}^\top \boldsymbol{\gamma} \\ \text{subject to} & \boldsymbol{\gamma} \in \mathbb{R}_+^R, \quad \boldsymbol{\delta} \in \mathbb{R}_+^K, \quad \boldsymbol{\Lambda} \in \mathbb{R}_+^{S\times T} \\ & \boldsymbol{A}^\top \boldsymbol{\gamma} + \sum_{s\in\mathcal{S}} \sum_{t\in\mathcal{T}} \Lambda_{st} \boldsymbol{g}_{st} = \sum_{k'\in\mathcal{K}} \boldsymbol{Q}(\boldsymbol{y}^{k'} - \boldsymbol{y}^k)\delta_{k'} \\ & \boldsymbol{\Lambda}\mathbf{e} = \boldsymbol{\beta} \end{array} \tag{7}$$

is feasible, which implies that strong duality holds. We conclude that the first constraint in (6) is satisfied if and only if there exists $(\boldsymbol{\gamma}, \boldsymbol{\delta}, \boldsymbol{\Lambda})$ feasible in (7) for which $\alpha - \boldsymbol{b}^\top \boldsymbol{\gamma} \geq 1$. A similar reformulation can be derived for the second constraint in (6), which concludes the proof. □

We close the section with the following result about the minimum probability with which a second-stage policy $\boldsymbol{y}^k$ is chosen under any distribution $\mathbb{P} \in \mathcal{P}$.

Proposition 2. *Let $(\boldsymbol{x}, \{\boldsymbol{y}^k\}_k)$ be a feasible decision in problem $(\mathcal{DPO}_K)$ with the ambiguity set (1) and the disutility function (2). For $k \in \mathcal{K}$, let $\mathcal{K}(k)$ be the subset of indices $k' \in \mathcal{K}$ for which there is a parameter realization $\boldsymbol{\xi} \in \Xi$ such that $\boldsymbol{\xi}^\top \boldsymbol{Q}\,(\boldsymbol{y}^k - \boldsymbol{y}^{k'}) > 0$. Then, the minimum probability with which policy $\boldsymbol{y}^k$, $k \in \mathcal{K}$, is chosen under any probability distribution $\mathbb{P} \in \mathcal{P}$ is given by the optimal value of the LP*

$$\begin{array}{ll} \textit{maximize} & \alpha - \boldsymbol{c}^\top \boldsymbol{\beta} \\ \textit{subject to} & \alpha \in \mathbb{R}, \quad \boldsymbol{\beta} \in \mathbb{R}_+^S, \quad \boldsymbol{\gamma} \in \mathbb{R}_+^R, \quad \boldsymbol{\Lambda} \in \mathbb{R}_+^{S\times T} \\ & \left.\begin{array}{l} \boldsymbol{\kappa}^{k'} \in \mathbb{R}_+^R, \quad \pi_{k'} \in \mathbb{R}_+, \quad \boldsymbol{\Phi}^{k'} \in \mathbb{R}_+^{S\times T} \\ \boldsymbol{b}^\top \boldsymbol{\gamma} + \alpha \leq 1, \quad \boldsymbol{b}^\top \boldsymbol{\kappa}^{k'} + \alpha \leq 0, \quad \boldsymbol{\Lambda}\mathbf{e} = \boldsymbol{\beta}, \quad \boldsymbol{\Phi}^{k'}\mathbf{e} = \boldsymbol{\beta} \\ \boldsymbol{A}^\top \boldsymbol{\kappa}^{k'} + \sum_{s\in\mathcal{S}} \sum_{t\in\mathcal{T}} \Phi_{st}^{k'} \boldsymbol{g}_{st} = \boldsymbol{Q}(\boldsymbol{y}^k - \boldsymbol{y}^{k'})\pi_{k'} \end{array}\right\} \quad \forall k' \in \mathcal{K}(k) \\ & \boldsymbol{A}^\top \boldsymbol{\gamma} + \sum_{s\in\mathcal{S}} \sum_{t\in\mathcal{T}} \Lambda_{st} \boldsymbol{g}_{st} = \mathbf{0}. \end{array}$$

Proof. The minimum probability with which policy $\boldsymbol{y}^k$ is chosen under any probability distribution $\mathbb{P} \in \mathcal{P}$ is given by the optimal value of the following moment problem.

$$\begin{array}{ll} \text{minimize} & \displaystyle\int_\Xi \mathbb{I}\left[\boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k \leq \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^{k'} \quad \forall k' \in \mathcal{K}\right] \mathrm{d}\mu(\boldsymbol{\xi}) \\ \text{subject to} & \mu \in \mathcal{M}_+(\mathbb{R}^Q) \\ & \displaystyle\int_\Xi \mathrm{d}\mu(\boldsymbol{\xi}) = 1 \\ & \displaystyle\int_\Xi \boldsymbol{g}(\boldsymbol{\xi})\,\mathrm{d}\mu(\boldsymbol{\xi}) \leq \boldsymbol{c}. \end{array}$$

Strong duality is guaranteed by Proposition 3.4 in [10], which is applicable since the ambiguity set $\mathcal{P}$ contains a Slater point. Thus, the dual problem

$$\begin{array}{ll} \text{maximize} & \alpha - \boldsymbol{c}^\top \boldsymbol{\beta} \\ \text{subject to} & \alpha \in \mathbb{R}, \quad \boldsymbol{\beta} \in \mathbb{R}_+^S \\ & \alpha - \boldsymbol{g}(\boldsymbol{\xi})^\top \boldsymbol{\beta} \leq \mathbb{I}\left[\boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k \leq \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^{k'} \quad \forall k' \in \mathcal{K}\right] \quad \forall \boldsymbol{\xi} \in \Xi \end{array}$$

attains the same optimal value. We can replace the semi-infinite constraint in this problem with the two constraints

$$\begin{array}{l} \alpha - \boldsymbol{g}(\boldsymbol{\xi})^\top \boldsymbol{\beta} \leq 1 \quad \forall \boldsymbol{\xi} \in \Xi \\ \alpha - \boldsymbol{g}(\boldsymbol{\xi})^\top \boldsymbol{\beta} \leq 0 \quad \forall \boldsymbol{\xi} \in \Xi \setminus \Xi_k, \end{array} \tag{8}$$

where $\Xi_k = \{\boldsymbol{\xi} \in \Xi : \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k \leq \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^{k'} \ \ \forall k' \in \mathcal{K}\}$. Note that

$$\Xi \setminus \Xi_k = \bigcup_{k'\in\mathcal{K}} \left\{\boldsymbol{\xi} \in \Xi : \boldsymbol{\xi}^\top \boldsymbol{Q}\,(\boldsymbol{y}^k - \boldsymbol{y}^{k'}) > 0\right\}.$$

Since the constraints in (8) are continuous in $\boldsymbol{\xi}$, we can replace the set $\Xi \setminus \Xi_k$ in (8) with

$$\begin{aligned} \mathrm{cl}\,(\Xi \setminus \Xi_k) &= \mathrm{cl} \bigcup_{k'\in\mathcal{K}} \left\{\boldsymbol{\xi} \in \Xi : \boldsymbol{\xi}^\top \boldsymbol{Q}\,(\boldsymbol{y}^k - \boldsymbol{y}^{k'}) > 0\right\} \\ &= \bigcup_{k'\in\mathcal{K}} \mathrm{cl}\left\{\boldsymbol{\xi} \in \Xi : \boldsymbol{\xi}^\top \boldsymbol{Q}\,(\boldsymbol{y}^k - \boldsymbol{y}^{k'}) > 0\right\} \\ &= \bigcup_{k'\in\mathcal{K}(k)} \left\{\boldsymbol{\xi} \in \Xi : \boldsymbol{\xi}^\top \boldsymbol{Q}\,(\boldsymbol{y}^k - \boldsymbol{y}^{k'}) \geq 0\right\}. \end{aligned}$$

Here, the second identity holds because $\mathcal{K}$ is finite, which implies that the union and the closure operators commute. The last identity follows from the fact that the k'-th set in the union is non-empty if and only if $k' \in \mathcal{K}(k)$. Hence, the semi-infinite constraints in (8) are satisfied if and only if the constraints

$$\begin{array}{l} \alpha - \boldsymbol{g}(\boldsymbol{\xi})^\top \boldsymbol{\beta} \leq 1 \quad \forall \boldsymbol{\xi} \in \Xi \\ \alpha - \boldsymbol{g}(\boldsymbol{\xi})^\top \boldsymbol{\beta} \leq 0 \quad \forall k' \in \mathcal{K}(k), \quad \forall \boldsymbol{\xi} \in \Xi : \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^k \geq \boldsymbol{\xi}^\top \boldsymbol{Q}\,\boldsymbol{y}^{k'} \end{array}$$

are satisfied. The result now follows if we apply similar reformulations to these semi-infinite constraints as in the proof of Proposition 1. □

4. Numerical experiments

We apply the K-adaptability approximation to a stylized two-stage version of the vertex packing problem. The optimization problems in this section are solved using the YALMIP modeling language [9] and Gurobi Optimizer 5.6 [6] with the default settings and a time limit of 7200 s.

Consider an undirected, node-weighted graph $G = (V, E, \boldsymbol{\xi})$ with nodes $V = \{1, \ldots, N\}$, edges $E \subseteq \{\{i, j\} : i, j \in V\}$ and weights $\xi_i \in \mathbb{R}_+$, $i \in V$. The vertex packing problem asks for a packing P, that is, a subset of nodes $P \subseteq V$ such that no pair of nodes in P is connected by an edge, that maximizes the sum of node weights $\sum_{i\in P} \xi_i$. We consider a two-stage distributionally robust variant of the problem where the node weights are modeled as a random vector $\tilde{\boldsymbol{\xi}}$ that is governed by an unknown probability distribution. The goal is to pre-commit to a subset of nodes P_1 before observing $\tilde{\boldsymbol{\xi}}$ and to complete P_1 to a packing P_2 after observing $\tilde{\boldsymbol{\xi}}$ so that $|P_1 \setminus P_2| + |P_2 \setminus P_1| \leq B$, where B denotes the budget of change, and the packing P_2 maximizes the sum of node weights.

Using the worst-case conditional value-at-risk as a risk measure, the problem can be formulated as a distributionally robust two-stage binary program with objective uncertainty:

$$\begin{array}{ll} \text{maximize} & \displaystyle\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{-CVaR}_\beta \left[\max_{\boldsymbol{y}\in\{0,1\}^N} \left\{ \tilde{\boldsymbol{\xi}}^\top \boldsymbol{y} : \|\boldsymbol{x} - \boldsymbol{y}\|_1 \leq B, \right.\right. \\ & \qquad \left.\left. y_i + y_j \leq 1 \quad \forall \{i, j\} \in E \right\} \right] \\ \text{subject to} & \boldsymbol{x} \in \{0, 1\}^N. \end{array}$$

In this formulation, the decisions $\boldsymbol{x}$ and $\boldsymbol{y}$ represent indicator functions for the sets P_1 and P_2, respectively, that is, we have $x_i = \mathbb{I}[i \in P_1]$ and $y_i = \mathbb{I}[i \in P_2]$, $i \in V$, and $\mathbb{P}$-CVaR$_\beta$ denotes the conditional value-at-risk at level β under the distribution $\mathbb{P} \in \mathcal{P}$.

For our numerical experiments, we generate random graphs with $N \in \{10, \ldots, 50\}$ nodes and $|E| = 5\,|V|$ edges. Note that the number of edges scales linearly with the number of nodes, which implies that the expected size of a maximum cardinality packing (*i.e.*, the maximum packing for unit node weights) grows proportionally with N. This ensures that the optimal packings do not degenerate to a trivial solution (*i.e.*, none or almost all nodes)

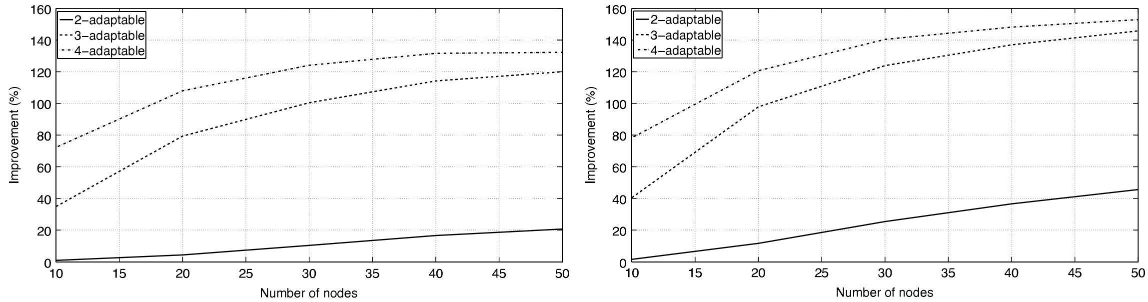

Fig. 1. Improvement of the best 2-, 3- and 4-adaptable solutions determined within the set time limit over the static solutions for the vertex packing problem with $B = 0.2N$ (left) and $B = 0.4N$ (right). The figures show the improvements for problems with $N = 10, 20, \ldots, 50$ nodes as averages over 100 instances.

Table 1
Summary of the results for the vertex packing problem. Each entry in the table documents the percentage of instances solved within the time limit, the average solution time for the instances solved within the time limit and the average optimality gap for the instances not solved to optimality. All results are averaged over 100 instances.

	K	Number of nodes N				
		10	20	30	40	50
$B = 0.2N$	2	100%/<1s/0%	100%/<1s/0%	100%/1s/0%	100%/17s/0%	100%/4m:12s/0%
	3	100%/<1s/0%	100%/<1s/0%	100%/12s/0%	100%/3m:14s/0%	74%/44m:39s/22.35%
	4	100%/<1s/0%	100%/2s/0%	100%/2m:11s/0%	65%/54m:27s/13.52%	0%/-/22.38%
$B = 0.4N$	2	100%/<1s/0%	100%/<1s/0%	100%/1s/0%	100%/9s/0%	100%/1m:13s/0%
	3	100%/<1s/0%	100%/<1s/0%	100%/3s/0%	100%/49s/0%	100%/12m:24s/0%
	4	100%/<1s/0%	100%/1s/0%	100%/36s/0%	100%/17m:33s/0%	10%/29m:05s/21.68%

Table 2
Minimum and maximum probabilities with which the policies $\boldsymbol{y}^k$, $k \in \mathcal{K}$, are chosen in the vertex packing problem. The probabilities are averaged over 100 instances, and the policies are ordered according to increasing maximum probabilities.

K	Number of nodes N		
	10	30	50
2	[0 0.32], [0.62 1]	[0 0.65], [0.32 1]	[0 0.92], [0.07 1]
3	[0 0.36], [0 0.66], [0.21 1]	[0 0.60], [0 0.88], [0.05 1]	[0 0.75], [0 0.91], [0.02 1]
4	[0 0.27], [0 0.46], [0 0.74], [0.11 1]	[0 0.34], [0 0.62], [0 0.86], [0.04 1]	[0 0.35], [0 0.75], [0 0.91], [0.02 1]

as N increases. We choose the following ambiguity set for the node weights.

$$\mathcal{P} = \left\{ \mathbb{P} \in \mathcal{M}_+(\mathbb{R}^N) : \mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{\xi}}\right] = \boldsymbol{\mu},\ \mathbb{E}_{\mathbb{P}}\left[\left|\tilde{\boldsymbol{\xi}} - \boldsymbol{\mu}\right|\right] \le 0.15\boldsymbol{\mu}, \right.$$
$$\left. \mathbb{E}_{\mathbb{P}}\left[\left|\mathbf{e}^\top(\tilde{\boldsymbol{\xi}} - \boldsymbol{\mu})\right|\right] \le 0.15 N^{-1/2}\mathbf{e}^\top\boldsymbol{\mu} \right\}.$$

Here, the average node weights $\boldsymbol{\mu}$ are chosen uniformly at random from the interval $[0, 10]^N$. The second condition in $\mathcal{P}$ imposes upper bounds on the mean absolute deviations of the individual node weights. The third condition is inspired by the central limit theorem, and it imposes an upper bound on the cumulative deviation of the node weights from their expected values [1, Section 2].

The K-adaptability formulation (4) corresponding to the vertex packing problem has $\mathcal{O}(KN)$ binary variables. Table 1 shows that most of the problem instances can be solved to optimality within the set time limit. The table also shows that problems with a smaller budget of change B are harder to solve. An investigation of the reports generated by Gurobi reveals that for small values of B, the solver requires a long time to determine good feasible solutions. Finally, the table presents estimates for the optimality gaps of those instances that could not be solved to optimality. These estimates are derived from a progressive (upper) bound on the optimal value of problem $(\mathcal{DPO})$ that results from disregarding the integrality requirement for the second-stage decisions $\boldsymbol{y}$, applying the classical min–max theorem to exchange the order of the maximization problem over $\mathbb{P} \in \mathcal{P}$ and the minimization problem over $\boldsymbol{y} \in \mathcal{Y}$ and subsequently dualizing the maximization problem.

Fig. 1 shows the improvement of the 2-, 3- and 4-adaptable solutions over the static solutions that take all decisions here-and-now. The figure reveals that the improvement increases with the number of nodes N, but it saturates as N increases. Indeed, one can show that for the considered problem class, the outperformance of the fully adaptable solutions to problem $(\mathcal{DPO})$ over the static solutions is bounded by a constant. The figure also shows that the improvement of the adaptable solutions increases with the budget of change B. This is intuitive as higher values of B give more flexibility to modify the vertex packing in the second stage when the node weights are known.

Table 2 shows the minimum and maximum probabilities with which the policies $\boldsymbol{y}^k$, $k \in \mathcal{K}$, are chosen. Note that in all instances, one policy has a maximum probability of 1 while all other policies have a minimum probability of 0. This is due to the fact that the ambiguity set $\mathcal{P}$ contains the Dirac distribution that places all probability mass on the expected value $\boldsymbol{\mu}$. For this distribution, there is always an optimal solution which selects a single policy with probability 1.

Acknowledgments

The authors are grateful to Rüdiger Schultz and the anonymous referee for their thoughtful comments which substantially improved the paper. This research was supported by the Swiss National Science Foundation grant BSCGI0_157733 and the EPSRC grants EP/M028240/1 and EP/M027856/1.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.orl.2015.10.006.

References

[1] C. Bandi, D. Bertsimas, Tractable stochastic analysis in high dimensions via robust optimization, Math. Program. B 134 (1) (2012) 23–70.
[2] A. Ben-Tal, M. Teboulle, An old-new concept of convex risk measures: The optimized certainty equivalent, Math. Finance 17 (3) (2007) 449–476.
[3] D. Bertsimas, C. Caramanis, Finite adaptibility in multistage linear optimization, IEEE Trans. Automat. Control 55 (12) (2010) 2751–2766.
[4] L.G. Epstein, A definition of uncertainty aversion, Rev. Econom. Stud. 66 (3) (1999) 579–608.
[5] I. Gilboa, D. Schmeidler, Maxmin expected utility with non-unique prior, J. Math. Econom. 18 (2) (1989) 141–153.
[6] Gurobi Optimization, Inc. Gurobi optimizer reference manual, 2014.
[7] G.A. Hanasusanto, D. Kuhn, W. Wiesemann, K-adaptability in two-stage robust binary programming, Oper. Res. 63 (4) (2015) 877–891.
[8] X. Li, K. Natarajan, C.-P. Teo, Z. Zheng, Distributionally robust mixed integer linear programs: Persistency model with applications, European J. Oper. Res. 233 (3) (2014) 459–473.
[9] J. Löfberg, YALMIP: A toolbox for modeling and optimization in MATLAB, in: IEEE International Symposium on Computer Aided Control Systems Design, 2004, pp. 284–289.
[10] A. Shapiro, On duality theory of conic linear problems, in: Semi-Infinite Programming, Kluwer Academic Publishers, 2001, pp. 135–165. (Chapter 7).
[11] W. Wiesemann, D. Kuhn, M. Sim, Distributionally robust convex optimization, Oper. Res. 62 (6) (2014) 1358–1376.