Minkowski Centers via Robust Optimization

Computation and Applications

den Hertog, D.; Pauphilet, J.; Yahya Soali, M.

DOI
10.1287/opre.2023.2448

Publication date
2024

Document Version
Final published version

Published in
Operations Research

Citation for published version (APA):
den Hertog, D., Pauphilet, J., & Yahya Soali, M. (2024). Minkowski Centers via Robust Optimization: Computation and Applications. *Operations Research*, *72*(5), 2135–2152. https://doi.org/10.1287/opre.2023.2448

https://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Vol. 72, No. 5, September–October 2024, pp. 2135–2152
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Methods

Minkowski Centers via Robust Optimization: Computation and Applications

Dick den Hertog,[a] **Jean Pauphilet,**[b,*] **Mohamed Yahya Soali**[c]

[a] Amsterdam Business School, University of Amsterdam, 1001 NL Amsterdam, Netherlands; [b] Management Science and Operations, London Business School, London NW1 4SA, United Kingdom; [c] Industrial Engineering and Operations Research, Columbia University, New York, New York 10027

*Corresponding author

Contact: d.denhertog@uva.nl, https://orcid.org/0000-0002-1829-855X (DdH); jpauphilet@london.edu, https://orcid.org/0000-0001-6352-0984 (JP); ms6352@columbia.edu, https://orcid.org/0009-0003-4150-343X (MYS)

Received: January 19, 2022
Revised: May 9, 2022
Accepted: March 16, 2023
Published Online in Articles in Advance:
April 19, 2023

Area of Review: Optimization

https://doi.org/10.1287/opre.2023.2448

Abstract. Centers of convex sets are geometric objects that have received extensive attention in the mathematical and optimization literature, both from a theoretical and practical standpoint. For instance, they serve as initialization points for many algorithms such as interior-point, hit-and-run, or cutting-planes methods. First, we observe that computing a Minkowski center of a convex set can be formulated as the solution of a robust optimization problem. As such, we can derive tractable formulations for computing Minkowski centers of polyhedra and convex hulls. Computationally, we illustrate that using Minkowski centers, instead of analytic or Chebyshev centers, improves the convergence of hit-and-run and cutting-plane algorithms. We also provide efficient numerical strategies for computing centers of the projection of polyhedra and of the intersection of two ellipsoids.

Supplemental Material: The online appendix is available at https://doi.org/10.1287/opre.2023.2448.

Keywords: Minkowski center • geometry • robust optimization

1. Introduction

Centers of convex sets have played a fundamental role in all areas of applied mathematics, especially in mathematical programming. Historically, the development of efficient linear optimization algorithms is deeply connected with the definition and computation of the center of a polytope. The ellipsoid algorithm solves linear optimization problems by constructing a volume-decreasing sequence of circumscribed ellipsoids (see Bland et al. (1981) for a review). This algorithm, proposed by Yudin and Nemirovskii (1976), sparked interest on computing the minimum-volume circumscribed ellipsoid of a polytope or a generic convex body (Todd 1982). Alternatively, Tarasov et al. (1988) proposed the inscribed ellipsoid method, where, at each step, one needs to compute numerically an approximation to the maximum-volume inscribed ellipsoid of a polytope. Because optimizing over ellipsoids is easier than over a general convex set, the minimum-volume circumscribed and the maximum-volume inscribed ellipsoids provide inner and outer approximations that can also be used to approximately solve optimization problems over a convex set. Karmarkar (1984) introduced another polynomial-time interior point algorithm for linear optimization. At each iteration, the algorithm constructs a mapping of the feasible set into a standard simplex and associates the current iterate with “the center” of the simplex, without providing a generic definition of center. Analytic centers were first defined and analyzed by Huard (1967) and further analyzed by Sonnevend (1986), Renegar (1988), and Jarre (1989). Modern interior point methods rely mostly on the analytic center due to its computational benefits (Roos et al. 2005).

Besides extremal ellipsoids problems, the Minkowski measure of symmetry (Minkowski 1911) has been used to support another geometric definition of the center of a convex set. Yet, compared with the other definitions, the Minkowski center has driven mostly theoretical interest and there is, to the best of our knowledge, no computational evidence on the tractability and the practical benefits of Minkowski centers. The present paper provides a first answer to these questions.

1.1. Contributions and Structure

This paper shows that Minkowski centers of a convex set are solutions of a robust optimization problem. Under this robust lens, we provide computationally tractable reformulations or approximations for a series of sets including polyhedra and projections of polyhedra. We can also derive known and new analytic expressions for the symmetry measure of simple sets by analyzing the optimization formulation directly, instead of the geometry of the set. We demonstrate numerically that Minkowski centers are viable alternatives to other centers, such as Chebyshev or analytic centers, and can speed up convergence of numerical algorithms.

After presenting the notations in Section 1.2 and the existing literature on centers of convex bodies in Section 1.3, the rest of the paper is organized as follows:

- We introduce Minkowski centers in Section 2 and connect them with existing definitions of centers. In particular, we show that Minkowski centers are special cases of Helly centers, like the centroid, the John or the volumetric center. We then derive a robust optimization formulation for computing Minkowksi centers of a convex set (Proposition 3). Under this lens, we derive tractable reformulations of this optimization problem for polyhedra and the convex hull of a finite number of points, and provide known and new analytical bounds in simple cases.
- Numerically, analytic centers are widely used in the initialization of many algorithms despite that they are analytical and not geometric. We demonstrate empirically in Section 3 that using Minkowski centers instead can provide substantial benefit in terms of algorithmic convergence, using the hit-and-run and the cutting-plane algorithms as illustrating examples.
- In Section 4, we consider the case of a convex set defined as the projection of a polyhedron. We show that computing Minkowski centers for such a set is equivalent to solving an *adjustable* robust optimization problem. We propose an approximation based on linear decision rules and evaluate its practical relevance on numerical simulations.
- Finally, in Section 5, we propose a numerical strategy for approximating a Minkowski center of the intersection of two ellipsoids. Our algorithm relies on a second-order cone (SOC) relaxation and bisection search. We also provide a (numerically verifiable) condition under which our approximation is tight, together with a constant factor approximation bound for our approach. We also discuss the extension to intersection of $m > 2$ ellipsoids.

1.2. Notations

We use nonbold face characters (x or λ) to denote scalars, lowercase bold faced characters ($\boldsymbol{x}$) to denote vectors, uppercase bold faced characters ($\boldsymbol{X}$) to denote matrices, and calligraphic characters such as $\mathcal{X}$ to denote sets. We let $\boldsymbol{e}$ (respectively, $\boldsymbol{0}$) denote the vector of all ones (respectively, zeros), with dimension implied by the context. We denote by $\boldsymbol{e}_i$ the unit vector with one at the ith coordinate and zero elsewhere; $\mathbb{R}$, $\mathbb{R}_+$, and $\mathbb{N}$ denote the set of real numbers, nonnegative real numbers, and nonnegative integers, respectively. For a positive integer $n \in \mathbb{N}$, we define $[n] := \{1, \ldots, n\}$. Given two n-dimensional vectors $\boldsymbol{x}, \boldsymbol{y}$, we use the notation $\boldsymbol{x}^\top \boldsymbol{y}$ for the inner product of $\boldsymbol{x}$ and $\boldsymbol{y}$, $\boldsymbol{x}^\top \boldsymbol{y} := \sum_{i \in [n]} x_i y_i$, and $\|\boldsymbol{x}\|$ for the Euclidean norm of $\boldsymbol{x}$, $\|\boldsymbol{x}\| := \sqrt{\boldsymbol{x}^\top \boldsymbol{x}}$. For $p \geq 1$, the p-norm of $\boldsymbol{x}$ is defined as $\|\boldsymbol{x}\|_p = (\sum_{i \in [n]} |x_i|^p)^{1/p}$, so that $\|\boldsymbol{x}\| = \|\boldsymbol{x}\|_2$.

1.3. Literature Review

In this section, we present the various definitions of centers that have been proposed in the applied mathematics literature.

Historically, the first definition of a center is the center of mass (or barycenter), used primarily in physics and motion geometry (Schwartz and Sharir 1988). The center of mass of a set is defined as the weighted arithmetic mean position of all its points, that is,

$$\frac{1}{\int_{\mathbb{R}^n} \mu(\boldsymbol{x})\, d\boldsymbol{x}} \int_{\mathbb{R}^n} \boldsymbol{x} \mu(\boldsymbol{x})\, d\boldsymbol{x},$$

where $\mu(\cdot)$ is a given mass density function over the set of interest $\mathcal{C}$. When μ is uniform, the center of mass is also called the centroid. In particular, the centroid of a finite number of m points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$, is $\frac{1}{m} \sum_{i \in [m]} \boldsymbol{x}_i$. Computing the centroid of a polytope is $\#P-$hard (Rademacher 2007), but it can be efficiently approximated via random sampling. In data science, the notion of centroid is the building block of the k-means clustering algorithm (Kanungo et al. 2002).

For convex sets, an important geometrical definition of a center is the notion of Helly center.

Definition 1. For a convex set $\mathcal{C}$, we say that $\boldsymbol{x}_H \in \mathcal{C}$ is a Helly center if for any chord $[\boldsymbol{u}, \boldsymbol{v}]$ passing through $\boldsymbol{x}_H$, we have

$$\frac{1}{n+1} \leq \frac{\|\boldsymbol{x}_H - \boldsymbol{u}\|}{\|\boldsymbol{v} - \boldsymbol{u}\|} \leq \frac{n}{n+1}.$$

Intuitively, the ratio of distances in Definition 1 measures how close $\boldsymbol{x}_H$ is to $\boldsymbol{u}$ and $\boldsymbol{v}$ (in relative terms). Helly centers are thus points that are sufficiently far away from the boundary. Klee (1963) proves that any convex compact body of $\mathbb{R}^n$ admits a Helly center, as a consequence of Helly's theorem. However, it is in general not unique. For instance, Barnes and Moretti (2005) prove that an ellipsoid admits an infinity of Helly centers (theorem 2.5).

Another class of centers encompasses centers defined via extremal ellipsoids (Güler and Gürtuna 2012). For instance, the center of the minimum-volume ellipsoid that contains a set $\mathcal{C}$ is referred to as the John (or Löwner-John) center of $\mathcal{C}$. The John center is well defined for convex bodies and unique (John 1948). Alternatively, the center of the maximum-volume ellipsoid contained in $\mathcal{C}$ is called the volumetric center of $\mathcal{C}$ (Vaidya 1996). However, even for polyhedra, known algorithms for finding the maximum-volume ellipsoid and its center require solving a semidefinite optimization problem (Boyd and Vandenberghe 2004, section 8.4.2).

Recently, Zhen and den Hertog (2018) use Fourier-Motzkin decomposition and adjustable robust optimization techniques to approximate it in a tractable fashion for projection of polyhedra of the form $\{x : \exists z \text{s.t.} A_x x + A_z z \le b\}$. If we further restrict our attention to isotropic ellipsoids, the center of a maximum-volume ball enclosed in $\mathcal{C}$ is often called a Chebyshev center of $\mathcal{C}$.

Definition 2. A Chebyshev center of a convex set $\mathcal{C}$ is a solution of the optimization problem

$$\max_{x \in \mathcal{C}, r \in \mathbb{R}_+} r \text{ s.t. } \mathcal{B}(x, r) \subseteq \mathcal{C},$$

where $\mathcal{B}(x, r)$ denotes the ball centered in x and of radius r (in Euclidean norm).

Chebyschev centers of a polyhedron can be computed by solving a linear optimization problem (Boyd and Vandenberghe 2004, sections 4.3.1 and 8.5.1). We should note that the definition of Chebyshev centers is ambiguous in the literature and that some authors, for example, Eldar et al. (2008) and Xia et al. (2021), alternatively define them as the centers of the minimum-volume circumscribed ball. To avoid any ambiguity, we will consistently use Definition 2 in this paper. Finally, the main limitation of centers defined via extremal ellipsoids is that they require the convex set $\mathcal{C}$ to be fully dimensional (or they require to restrict our attention to ellipsoids in the affine hull of $\mathcal{C}$).

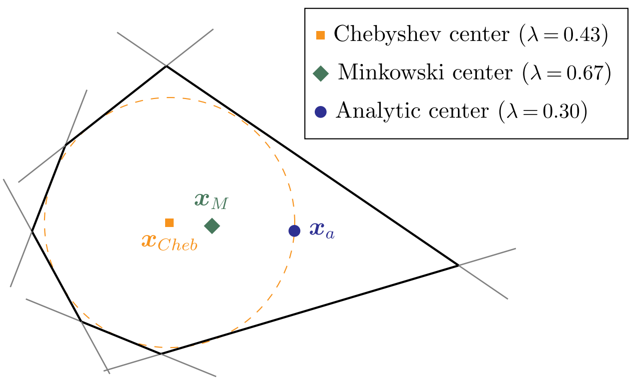
Finally, the most used definition of a center in optimization is certainly the analytic center.

Definition 3. The analytic center of the convex set $\mathcal{C} = \{x : Ax = b;\ f_i(x) \le 0,\ \ \forall i \in [m]\}$ is the solution of the optimization problem

$$\max_{x} \sum_{i} \log(-f_i(x)) \text{ s.t. } Ax = b.$$

The previous maximization problem aims to find a strictly feasible point $x \in \mathcal{C}$ with the largest sum of log-slacks. When $\mathcal{C}$ is bounded, the logarithmic barrier terms $\log(-f_i(x))$ are bounded above, the optimization problem is well defined. For polyhedra, the analytic center, when it exists, is unique. Being defined as the solutions of a convex optimization problem, analytic centers can be computed in a tractable fashion. However, a major deficiency of this definition is that it is not a geometric concept but rather depends on the analytical description of the set $\mathcal{C}$. For example, the analytic center of the n-dimensional standard simplex defined as $\{x \in \mathbb{R}^n_+ : e^\top x = 1\}$ is the vector $\frac{1}{n}e$, but the analytic center of the geometrically equivalent set $\{x \in \mathbb{R}^n_+ : e^\top x \le 1, e^\top x \ge 1\}$ does not exist. Similarly, duplicating or adding redundant constraints in the description of $\mathcal{C}$ pushes the analytic center arbitrarily close to the boundary (Caron et al. 2002). Yet, the analytic center remains very popular and a cornerstone in optimization algorithms.

Figure 1. (Color online) Example of the Chebyshev (x_{Cheb}), Analytic (x_a), and Minkowski (x_M) Center of a Two-Dimensional Polyhedron



From a geometric perspective, analytic or Chebyshev centers are not necessarily Helly centers, as formally stated here.

Proposition 1. *Chebyshev or analytic centers are not necessarily Helly centers.*

The proof of Proposition 1 relies on counter-examples provided in Online Appendix EC.1. This negative result motivates our investigation into alternative definitions of centers. In the rest of this paper, we study Minkowski centers, which are special cases of Helly center—as we will prove in the next section. For illustration purposes, Figure 1 represents the Chebyshev, analytic, and Minkowski centers of a two-dimensional polytope. The presence of multiple constraints on the left pushes the analytic center to the right, while the pointy shape of the polyhedron settles the Chebyshev center on the left-hand side. The Minkowski center, on the other hand, appears more "central."

2. Minkowski Center and Robust Optimization Formulation

In this paper, we study Minkowski centers of a closed, bounded, and convex body $\mathcal{C} \subseteq \mathbb{R}^n$. Minkowski centers are related to the notion of symmetry of the set. Let us first define the symmetry of $\mathcal{C}$ with respect to a point $x \in \mathcal{C}$ as

$$\text{sym}(x, \mathcal{C}) := \max_{\lambda \ge 0} \lambda \text{ s.t. } \ x + \lambda(x - y) \in \mathcal{C},\ \ \forall y \in \mathcal{C}.$$

This measure of symmetry, initially proposed by Minkowski, intuitively states that $\text{sym}(x, \mathcal{C})$ is the largest scalar λ such that every point $y \in \mathcal{C}$ can be reflected through x by the factor λ and still lies in $\mathcal{C}$. Among other properties, we have $\text{sym}(x, \mathcal{C}) \le 1$. We refer to Belloni and Freund (2007) for an analysis of some fundamental properties of $\text{sym}(x, \mathcal{C})$. Then, a Minkowski center is defined as a point $x \in \mathcal{C}$ maximizing symmetry.

Definition 4. A point $x^\star$ is called a Minkowski center or symmetric point of $\mathcal{C}$ if $x^\star$ is a solution of the optimization problem $\max_{x \in \mathcal{C}} \operatorname{sym}(x, \mathcal{C})$. The optimal objective value, $\operatorname{sym}(\mathcal{C}) := \operatorname{sym}(x^\star, \mathcal{C})$, is called the symmetry of $\mathcal{C}$.

In particular, Minkowski centers are not necessarily unique (for instance, Minkowski centers of $\{x \in [0,1]^3 : x_1 + x_2 \le 1\}$ are all points of the form $(1/3, 1/3, t)$ for $t \in [1/3, 2/3]$) and the set $\mathcal{C}$ is symmetric with respect to some x_0 (i.e., $\forall x \in \mathcal{C}, 2x_0 - x \in \mathcal{C}$) if and only if $\operatorname{sym}(\mathcal{C}) = 1$.

2.1. Minkowski Centers are Helly Centers

Here, we connect the definition of Minkowski centers with the notion of Helly centers. We first provide a sufficient condition for a point x to be a Helly center.

Proposition 2. *If $x \in \mathcal{C}$ satisfies $\frac{1}{n} \le \operatorname{sym}(x, \mathcal{C})$, then x is a Helly center of $\mathcal{C}$.*

The proof of Proposition 2 is provided in Online Appendix EC.2.1. From Proposition 2, we can prove that Minkowski centers, among other definitions of centers, are special cases of Helly centers.

Corollary 1. *If $\mathcal{C}$ is full dimensional,* (a) *the centroid, x_c,* (b) *the John center, x_J,* (c) *the volumetric center, x_v,* (d) *any Minkowski center, x_M, are Helly centers.*

Proof. We prove that the symmetry of $\mathcal{C}$ at each center is at least $1/n$. The results then follows from Proposition 2. (a) Hammer (1951) proved that $\operatorname{sym}(x_c, \mathcal{C}) \ge 1/n$. (b) The John center is the center of the minimum-volume circumscribed ellipsoid $\mathcal{E}, \mathcal{C} \subseteq \mathcal{E}$. John (1948) showed that $(1/n)\mathcal{E} \subseteq \mathcal{C}$ (Theorem 3), which implies that $\operatorname{sym}(x_J, \mathcal{C}) \ge 1/n$. (c) Similarly, the maximum-volume inscribed ellipsoid (whose center is the volumetric center) satisfies $\mathcal{E}' \subseteq \mathcal{C} \subseteq n\mathcal{E}'$ so $\operatorname{sym}(x_v, \mathcal{C}) \ge 1/n$. (d) Because a Minkowski center maximizes symmetry, $\operatorname{sym}(x_M, \mathcal{C}) \ge \operatorname{sym}(x_c, \mathcal{C}) \ge 1/n$. □

2.2. Robust Optimization Formulation

As a starting point to our analysis, we would like to emphasize that Minkowski centers are the solution of a robust optimization problem. From Definition 4, we can obviously write a Minkowski center as the solution of

$$\max_{x \in \mathcal{C}, \lambda \ge 0} \lambda \text{ s.t. } x + \lambda(x - y) \in \mathcal{C}, \ (\forall y \in \mathcal{C}), \tag{1}$$

which resembles a robust optimization problem where the set $\mathcal{C}$ defines both the uncertainty set and the constraints. However, the constraints involve products of decision variables, λx, hence might be nonconvex in (λ, x). Still, we can reformulate the previous optimization problem into one that is convex in its decision variables and uncertain parameters:

Proposition 3. *Assume that $\mathcal{C}$ can be described via linear equality constraints and m convex inequality constraints, that is, $\mathcal{C} = \{x \mid Ax = b;\ f_i(x) \le 0,\ \forall i \in [m]\}$. Consider $(w^\star, \lambda^\star)$, solutions of the following robust convex optimization problem:*

$$\begin{aligned} \max_{w, \lambda \ge 0} \lambda \text{ s.t. } \quad & Aw = (1+\lambda)b, \\ & (1+\lambda) f_i\left(\frac{w}{1+\lambda}\right) \le 0, \qquad \forall i \in [m], \\ & f_i(w - \lambda y) \le 0, \quad \forall y \in \mathcal{C}, \ \forall i \in [m]. \end{aligned} \tag{2}$$

Then, the point $x^\star := w^\star/(1+\lambda^\star)$ is a Minkowski center of $\mathcal{C}$ (with symmetry measure $\lambda^\star$).

Note that (2) is a robust optimization problem with a linear objective and constraints that are convex in the decision variables (w, λ) for a fixed y, and convex in the uncertain parameters y for a fixed (w, λ). This class of robust constraints can be approximated using the so-called reformulation-perspectification technique (Bertsimas et al. 2022).

Proof. Because $\lambda \ge 0, 1 + \lambda > 0$ and we can consider the bijective change of variable $(x, \lambda) \longmapsto (w, \lambda)$ with $w = (1+\lambda)x$. Problem (1) becomes

$$\begin{aligned} \max_{w, \lambda \ge 0} \lambda \text{ s.t. } \quad & \frac{w}{1+\lambda} \in \mathcal{C}, \\ & w - \lambda y \in \mathcal{C}, \ \forall y \in \mathcal{C}. \end{aligned} \tag{3}$$

To enforce $w/(1+\lambda) \in \mathcal{C}$, we need to impose

$$A \frac{w}{1+\lambda} = b \Leftrightarrow Aw = (1+\lambda)b,$$

$$f_i\left(\frac{w}{1+\lambda}\right) \le 0 \Leftrightarrow (1+\lambda) f_i\left(\frac{w}{1+\lambda}\right) \le 0, \quad \forall i \in [m].$$

Observe that $(x, t) \longmapsto t f_i(x/t)$ is the perspective function of f_i and is jointly convex in (x, t) over $\{(x, t) \in \mathbb{R}^n \times \mathbb{R}_+ : x/t \in \operatorname{dom}(f_i)\}$ (Boyd and Vandenberghe 2004, section 3.2.6). Therefore, all constraints are convex constraints in (w, λ).

Regarding the robust constraints, $w - \lambda y \in \mathcal{C},\ \forall y \in \mathcal{C}$, we consider the equality and inequality constraints separately. First, (w, λ) should satisfy $Aw - \lambda Ay = b,\ \forall y \in \mathcal{C}$. However, because $Ay = b$ for $y \in \mathcal{C}$, these constraints are equivalent to $Aw = (1+\lambda)b$, which are already enforced. Second, the inequality constraints can be written as $f_i(w - \lambda y) \le 0,\ \forall y \in \mathcal{C}$, which are robust constraints, convex in the decision variables and convex in the uncertain parameters y. □

This observation prompts us to investigate whether tools and techniques developed for robust optimization problems could be usefully and successfully applied to compute Minkowski centers of convex sets.

2.3. Tractable Reformulations for Polyhedra

In robust optimization, tractable reformulations are obtained when the robust constraints are concave in the uncertain parameter (Ben-Tal et al. 2015). When they are convex in the uncertain parameter, like in (2), even computing the worst-case scenario, that is, solving $\max_{y \in \mathcal{C}} f_i(\boldsymbol{w} - \lambda \boldsymbol{y})$ for a fixed $(\boldsymbol{w}, \lambda)$, is challenging. Accordingly, we first consider the easy case where the f_is are linear and hence both convex and concave.

First, we consider the case where $\mathcal{C}$ is described via linear constraints.

Proposition 4. *Consider* $\mathcal{C} = \{\boldsymbol{x} | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b};\ \boldsymbol{C}\boldsymbol{x} \le \boldsymbol{d}\}$, *where* $\boldsymbol{C} \in \mathbb{R}^{m \times n}$, $\boldsymbol{d} \in \mathbb{R}^m$. *For* $i \in [m]$, *define* $\delta_i := \min_{\boldsymbol{y} \in \mathcal{C}} \boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{y}$. *Then,* (2) *is equivalent to*

$$\max_{\boldsymbol{w}, \lambda \ge 0} \lambda \text{ s.t. } \quad \boldsymbol{A}\boldsymbol{w} = (1+\lambda)\boldsymbol{b}, \quad \boldsymbol{C}\boldsymbol{w} - \lambda \boldsymbol{\delta} \le \boldsymbol{d}. \tag{4}$$

Remark 1. Because we assume throughout the paper that the set $\mathcal{C}$ is bounded, we know that $\delta_i > -\infty$ for all $i \in [m]$. For an unbounded polyhedron $\mathcal{C}$, however, the behavior of Problem (4) depends on the recession cone of $\mathcal{C}$. Let us consider a decomposition of $\mathcal{C}$ into $\mathcal{C} = \mathcal{C}_0 + \mathcal{H}$, where $\mathcal{C}_0$ is a convex bounded set and $\mathcal{H}$ is the recession cone of $\mathcal{C}$. If $\mathcal{H}$ is not a subspace, then there exists an index i such that $\delta_i = -\infty$. Consequently, the set of feasible (and a fortiori optimal) solutions to (4) is $\mathcal{C} \times \{0\}$, and all points in $\mathcal{C}$ are Minkowski centers. If $\mathcal{H}$ is a subspace, the δ_is are all finite, but the set of Minkowski centers of $\mathcal{C}$ (and the set of optimal solutions to Problem 4) is invariant by translation by any $\boldsymbol{h} \in \mathcal{H}$.

Proof. From Proposition 3, we know that a Minkowski center can be obtained by rescaling the solution of the following optimization problem:

$$\begin{aligned} \max_{\boldsymbol{w}, \lambda \ge 0} \lambda \text{ s.t. } \quad & \boldsymbol{A}\boldsymbol{w} = (1+\lambda)\boldsymbol{b}, \\ & \boldsymbol{C}\boldsymbol{w} \le (1+\lambda)\boldsymbol{d}, \\ & \boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{w} - \lambda \boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{y} \le \boldsymbol{e}_i^\top \boldsymbol{d}, \quad \forall \boldsymbol{y} \in \mathcal{C}, \forall i \in [m]. \end{aligned}$$

The ith robust constraint, $i \in [m]$, is equivalent to

$$\begin{aligned} & \boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{w} + \max_{\boldsymbol{y} \in \mathcal{C}} \{-\lambda \boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{y}\} \le \boldsymbol{e}_i^\top \boldsymbol{d} \\ \Leftrightarrow\ & \boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{w} - \lambda \min_{\boldsymbol{y} \in \mathcal{C}} \{\boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{y}\} \le \boldsymbol{e}_i^\top \boldsymbol{d}, \end{aligned}$$

where the equivalence follows from the fact that $\lambda > 0$.

By definition of $\mathcal{C}$, note that $\delta_i := \min_{\boldsymbol{y} \in \mathcal{C}} \boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{y} \le d_i$, so that the constraints $\boldsymbol{C}\boldsymbol{w} - \lambda \boldsymbol{\delta} \le \boldsymbol{d}$ imply $\boldsymbol{C}\boldsymbol{w} \le (1+\lambda)\boldsymbol{d}$, which is then redundant with the robust constraint. □

According to Proposition 4, computing a Minkowski center of a polyhedron can be achieved by solving $m+1$ linear optimization problems, including m optimization problems over the same feasible set $\mathcal{C}$. Proposition 4 hence recovers the numerical approach presented in Belloni and Freund (2007, section 5.2), yet from an optimization perspective. Our approach is also numerically more efficient. Indeed, the number of optimization problems to be solved, m, does not depend on the number of equality constraints but only on the number of linear inequalities in the description of $\mathcal{C}$. On the contrary, the approach of Belloni and Freund (2007) applies to $\mathcal{C}$ described as $\mathcal{C} = \{\boldsymbol{x} | \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}; -\boldsymbol{A}\boldsymbol{x} \le -\boldsymbol{b};\ \boldsymbol{C}\boldsymbol{x} \le \boldsymbol{d}\}$, which is more prodigal in linear inequalities.

Second, we consider the case where $\mathcal{C}$ is described as the convex hull of a finite number of points. Consider m points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \in \mathbb{R}^n$ and $\mathcal{C} = \text{conv}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$. For notation convenience, let us define $\Lambda_m := \{\boldsymbol{\lambda} \in \mathbb{R}_+^m | \sum_{i \in [m]} \lambda_i = 1\}$, so that $\mathcal{C} = \{\sum_{i \in [m]} \lambda_i \boldsymbol{x}_i, | \boldsymbol{\lambda} \in \Lambda_m\}$.

Proposition 5. *Consider m points* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \in \mathbb{R}^n$ *and* $\mathcal{C} = \text{conv}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$. *The optimization problem* (2) *is equivalent to*

$$\max_{\substack{\boldsymbol{w}, \lambda \ge 0, \\ \boldsymbol{\nu}^1, \ldots, \boldsymbol{\nu}^m \in \Lambda_m}} \lambda \text{ s.t. } \ \boldsymbol{w} = \lambda \boldsymbol{x}_i + \sum_{j \in [m]} \nu_j^i \boldsymbol{x}_j, \quad \forall i \in [m].$$

Proposition 5 (proved in Online Appendix EC.2.2) recovers exactly the result provided in Belloni and Freund (2007, section 5.1). Unfortunately, this formulation involves in the order of m^2 decision variables and constraints, so column-and-constraint generation procedures could be investigated to improve practical tractability.

2.4. Analytic Expressions for Simple Sets

Deriving analytic expressions or bounds for the symmetry measure of a set can be of theoretical interest. For instance, in a robust optimization context, Bertsimas et al. (2011b) derive closed-form expression for the symmetry measure of many uncertainty sets by using the following result of Belloni and Freund (2007, equation (40)):

Lemma 1. *Consider* $\mathcal{C} = \{\boldsymbol{x} | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b};\ \boldsymbol{C}\boldsymbol{x} \le \boldsymbol{d}\}$, *where* $\boldsymbol{C} \in \mathbb{R}^{m \times n}$, $\boldsymbol{d} \in \mathbb{R}^m$. *For* $i \in [m]$, *define* $\delta_i^{:} = \min_{\boldsymbol{y} \in \mathcal{C}} \boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{y}$. *Then, for any* $\boldsymbol{x} \in \mathcal{C}$,

$$\text{sym}(\boldsymbol{x}, S) = \min_{i \in [m]} \frac{d_i - \boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{x}}{\boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{x} - \delta_i},$$

with the convention that $0/0 = +\infty$.

In particular, Lemma 1 remains valid if $\mathcal{C}$ is described as the intersection of an infinite number of half-spaces. Based on this observation, Bertsimas et al. (2015) are able to derive explicit formulae for the symmetry measure of some nonpolyhedral sets.

In this section, we give new, direct, and simple proofs for some of these results. Our proof technique relies on analyzing robust optimization Formulation (3) directly and naturally leads to generalizations to a broader class of sets than previously studied. In particular, we will

Table 1. Analytical Expression of the Minkowski Symmetry Measure and Minkowski Center of Simple Sets

No.	Convex set	Symmetry measure	Minkowski center
1	$p-$norm unit ball $\mathcal{B}_p^+ = \{\boldsymbol{x} \geq \boldsymbol{0} \,\|\, \|\boldsymbol{x}\|_p \leq 1\}$	$\frac{1}{n^{1/p}}$	$\frac{1}{n^{1/p}+1}\boldsymbol{e}$
2	Standard simplex $\Delta = \left\{\boldsymbol{x} \geq \boldsymbol{0} \,\|\, \sum_{i\in[n]} x_i \leq 1\right\}$	$\frac{1}{n}$	$\frac{1}{n+1}\boldsymbol{e}$
3	Budgeted uncertainty set, equal weights $\Delta_\gamma^e = \left\{\boldsymbol{x} \in [0,1]^m \,\|\, \sum_{i\in[n]} x_i \leq \gamma\right\}$, with $\gamma \leq n$	$\frac{\gamma}{n\min(1,\gamma)}$	$\frac{\gamma\min(1,\gamma)}{\gamma+n\min(1,\gamma)}\boldsymbol{e}$
$\boxed{4}$	Budgeted uncertainty set $\Delta_\gamma = \left\{\boldsymbol{x} \in [0,1]^n \,\|\, \sum_{i\in[n]} u_i x_i \leq \gamma\right\}$, with $\gamma \leq \sum_i u_i$, and $u_i \geq 0$	$\frac{\gamma}{\sum_{i\in[n]}\min(u_i,\gamma)}$	$\frac{\gamma}{\gamma+\sum_{i\in[n]}\min(u_i,\gamma)}\boldsymbol{e}$
$\boxed{5}$	p-norm ellipsoidal set $\mathcal{E}_p^+ = \left\{\boldsymbol{x} \in \mathbb{R}_+^n \,\|\, \|\boldsymbol{A}\boldsymbol{x}\|_p \leq 1\right\}$, with $\boldsymbol{A} \in \mathbb{R}_+^{m\times n}$	$\lambda^\star = \frac{1}{\|\boldsymbol{A}\boldsymbol{y}^\star\|_p}$ with $y_i^\star = \frac{1}{\|\boldsymbol{A}^\top \boldsymbol{e}_i\|_p}, i \in [n]$	$\frac{\lambda^\star}{1+\lambda^\star}\boldsymbol{y}^\star$

Note. A box indicates results not already derived in the literature.

consider two special structures, namely permutation-invariant sets and packing constrained sets. Some examples of convex sets and their symmetry measures are reported in Table 1.

First, we can easily compute the Minkowski measures of sets that are permutation-invariant. Indeed, in this case, (3) simplifies into a two-dimensional problem.

Lemma 2. *Assume that $\mathcal{C}$ is permutation-invariant, that is, for any $\boldsymbol{x} \in \mathcal{C}$ and any permutation σ, $\boldsymbol{x}_\sigma := (x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \in \mathcal{C}$. Then, there exists an optimal solution to (3) satisfying $\boldsymbol{w} = t\boldsymbol{e}$ for some $t \in \mathbb{R}$.*

Proof. Consider a feasible solution for (3), $(\lambda, \boldsymbol{w})$. For any permutation σ, $(\lambda, \boldsymbol{w}_\sigma)$ is also feasible, with same objective value. Define $\bar{\boldsymbol{w}} = \frac{1}{n!}\sum_{\sigma\in\Sigma_n} \boldsymbol{w}_\sigma$, where Σ_n is the set of all permutations of $[n]$. Then $(\lambda, \bar{\boldsymbol{w}})$ is also feasible with objective value λ. Applying this construction with an optimal solution $\boldsymbol{w}$ yields the result. □

To illustrate the implications of this observation, we consider the intersection of the p-norm unit ball and the nonnegative orthant.

Proposition 6. *Consider $\mathcal{B}_p^+ = \{\boldsymbol{x} \in \mathbb{R}_+^n \,|\, \|\boldsymbol{x}\|_p \leq 1\}$. Then, $(\lambda^\star, \boldsymbol{w}^\star) = \left(\frac{1}{n^{1/p}}, \frac{1}{n^{1/p}}\boldsymbol{e}\right)$ is an optimal solution of (3).*

The proof is deferred to Online Appendix EC.2.3 and relies directly on applying Lemma 2 to $\mathcal{B}_p^+$.

We now consider a broad class of polyhedra referred to as packing constrained sets, that is, sets of the form $\mathcal{P} := \{\boldsymbol{x} \geq \boldsymbol{0} \,|\, \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$, where $\boldsymbol{A} \in \mathbb{R}_+^{m\times n}$ is a matrix with nonnegative entries and $\boldsymbol{b} \in \mathbb{R}_+^m$. Among others, such sets appear in the studied multidimensional knapsack problem (Kellerer et al. 2004) or in robust optimization with budgeted uncertainty set (Bertsimas and Sim 2004), intersections of budgeted uncertainty sets, Central Limit Theorem sets (Bandi and Bertsimas 2012), and inclusion-constrained budgeted sets (Gounaris et al. 2016).

Proposition 7. *Consider $\mathcal{P} := \{\boldsymbol{x} \geq \boldsymbol{0} \,|\, \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$, with $\boldsymbol{A} \in \mathbb{R}_+^{m\times n}$ and $\boldsymbol{b} \in \mathbb{R}_+^m$. For $i \in [n]$, define*

$$y_i^\star := \max_{\boldsymbol{y}\in\mathcal{P}} \boldsymbol{e_i}^\top \boldsymbol{y} = \min_{j\in[m]} \left(\frac{b_j}{A_{ji}}\right),$$

with the convention $0/0 = +\infty$. The Minkowski measure and a scaled Minkowski center of $\mathcal{P}$ are

$$\lambda^\star = \min_{j\in[m]} \left(\frac{b_j}{\boldsymbol{e_j}^\top \boldsymbol{A}\boldsymbol{y}^\star}\right), \quad \boldsymbol{w}^\star = \lambda^\star \boldsymbol{y}^\star.$$

Among others, we can readily apply Proposition 7 to the budgeted uncertainty set. For example, for the budgeted uncertainty set with equal weights, $\Delta_\gamma^e = \{\boldsymbol{x} \in [0,1]^n \,|\, \sum_{i\in[n]} x_i \leq \gamma\}$ with $\gamma \geq 0$, we have $y_i^\star = \min(1,\gamma)$. If $\gamma \geq n$, we obtain $\lambda^\star = 1$, which is intuitive because in this case the budget constraint is redundant and $\Delta_\gamma^e = [0,1]^n$ is perfectly symmetric. If $\gamma \leq n$, we have $\lambda^\star = \frac{\gamma}{n\min(1,\gamma)}$. In particular, if $\gamma \leq 1$, we have $\lambda^\star = 1/n$, which is consistent with the fact that Δ_γ^e corresponds to a scaled simplex in this case. In the less trivial case where $1 \leq \gamma \leq n$, we obtain $\lambda^\star = \gamma/n$. A similar discussion can be conducted for a generic budgeted uncertainty set.

Furthermore, a similar line of proof can be applied to the intersection of a class of generalized ellipsoids with the nonnegative orthant, that is, sets of the form $\mathcal{E}_p^+ := \{\boldsymbol{x} \geq \boldsymbol{0} \,|\, \|\boldsymbol{A}\boldsymbol{x}\|_p \leq 1\}$, where $\boldsymbol{A} \in \mathbb{R}_+^{m\times n}$, as reported in Table 1. These "nonnegative" ellipsoids are also important uncertainty sets in the literature and have been used as baselines in many robust optimization settings (Bertsimas et al. 2011a). The corresponding proofs can be found in Online Appendices EC.2.4 and EC.2.5.

2.5. Choosing Among Minkowski Centers

As previously discussed, Minkowksi centers are not uniquely defined, not even for bounded sets. When $\mathcal{C}$ is a compact, convex set with a nonempty interior, Belloni and Freund (2007, proposition 6) prove that the set of its Minkowksi centers is a compact convex set with empty interior. Under our robust optimization lens, multiplicity of Minkowski centers relates to the multiplicity of robust optimal solutions. Indeed, it has been observed (Iancu and Trichakis 2014) that different robust optimal solutions, although leading to the same worst-case cost, can provide very different average performance. In this section, we propose two methods to choose one center among the set of all Minkowksi centers and describe them in the particular case of polyhedra.

The first method computes a Minkowski center of the set of Minkowski centers. Let $\lambda^\star$ be the objective value of (3). The set of Minkowski centers of $\mathcal{C}$ can thus be described as

$$\mathcal{M}(\mathcal{C}) = \left\{ \boldsymbol{x} \;\middle|\; \begin{array}{l} \boldsymbol{x} \in \mathcal{C} \\ (1+\lambda^\star)\boldsymbol{x} - \lambda^\star \boldsymbol{y} \in \mathcal{C}, \quad \forall \boldsymbol{y} \in \mathcal{C} \end{array} \right\}.$$

Hence, by applying Proposition 3 to $\mathcal{M}(\mathcal{C})$, we can obtain a Minkowski center of $\mathcal{M}(\mathcal{C})$ by solving the following optimization problem:

$$\max_{\boldsymbol{v},\,\mu \ge 0} \mu \;\; \text{s.t.} \;\; \boldsymbol{v}/(1+\mu) \in \mathcal{M}(\mathcal{C}),$$
$$\boldsymbol{v} - \mu \boldsymbol{z} \in \mathcal{M}(\mathcal{C}), \;\; \forall \boldsymbol{z} \in \mathcal{M}(\mathcal{C}).$$

The difficulty in the previous formulation is that the description of $\mathcal{M}(\mathcal{C})$ contains robust constraints and appears at three different places in the optimization problem: as constraints on $\boldsymbol{v}/(1+\mu)$, as constraints in the uncertainty set ($\boldsymbol{z} \in \mathcal{M}(\mathcal{C})$), and as constraints that need to be "robustified" ($\boldsymbol{v} - \mu\boldsymbol{z} \in \mathcal{M}(\mathcal{C})$). Fortunately, we can obtain a tractable formulation in the case of polyhedra.

Proposition 8. *Assume* $\mathcal{C} = \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b};\ \boldsymbol{C}\boldsymbol{x} \le \boldsymbol{d}\}$ *and denote* $\lambda^\star$ *the objective value of* (4). *Define* $\delta_i := \min_{\boldsymbol{y} \in \mathcal{C}} \boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{y}$ *and* $\tilde{\delta}_i := \min_{\boldsymbol{y} \in \mathcal{M}(\mathcal{C})} \boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{y}$ *for* $i \in [m]$. *The point* $\boldsymbol{v}^\star/(1+\mu^\star)$, *with* $(\boldsymbol{v}^\star, \mu^\star)$ *solutions of*

$$\max_{\boldsymbol{v},\,\mu \ge 0} \mu \;\; \text{s.t.} \;\; \boldsymbol{A}\boldsymbol{v} = (1+\mu)\boldsymbol{b},$$
$$(1+\lambda^\star)(\boldsymbol{C}\boldsymbol{v} - \mu\tilde{\boldsymbol{\delta}}) \le \boldsymbol{d} + \lambda^\star \boldsymbol{\delta},$$

is a Minkowski center of $\mathcal{M}(\mathcal{C})$.

Proof. From Proposition 4, we have that the set of Minkowski centers of $\mathcal{C}$ is a polyhedron $\mathcal{M}(\mathcal{C}) = \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b};\ \boldsymbol{C}\boldsymbol{x} \le \tilde{\boldsymbol{d}}\}$, with

$$\tilde{\boldsymbol{d}} := \frac{1}{1+\lambda^\star}(\boldsymbol{d} + \lambda^\star \boldsymbol{\delta}).$$

In other words, $\mathcal{M}(\mathcal{C})$ is also a polyhedron defined with the same equality constraints as $\mathcal{C}$ and the same inequality constraints except with a different right-hand side vector $\tilde{\boldsymbol{d}}$.

Applying Proposition 4 to $\mathcal{M}(\mathcal{C})$, we can obtain a Minkowski center of $\mathcal{M}(\mathcal{C})$ by rescaling the solution of

$$\max_{\boldsymbol{v},\,\mu \ge 0} \mu \;\; \text{s.t.} \;\; \boldsymbol{A}\boldsymbol{v} = (1+\mu)\boldsymbol{b}, \quad \boldsymbol{C}\boldsymbol{v} - \mu\tilde{\boldsymbol{\delta}} \le \tilde{\boldsymbol{d}}. \;\; \square$$

Remark 2. For an unbounded set $\mathcal{C}$ (which admits an infinity of Minkowski centers), the behavior of our approach would again depend on $\mathcal{H}$, the recession cone of $\mathcal{C}$. If it is not a subspace, then $\mathcal{M}(\mathcal{C}) = \mathcal{C}$ and computing a Minkowski center of $\mathcal{M}(\mathcal{C})$ is not useful. If it is a subspace, then $\mathcal{M}(\mathcal{C}) = \mathcal{M}(\mathcal{C}_0) + \mathcal{H}$ and $\mathcal{H}$ is also the recession cone of $\mathcal{M}(\mathcal{C})$, so Proposition 8 can be used to identify better Minkowski centers, but could never lead to a unique one.

Because Minkowski centers can be viewed as robust optimal solutions of a given optimization problem, the second method we propose to select one center is to consider Pareto robust optimal solutions, as defined in Iancu and Trichakis (2014).

Definition 5. Consider a polyhedron $\mathcal{C} = \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b};\ \boldsymbol{C}\boldsymbol{x} \le \boldsymbol{d}\}$ with m inequality constraints. Denote $\delta_i := \min_{\boldsymbol{y} \in \mathcal{C}} \boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{y}$ for $i \in [m]$ and $\lambda^\star$ the objective value of (4). Then, we call a solution of the optimization problem

$$\max_{\boldsymbol{x}} \boldsymbol{v}^\top(\boldsymbol{d} - (1+\lambda^\star)\boldsymbol{C}\boldsymbol{x} + \lambda^\star \boldsymbol{C}\bar{\boldsymbol{y}})$$
$$\text{s.t.} \;\; \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \; (1+\lambda^\star)\boldsymbol{C}\boldsymbol{x} \le \boldsymbol{d} + \lambda^\star \boldsymbol{\delta},$$

for some $\bar{\boldsymbol{y}}$ in the relative interior of $\mathcal{C}$ and some valuation of the constraints $\boldsymbol{v} \in \mathbb{R}^m_+$, a Pareto-optimal Minkowski center of $\mathcal{C}$.

In other words, a Pareto-optimal Minkowski center is a center that maximizes the penalized sum of the slacks in the constraints $(1+\lambda^\star)\boldsymbol{C}\boldsymbol{x} - \lambda^\star \boldsymbol{C}\bar{\boldsymbol{y}} \le \boldsymbol{d}$, at some predefined point $\bar{\boldsymbol{y}}$.

Remark 3. If $\mathcal{C}$ is unbounded and its recession cone is not a subspace, then we can find a nontrivial Pareto-optimal Minkowski center of $\mathcal{C}$ whenever the valuation vector $\boldsymbol{v}$ is chosen so that $v_i = 0$ if $\delta_i = -\infty$. If $\mathcal{C}$ is unbounded and its recession cone is a subspace, however, the set of Pareto-optimal Minkowski centers is invariant by translation by $\mathcal{H}$; hence, it is unbounded.

3. Practical Benefits of Minkowski Centers

Although Minkowski centers have mostly been regarded as theoretical objects, the previous section shows that it can be expressed as the solution of a tractable linear optimization problem for polyhedra. In this section, we illustrate numerically the tractability and potential practical benefits from using Minkowski centers (instead of available alternatives) in two popular algorithms.

Table 2. Median (and Interquartile Range) for Runtime and Three Performance Metrics for Analytic, Chebyshev, and Minkowski Centers

Method	Runtime	Symmetry measure	Depth	Average sum of log slacks
Chebyshev	0.014 (0.036)	0.0 (0.004)	0.025 (0.598)	−1.584 (20.371)
Analytic	0.231 (0.349)	0.009 (0.075)	0.0 (0.148)	0.066 (7.172)
Minkowski	5.308 (23.278)	0.056 (0.13)	0.0 (0.109)	−0.756 (5.322)

Note. We report runtime (in seconds), the symmetry measure $\text{sym}(\boldsymbol{x}, \mathcal{C})$ and the average sum of log-slacks $\frac{1}{m}\sum \log(d_i - \boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{x})$.

3.1. Computational Tractability

We first evaluate the numerical scalability of computing Minkowski centers of polyhedra and how it compares to other known and used centers, namely analytic and Chebyshev centers, on 78 polyhedra from the NETLIB library (Gay 1985) and 37 from the MIPLIB library. As reported in Table 2, these computational times are one order of magnitude higher than those needed to compute analytic and Chebyshev centers. To the best of our knowledge, our paper is the first to investigate the numerical tractability of Minkowski centers, although it has been extensively used for theoretical purposes. We also report measures of centrality: the measure of symmetry $\text{sym}(\boldsymbol{x}, \mathcal{C})$, the depth, and the average sum of log-slacks $\frac{1}{m}\sum_{i\in[m]}\log(d_i - \boldsymbol{e}_i^\top \boldsymbol{C}\boldsymbol{x})$. These three metrics are maximized (by definition) by Minkowski, Chebyshev, and analytic centers, respectively. By reporting these measures, we want to emphasize how complex and ambiguous it is to properly define a center of a set and how varied the current definitions are. In the next two sections, we adopt a more pragmatic approach and evaluate the benefit from using Minkowski centers as initialization points of two numerical algorithms.

3.2. Hit-and-Run Algorithm

Hit-and-run is a standard algorithm for sampling random points from an arbitrary density on a high dimensional Euclidian space (see Chen and Schmeiser (1993) for a comparison of sampling schemes), initially proposed by Smith (1984). In particular here, we apply it to sample points uniformly over a polyhedron $\mathcal{P}$.

The hit-and-run (HAR) algorithm starts at an initial point $\boldsymbol{x}_0 \in \mathcal{P}$ and generates a sequence $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ in $\mathcal{P}$ with random increments $\boldsymbol{x}_{m+1} - \boldsymbol{x}_m$. Precisely, at step m, we generate a random direction $\boldsymbol{d}_m$. The line passing through $\boldsymbol{x}_m$ with direction $\boldsymbol{d}_m$ *hits* the boundary of $\mathcal{P}$ at some points $\boldsymbol{y}_m^-$ and $\boldsymbol{y}_m^+$. We sample $\boldsymbol{x}_{m+1}$ uniformly over the segment $[\boldsymbol{y}_m^-, \boldsymbol{y}_m^+]$. Algorithm 2 in Online Appendix EC.3.1 describes the algorithm in pseudocode for a polyhedron $\mathcal{P}$ described as the intersection of halfspaces (see Bélisle et al. (1993) for its extension to generic compact convex sets). It was later shown to have polynomial mixing time for sampling from convex sets (Lovász 1999) and seems to be much faster in practice. We refer to Bélisle et al. (1998) for a careful review of the literature.

The sequence of points generated, $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$, is an ergodic Markov chain that geometrically converges to the uniform distribution over $\mathcal{P}$ (Chen and Schmeiser 1993, sections 2 and 3). To estimate the expected value of some functional of $\tilde{x}$, $\mathbb{E}[h(\tilde{x})]$, using N uniformly sampled points from $\mathcal{P}$, two options are possible: (a) run Algorithm 2 with m steps, N times and consider $\{\boldsymbol{x}_m^{(i)}, i \in [N]\}$ or (b) run Algorithm 2 with $m \times N$ steps and consider $\{\boldsymbol{x}_{im}, i \in [N]\}$. Generally speaking, for a fixed value of $m \times N$, option (b) will provide better point estimates but worse standard errors, due to auto-correlations between the samples (Chen and Schmeiser 1993, section 5.1). In any case, it is crucial that the distribution of the sequence generated by the algorithm converges as fast as possible (in terms of number of steps m) toward the uniform distribution. Intuitively, starting from a "central" point $\boldsymbol{x}_0$ should speed up convergence.

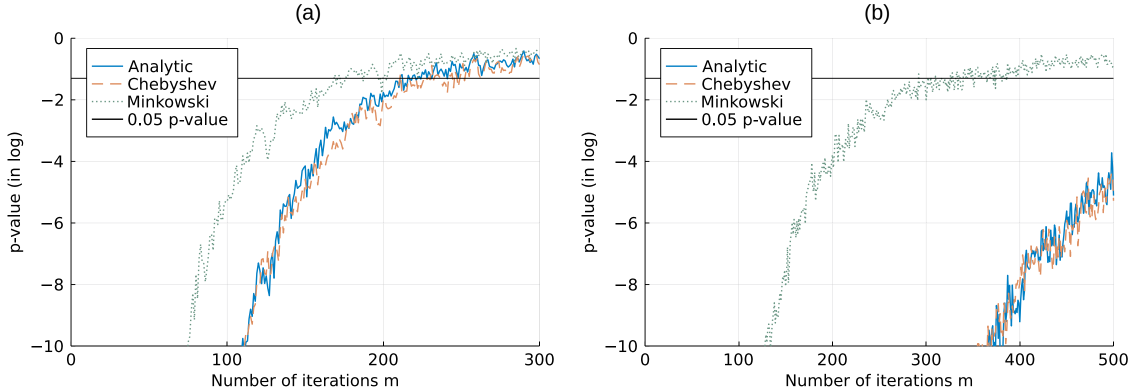
Formally, we want to test the null hypothesis:

$$(H_0^m) : \tilde{\boldsymbol{x}}_m \text{ is uniformly distributed on } \mathcal{P},$$

using an independent identically distributed random sample of size $n = 5,000$. Díaz et al. (2006) develop a method for testing this hypothesis called the distance to boundary (DB) test.

For a compact subset of $\mathbb{R}^n$, $\mathcal{C}$, define the distance of any point $\boldsymbol{x}$ to the boundary as $D(\boldsymbol{x}, \partial\mathcal{C}) = \min\{\|\boldsymbol{x} - \boldsymbol{y}\| : \boldsymbol{y} \in \partial\mathcal{C}\}$ and denote by R the maximum distance to the boundary that can be attained on $\mathcal{C}$, that is, $R = \max\{D(\boldsymbol{x}, \partial\mathcal{C}) : \boldsymbol{x} \in \mathcal{C}\}$. The quantity R is sometimes called the depth of $\mathcal{C}$ and $D(\boldsymbol{x}, \delta\mathcal{C})/R$ the relative depth at $\boldsymbol{x}$. For a wide class of sets, namely sets that are "invariant by erosion," Díaz et al. (2006) show that, under (H_0^m), the relative depth $\tilde{y}_m = D(\tilde{\boldsymbol{x}}_m, \delta\mathcal{S})/R$ follows a beta distribution with parameters $(1, d)$, that is, its cumulative distribution function is $y \longmapsto 1 - (1 - y)^d$, for $y \in [0, 1]$. Accordingly, we can test (H_0^m) by testing whether $\tilde{y}_m$ follows the right distribution via a Kolmogorov-Smirnov test. In particular, this result holds for a convex polyhedron circumscribed to a ball, that is, defined as the intersection of halfspaces that are all tangent to a ball. We will use this type of polyhedra in our experiments.

Figure 2. (Color online) p Value of a DB Test for the Hit-and-Run Algorithm, as the Number of Interactions m Increases



Notes. Results are averaged over 20 random polyhedra defined as the intersection of 10 halfspaces. (a) $n = 50$. (b) $n = 100$.

For our experiment, we generate random convex polyhedra circumscribed to a ball in dimension $n \in \{10, 20, 50, 100\}$ (see Algorithm 3 in Online Appendix EC.3). We run the HAR algorithm with different initial points x_0. In particular, we compare the Minkowski, Chebyshev, and analytic centers. Figure 2 represents the p value of the DB test for (H_0^m) as a function of the number of steps m. Recall that one can reject the null hypothesis (H_0^m) (i.e., conclude that the sample is not uniformly distributed) when the p value is low. We also display a 0.05 cutoff. We observe that the hit-and-run algorithm initialized with a Minkowski center converges faster to a uniform distribution than when initialized with either analytic or Chebyshev centers. In particular, the benefit from a Minkowski center increases as the dimension of the space n increases.

We also compute the number of iterations m required for Algorithm 2 to achieve a p value of 0.05 for each initialization point. Table 3 reports the average number of additional iterations required with the analytic center versus the Minkowski center. We had to limit the total number of iterations in the HAR algorithm (we used 500 in our experiments). So, on some instances, some methods failed to reach the 0.05 p value target within the limit. In these cases, to allow for a fair comparison, we compare the number of iterations required to achieve a lower p value target, equal to the best p value achieved by the worst performing method. Table 3 confirms the benefit from the Minkowski center, especially in high dimensions. In low dimension, we observe that the analytic center seems to perform better when more constraints define the polyhedron. We confirm these findings by doing a regression analysis of the number of additional iterations (in log terms) as a function of the dimension n and the number of halfspaces defining the polyhedron (see Table EC.2 in Online Appendix EC.4.1). In Online Appendix EC.4.1, Table EC.1, we conduct a similar analysis for the Chebyshev center and observe similar (although marginally stronger) benefits.

3.3. Cutting-Plane Algorithm

Cutting-plane methods (CPMs) are a broad family of algorithms for solving convex or quasiconvex nondifferentiable optimization problems (see Elhedhli et al. (2009) for a comprehensive overview). To motivate the use of Minkowski centers within CPM schemes, we consider in this section the basic implementation of a CPM algorithm to minimize a piece-wise linear convex function over an intersection of ellipsoids and evaluate its performance on random instances by following the methodology of Boyd et al. (2008), depending on the definition of center used.

We consider a generic problem of the form

$$\min_{x,t} \; t \;\; \text{s.t.} \;\; (x,t) \in \mathcal{C}, \tag{5}$$

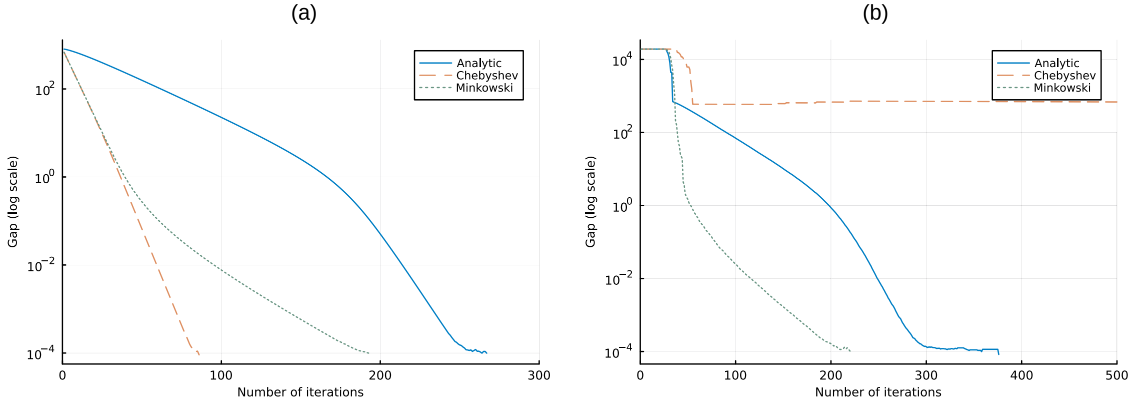
where $\mathcal{C}$ is a convex set. Typically, (5) arises as the epigraph formulation of a constrained minimization problem.

Table 3. Number of Additional Iterations Required by Algorithm 2 When Initialized with the Analytic Center vs. Minkowski Center

	No. halfspaces (p)				
Dimension (n)	10	20	30	40	50
10	0.3 (0.3)	−1.0 (0.5)	−3.1 (0.5)	−2.8 (0.7)	−3.2 (0.6)
20	4.1 (1.3)	3.8 (1.1)	1.6 (1.3)	−4.0 (1.2)	−5.6 (1.3)
50	47.9 (5.3)	69.5 (4.3)	61.8 (5.5)	54.9 (4.0)	44.7 (4.6)
100	283.6 (8.9)	362.1 (4.9)	362.0 (7.4)	375.4 (7.4)	376.1 (6.7)

Note. We report the average number over 20 random polyhedra (and standard errors).

Figure 3. (Color online) Convergence Profile (in Terms of Iterations) of the CPM for Different Query Points



Notes. Results are averaged over 20 random instances in dimension $n = 20$ with $m = 100$ linear pieces. (a) $k = 0$. (b) $k = 10$.

In our implementation, we will consider the minimization of a piecewise linear function over an intersection of ellipsoids, that is,

$$\min_{x,t} t \text{ s.t. } \boldsymbol{a}_i^\top \boldsymbol{x} + b_i \le t,\ \forall i \in [m],$$

$$\|\boldsymbol{F}_i \boldsymbol{x} + \boldsymbol{g}_i\|_2 \le 1,\ \forall i \in [k].$$

To apply the CPM described in Algorithm 1, three ingredients are needed: First, the ability to test whether the current solution is feasible, $(\boldsymbol{x}_k, t_k) \in \mathcal{C}$. Second, an oracle that, given an infeasible solution $(\boldsymbol{x}_k, t_k)$, provides a hyperplane that separates the current solution from the feasible set $\mathcal{C}$. In our case, we will simply consider the most violated constraint and obtain a separating hyperplane by linearizing it around the current solution. Finally, and most relevant to our experiments, we need a `query` function that returns a point from a given polyhedron. From a convergence perspective, it is understood that the query point should be "central" so that the volume of $\mathcal{P}_k$ decreases fast. In our experiments, we will numerically compare the convergence of this algorithm when a Minkowski, analytic, or Chebyshev center is used as a query point. We shall denote the variants as MC-, AC-, and CC-CPM algorithms. Regarding the termination criterion, we impose a limit on the total number of iterations (600 for MC-CPM and 60,000 for AC- and CC-CPM in our experiments) and the bound gap (10^{-4}). We consider instances in $n \in \{10, 20, 50\}$ dimensions, with $m \in \{100, 200, 500\}$ linear pieces, and $k \in \{0, 1, 5, 10\}$ ellipsoids. As in Boyd et al. (2008), instances are generated randomly by sampling the entries of $\boldsymbol{a}_i$ and b_i independently from a standard normal distribution. We sample $\boldsymbol{F}_i \sim \mathcal{U}([0,1])^{n/2\times n}$ and $\boldsymbol{g}_i$ uniformly in the unit ball, so that $\boldsymbol{x} = \boldsymbol{0}$ is always feasible. We use $\mathcal{P}_0 = \{(\boldsymbol{x}, t) : \boldsymbol{a}_i^\top \boldsymbol{x} + b_i \le t,\ \forall i \in [m]\}$ as our initial polyhedron.

Algorithm 1 (CPM for Solving (5))

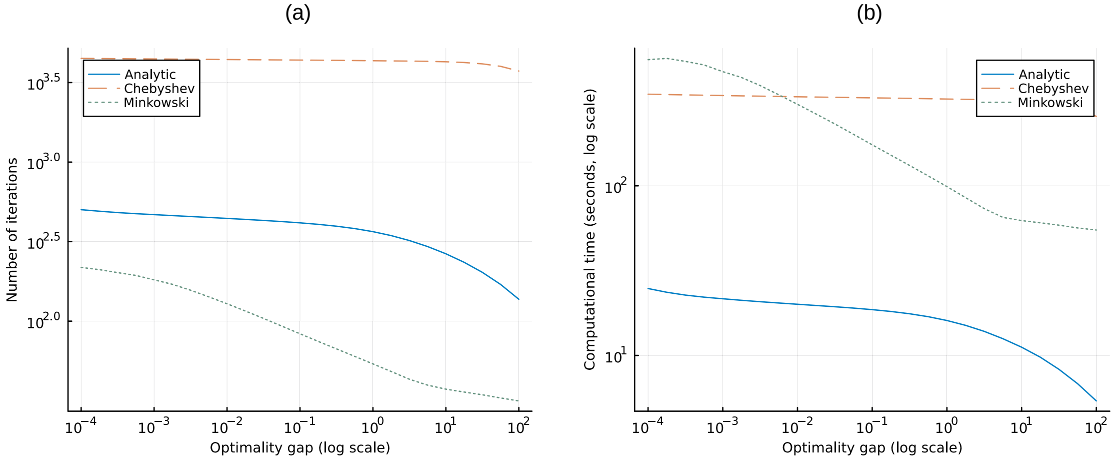
Input: Initial polytope $\mathcal{P}_0$ enclosing $\mathcal{C}$.
Output: A solution $\boldsymbol{x}$ to (5).

```
query a point $(\boldsymbol{x}_0, t_0) \in \mathcal{P}_0$.
while termination criterion not met do
    if $(\boldsymbol{x}_k, t_k) \in \mathcal{C}$ then
        set $\mathcal{P}_{k+1} = \mathcal{P}_k \cap \{(\boldsymbol{x}, t) | t \le t_k\}$.
    else
        an oracle finds a separating hyperplane, that is, $(\boldsymbol{a}, a, b)$ s.t. $\boldsymbol{a}^\top \boldsymbol{x}_k + a t_k > b$ but $\mathcal{C} \subseteq \{(\boldsymbol{x}, t) | \boldsymbol{a}^\top \boldsymbol{x} + at \le b\}$.
        set $\mathcal{P}_{k+1} = \mathcal{P}_k \cap \{(\boldsymbol{x}, t) | \boldsymbol{a}^\top \boldsymbol{x} + at \le b\}$.
    query $(\boldsymbol{x}_{k+1}, t_{k+1}) \in \mathcal{P}_{k+1}$.
```

Figure 3 displays the convergence profile (in terms of the number of iterations) of the suboptimality gap, averaged over 20 instances in dimension $n = 50$ with $m = 500$ linear pieces. The left and right panels report results for instances with $k = 0$ and $k = 10$ second-order cone constraints, respectively. First, we should emphasize that no method is a clear winner and that the quality of a query point depends heavily on the instance. For example, in the linear case ($k = 0$), the CPM with Chebyshev center demonstrates the fastest convergence, whereas it requires two orders of magnitude more iterations than MC- and AC-CPM with $k = 10$ SOC constraints. Second, we observe that Minkowski centers provide a competitive alternative to analytic or Chebyshev centers. In particular, when $k = 10$, it nearly halves the number of iterations compared with using analytic centers. However, in Figure 3, convergence is measured in terms of number of iterations. Because computing Minkowski centers requires one to two orders of magnitude more time and because the query function is invoked at each iteration, the reduction in number of iterations does not translate into a reduction in computational time (see Figure EC.1 in Online Appendix EC.4.2).

To verify this finding across various problem sizes, we compare, for different targets of optimality gap, the total number of iterations and total computational time needed by each algorithm to achieve this gap, averaged over all instances generated. As displayed in Figure 4, we observe that MC-CPM requires fewer iterations than other query strategies on average, across all tolerance

Figure 4. (Color online) Average Convergence Speed of the CPM for Different Query Points and Different Target Optimality Gaps



Notes. (a) Number of iterations. (b) Computational time.

level. However, given the additional computational burden, it is usually slower than AC-CPM, but competes with CC-CPM for moderate optimality gaps. Although these metrics are averaged across all instances, we observed in Figure 3 that the performance (and especially that of CC-CPM) depends greatly on the presence of SOC constraints, among other problem dimensions. To capture the effect of n, m, and k on the performance of each method, we report the average number of iterations and computational time to achieve a 10^{-2} optimality gap for each method, for all values of n and m, and some values of k, in Online Appendix EC.4.2.

Contrasting our findings on instances of the HAR algorithm (Section 3.2) and the CPM (Section 3.3), we conclude that there is no absolute best definition of centers but that the convergence of numerical algorithms can be significantly (and positively) impacted by choosing the "right" one for the algorithm and the particular instance at hand. On this matter, we argue that Minkowski centers, which have been largely overlooked for computational purposes, should be considered as a potential candidate. Given the additional burden of computing Minkowski centers, they could be most impactful for algorithms that query a strictly feasible point once (e.g., at initialization). Given the drastic reduction in the number of iterations we observe for the CPM, further work could also investigate the design of tailored solvers for computing Minkowski centers more efficiently.

4. Tractable Approximations for Projections of Polyhedra

In this section, we consider the important case where the convex set $\mathcal{C}$ is the projection of a polyhedron. Precisely, we consider a polyhedron

$$\mathcal{P} = \{(\boldsymbol{x}, \boldsymbol{z}) \in \mathbb{R}^{n_x+n_z} \,|\, \boldsymbol{A}_x\boldsymbol{x} + \boldsymbol{A}_z\boldsymbol{z} = \boldsymbol{b},\ \boldsymbol{C}_x\boldsymbol{x} + \boldsymbol{C}_z\boldsymbol{z} \le \boldsymbol{d}\},$$

and its projection onto the $\boldsymbol{x}$−space, that is, $\mathcal{P}_x = \{\boldsymbol{x} \in \mathbb{R}^{n_x} \,|\, \exists \boldsymbol{z} \in \mathbb{R}^{n_z} \text{s.t.} (\boldsymbol{x}, \boldsymbol{z}) \in \mathcal{P}\}$. In optimization, and combinatorial optimization in particular, such definition of sets as polyhedral projections are commonly referred to as extended or lifted formulations (Conforti et al. 2010).

The general approach for computing a (Minkowski) center for $\mathcal{P}_x$ would be to first derive an explicit algebraic description of $\mathcal{P}_x$ which does not rely on any additional variables $\boldsymbol{z}$, for instance by using Fourier-Motzkin elimination (FME; Motzkin 1936). However, the number of constraints resulting from this procedure grows exponentially in n_z. Moreover, FME introduces many redundant constraints that would need to be identified and removed or might negatively impact the quality of the analytic center. Hence, an algorithm that could compute a center of $\mathcal{P}_x$ by working directly on its lifted description would be extremely tractable and valuable.

Also, the projection of a Minkowski center of $\mathcal{P}$ seems like a natural candidate for a Minkowski center of $\mathcal{P}_x$. However, we show that this approach fails.

Lemma 3. *The projection onto the $\boldsymbol{x}$-space of a Minkowski center of $\mathcal{P}$ is not necessarily a Minkowski center of $\mathcal{P}_x$.*

Proof. Our proof is based on the following counterexample. In dimension n, consider the set $\mathcal{P}_n = \{\boldsymbol{x} \in \mathbb{R}^n_+ \,|\, \boldsymbol{e}^\top \boldsymbol{x} \le 1\}$. The Minkowski center of $\mathcal{P}_n$ is the vector $\frac{1}{n+1}\boldsymbol{e}$ and its measure of symmetry is $\frac{1}{n}$. In particular, $\mathcal{P}_1 = [0, 1]$ and its center is $1/2$. If we consider the projection of $\mathcal{P}_n$ onto the first coordinate, we recover $\mathcal{P}_1$. However, the projection of the Minkowski center is $1/(n+1) \neq 1/2$ for $n \ge 2$. □

Actually, the proof of Lemma 3 shows that the projection of the Minkowski center of $\mathcal{P}_n$ onto the first coordinate is not even a Helly center of the set $\mathcal{P}_1$.

Furthermore, one can show that projection can only improve symmetry.

Lemma 4. For any polyhedron $\mathcal{P}$, $\text{sym}(\mathcal{P}) \le \text{sym}(\mathcal{P}_x)$.

Proof. Consider a center of Minkowski of $\mathcal{P}$, (x, z). Then, $\text{sym}(x, \mathcal{P}_x) \ge \text{sym}(\mathcal{P})$. □

4.1. Adjustable Robust Optimization Reformulation

We now derive an analog of Proposition 3 that applies to the case where the set is described as the projection of a polyhedron directly. In this case, computing a Minkowski center for $\mathcal{P}_x$ is equivalent to an adjustable robust optimization (ARO) problem.

Proposition 9. *Consider the set*

$$\mathcal{P}_x = \{x \in \mathbb{R}^{n_x} \mid \exists z \in \mathbb{R}^{n_z} : A_x x + A_z z = b, \\ C_x x + C_z z \le d\}.$$

Let $(w^\star, z_w^\star, \lambda^\star)$ be solution of the adjustable robust optimization problem

$$\max_{w, z_w, \lambda \ge 0} \lambda \text{ s.t. } \quad A_x w + A_z z_w = (1+\lambda) b, \\ C_x w + C_z z_w \le (1+\lambda) d, \\ \forall (y, z_y) \in \mathcal{P}, \exists z : (w - \lambda y, z) \in \mathcal{P}. \quad (6)$$

Then $x^\star = w^\star/(1+\lambda^\star)$ is a Minkowski center for $\mathcal{P}_x$.

Proof. From the proof of Proposition 3, we know that the result holds with $(w^\star, \lambda^\star)$ solution of

$$\max_{w, \lambda \ge 0} \lambda \text{ s.t. } \frac{w}{1+\lambda} \in \mathcal{P}_x, \\ w - \lambda y \in \mathcal{P}_x, \ \forall y \in \mathcal{P}_x.$$

By an appropriate rescaling of the additional variables,

$$\frac{w}{1+\lambda} \in \mathcal{P}_x \Leftrightarrow \exists z_w : A_x w + A_z z_w = (1+\lambda) b, \\ C_x w + C_z z_w \le (1+\lambda) d.$$

Finally, the robust constraints can be rewritten as $\forall (y, z_y) \in \mathcal{P}, \exists z : (w - \lambda y, z) \in \mathcal{P}$. □

The term "adjustable" comes from the fact that in the robust constraints, the additional variable z, needed to certify that $w - \lambda y \in \mathcal{P}_x$, can be adjusted to the uncertain parameter y. Effectively, z is a function of y (and potentially of z_y as well). Linear ARO optimization problems are NP-hard in general, as stated in Ben-Tal et al. (2004) and proved in Guslitser (2002, section 3.3). Instead of solving (6) exactly, we can obtain tractable approximations by restricting our attention to parametrized functional forms for z (as a function of y and z_y).

For instance, restricting our attention to z of the form $z = z_w - \lambda z_y$, is equivalent to computing a Minkowski center for $\mathcal{P}$ and taking its projection onto the x-space. Exploring a larger class of policies might lead to stronger formulations and better approximations. In the following section, we propose restricting the scope to general affine decision rules to derive tractable approximations of Minkowski centers, and a lower bound on the symmetry measure of $\mathcal{P}_x$.

Remark 4. The optimization formulations and algorithms we present in this section are described for evaluating (or approximating) the symmetry measure of the set $\mathcal{P}_x$. By adding the linear constraints $w = (1+\lambda)x$, they can be used to evaluate $\text{sym}(x, \mathcal{P}_x)$ for a given x as well.

4.2. Converging Exact Algorithm

The adjustable robust optimization problem (6) can be approximated by replacing the uncertainty set $\mathcal{P}$ by a finite number of scenarios and solving the fully adjustable problem (6) on this discrete uncertainty set. Precisely, given $y^{(1)}, \ldots, y^{(k)} \in \mathcal{P}_x$, we solve

$$\max_{w, z_w, z^{(1)}, \ldots, z^{(k)}, \lambda \ge 0} \lambda \text{ s.t. } \quad A_x w + A_z z_w = (1+\lambda) b, \\ C_x w + C_z z_w \le (1+\lambda) d, \\ \forall i \in [k], (w - \lambda y^{(i)}, z^{(i)}) \in \mathcal{P}. \quad (7)$$

Because (7) is less constrained than (6), its value provides an upper bound on $\text{sym}(\mathcal{P}_x)$. This generic approach in adjustable robust optimization was first presented by Hadjiyiannis et al. (2011) and is known to guarantee tight upper bounds for a large class of ARO problems and applications. We will later refer to this method via the acronym HGK, the initials of the authors.

Zeng and Zhao (2013) proposed a column-and-constraint generation scheme that converges to the exact robust solution of (6) by solving a sequence of problems of the form (7). Their method relies on an oracle that, for any solution (w, λ), can either certify that it satisfies the robust constraint "$\forall (y, z_y) \in \mathcal{P}, \exists z : (w - \lambda y, z) \in \mathcal{P}$," or returns a scenario $\bar{y} \in \mathcal{P}_x$ for which no such z exist. In Online Appendix EC.3.3, we describe how this oracle amounts to solving a nonconvex maximization problem. At each iteration of their algorithm, a robust optimization problem with finite uncertainty of Form (7) is solved. If the resulting solution is robust feasible, then it is an optimal solution to the original problem (6). Otherwise, the oracle provides a new scenario $\bar{y}$ to be added to (7). This algorithm is guaranteed to converge in a finite number of iterations, yet that can be linear in the number of extreme points of $\mathcal{P}_x$ (Zeng and Zhao 2013, proposition 2). In practice, we typically impose a limit on the number of iterations or total computational time.

In our setting, if the algorithm does not converge within the allocated budget, then the resulting solution (w, λ) is not robust feasible. In other words, the point $x := w/(1+\lambda)$ belongs to $\mathcal{P}_x$ and is a candidate for being a Minkowski center, but the value of λ provided is an upper bound on both $\mathrm{sym}(x, \mathcal{P}_x)$ and $\mathrm{sym}(\mathcal{P}_x)$.

Alternatively, Zhen et al. (2018) proposed a general method based on FME to solve adjustable linear robust optimization problems like (6). In our case, their approach is equivalent to using FME to derive an explicit description of $\mathcal{P}_x$ and then compute the Minkowski center of $\mathcal{P}_x$ using the formulations from Section 2.

4.3. Approximations with Computationable Suboptimality Gaps

We restrict our attention to adjustable variables of the form

$$z = Yy + Zz_y + z_0,$$

where Y, Z, z_0 are here-and-now decision variables. For instance, taking $Y = 0$, $Z = -\lambda I$, and $z_0 = z_w$ recovers the projection of a Minkowski center of $\mathcal{P}$. Among others, such linear decision rules (LDRs in short, first proposed by Ben-Tal et al. 2004) are simple, tractable, and often enjoy strong empirical and theoretical performance for adjustable robust optimization problems (Bertsimas et al. 2010, Bertsimas and Goyal 2012, Housni and Goyal 2021). All in all, we solve

$$\begin{aligned} \max_{w, z_w, Z, Y, z_0, \lambda \geq 0} \lambda \text{ s.t. } & A_x w + A_z z_w = (1+\lambda) b, \\ & C_x w + C_z z_w \leq (1+\lambda) d, \\ & \forall (y, z_y) \in \mathcal{P}, \, (w - \lambda y, Yy + Zz_y + z_0) \in \mathcal{P}. \end{aligned} \tag{8}$$

The objective value of the above optimization problem $\lambda^\star_{LDR}$ provides a lower bound on the actual symmetry of $\mathcal{P}_x$, that is, $\lambda^\star_{LDR} \leq \mathrm{sym}(\mathcal{P}_x)$. Among others, Bertsimas et al. (2010) and Ben-Ameur et al. (2018) show that linear decision rules are optimal (hence, the inequality is tight) when the uncertainty set (here, $\mathcal{P}$) is a standard simplex.

The robust constraints in (8) can be written explicitly

$$\begin{aligned} \forall (y, z_y) \in \mathcal{P}, \, A_x w - \lambda A_x y + A_z Yy + A_z Z z_y + A_z z_0 &= b, \\ C_x w - \lambda C_x y + C_z Yy + C_z Z z_y + C_z z_0 &\leq d. \end{aligned}$$

They can then be enforced numerically either by adopting a cutting-plane approach or by computing the robust counterpart of each constraint separately via strong duality (Bertsimas et al. 2016), thus leading to a linear optimization problem. We implement the later approach for our numerical experiments.

To measure the quality of the approximation provided by using linear decision rules, we also derive an upper bound on $\mathrm{sym}(\mathcal{P}_x)$ using the generic approach described in the previous section. To identify the scenarios $y^j, j \in [k]$, Hadjiyiannis et al. (2011) suggest considering each robust constraint and compute the binding scenarios for each of them, the decision variables being fixed. In our implementation, we follow their recommendation to obtain a finite number of scenarios and solve one instance of (7) to obtain a valid upper bound. We denote the objective value of (7) λ^k_{HGK} or $\lambda^k_{HGK}(x)$ if we upper bound $\mathrm{sym}(x, \mathcal{P}_x)$ by adding the constraints $w = (1+\lambda)x$.

Remark 5. In Online Appendix EC.3.4, we describe a potentially stronger approximation based on quadratic decision rules. Unfortunately, the robust counterpart of these problems can only be approximated by large semidefinite optimization problems, which we were unable to solve to reasonable accuracy, even for the smallest polyhedron we considered (dimension $n = 10$).

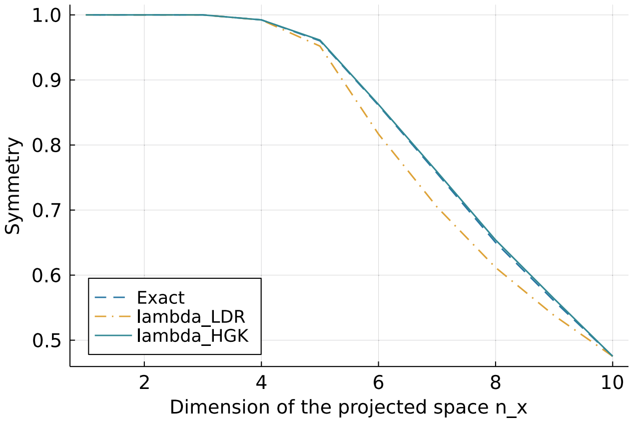
4.4. Numerical Experiments

In this section, we evaluate the performance of our method for computing approximate values of the Minkowski measure (lower bounds via (8) and upper bounds via (7)) for polytopic projections.

First, we evaluate the quality of our ARO-based approximation on small instances where exact methods apply. We generate random polyhedra, following the same generation methodology as Section 3.2, in $n = 10$ dimensions and using $p = 10$ linear inequalities. For each polyhedron, we consider its projection onto the first n_x coordinates, $n_x \in [n]$. Hence, $n - n_x$ corresponds to the number of dimensions eliminated. We compute the approximate Minkowski center obtained by solving (8) and obtain λ^*_{LDR}, and λ^*_{HGK}. As a comparison, we apply two exact approaches: We perform a FME procedure to obtain an explicit description of the projected polyhedron and then compute its Minkowski center by solving (4). Alternatively, we implement the column-and-constraint generation (C&CG) algorithm of Zeng and Zhao (2013). We impose a 1,000-iteration (respectively, 2,000-iteration) limit and use an exact spatial-branch-and-bound solver (respectively, the heuristic linearization technique of Bertsimas et al. 2012) as the oracle.

Figure 5 compares the lower and upper bounds, λ^*_{LDR} and λ^*_{HGK}, with the exact value of $\mathrm{sym}(\mathcal{P}_x)$ for different values of n_x. Notice that $n_x = n = 10$ corresponds to the case $\mathcal{P} = \mathcal{P}_x$ so we naturally expect $\lambda^*_{LDR} = \lambda^*_{HGK} = \mathrm{sym}(\mathcal{P}_x)$. At the other extreme, when $n_x = 1$, $\mathcal{P}_x$ is a segment, which is perfectly symmetric so one should conclude that $\mathrm{sym}(\mathcal{P}_x) = 1$. First, we observe that the lower bound, the upper bound, and the exact value of symmetry $\mathrm{sym}(\mathcal{P}_x)$ are nonincreasing with n_x. In other words, projecting increases symmetry, which validates experimentally Lemma 4. We also observe on Figure 5 that our adjustable robust optimization approach provides valid

Figure 5. (Color online) Comparison of λ^*_{LDR}, λ^*_{HGK}, sym$(\mathcal{P}_x)$ for Different Values of n_x



Note. Results are averaged over 20 polyhedra in dimension $n = 10$.

and small (within 5%) intervals on sym$(\mathcal{P}_x)$. In particular, the upper bound derived from (7) is almost tight. Further improvement should thus mainly come from improving the lower bound. However, the width of the interval $[\lambda^*_{LDR}, \lambda^*_{HGK}]$ does not necessarily imply a bound on the distance between the returned solution and the set of Minkowski centers for $\mathcal{P}_x$, although intuition suggests the tighter the interval the closer the solution is to an actual Minkowski center (see Figure EC.3 in Online Appendix EC.4.3).

In addition to providing high-quality solutions, the adjustable robust optimization approach is also significantly more computationally tractable than the exact approaches, as reported in Table 4. After accounting for the time required by the FME procedure, the ARO approach is 10^4 times faster than the FME-based approach. We should also mention that FME requires substantial memory and we could not perform simulations on larger instances with 16 GB of RAM. The C&CG algorithm terminates in a short number of iterations for $n_x = 1, 2, 3$, yet requires one to two orders of magnitude more time than ARO. For larger values of n_x, however, the method often fails to converge within 1,000 (respectively, 2,000) iterations, which represent three to nine hours of computation. We should emphasize here that, when terminated early, the C&CG algorithm cannot provide any suboptimality gap on the incumbent solution it returns.

Table 4. Average Runtimes (in Seconds) for the Adjustable Robust Optimization Approach (Both Lower and Upper Bounds) Compared with Two Exact Approaches: FME Followed by Solving (4) and Column-and-Constraint Generation (C&CG) Approach of Zeng and Zhao (2013) with an Exact or Heuristic Oracle

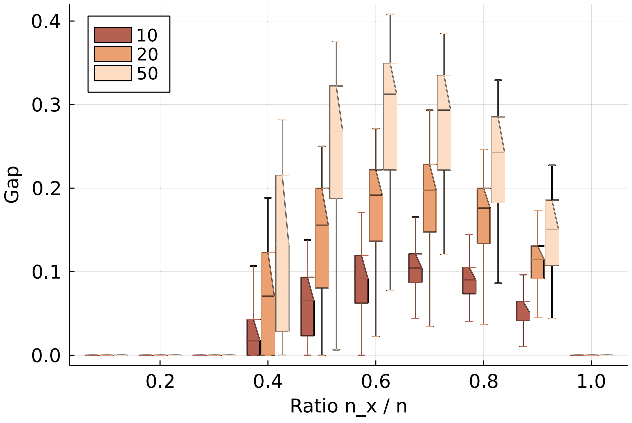
	ARO		FME-based method		C&CG method	
n_x	(8)	(7)	FME	(4)	Exact	Heuristic
10	0.02	0.05	0.0	0.01	29,107.41	20,840.75
9	0.03	0.06	7.50	0.01	25,946.54	37,127.03
8	0.03	0.06	22.72	0.02	21,612.46	27,550.29
7	0.03	0.06	242.16	0.02	14,063.92	28,630.70
6	0.03	0.06	337.29	0.01	15,031.63	31,139.72
5	0.03	0.05	346.70	0.00	14,027.92	21,724.48
4	0.03	0.06	347.67	0.00	19,228.26	11,924.57
3	0.04	0.06	347.69	0.00	1.03	1.93
2	0.04	0.06	347.69	0.00	0.61	0.95
1	0.04	0.06	347.69	0.00	0.46	0.30

Note. Results are averaged over 20 iterations.

We conduct further experiments in higher dimensions, $n \in \{10, 20, 50\}$, and for polyhedra defined with $p \in \{10, 20, 30, 40, 50\}$ inequalities. Figure 6 represents the distribution of the gap $(\lambda^*_{HGK} - \lambda^*_{LDR})/\lambda^*_{HGK}$ for varying values of n_x/n and varying values of n. We observe a similar qualitative behavior as in Figure 5: The width of the interval $[\lambda^*_{LDR}, \lambda^*_{HGK}]$ increases and then decreases with n_x. A more detailed regression analysis (Table EC.6 in Online Appendix EC.4.3) suggests the gap scales as $0.9(n_x/n) - 0.7(n_x/n)^2$, hence maximized for $n_x/n \approx 0.64$, which is consistent with our observations. We also observe that the gap increases with the total dimension n and with the number of inequality constraints defining $\mathcal{P}$.

Regarding computational time, we observe that the effort required for solving (8) is fairly independent of the number of linear inequalities p but depends primarily on the dimension of the projected and of the full space, n_x and n, respectively. On the contrary, solving (7) primarily depends on p and not on n_x/n, which is intuitive because the number of constraints p directly impacts the number of binding scenarios involved in (7). Tables EC.7 and EC.8 in Online Appendix EC.4.3 summarize the average computational time required for both problems for varying input sizes.

Figure 6. (Color online) Distribution (Box Plot) of the Gap $(\lambda^*_{HGK} - \lambda^*_{LDR})/\lambda^*_{HGK}$ for Varying Values of n_x/n and Varying Values of n



5. Intersection of Two Ellipsoids

For $i = 1, 2$, we define the ellipsoid $\mathcal{E}_i = \{x \in \mathbb{R}^n \mid \|\boldsymbol{A}_i(\boldsymbol{x} - \boldsymbol{x}_i)\| \le 1\}$, where $\boldsymbol{x}_i \in \mathbb{R}^n$ and $\boldsymbol{A}_i \in \mathbb{R}^{n\times n}$. We are interested in computing a Minkowski center of the intersection of these two ellipsoids, $\mathcal{E}_1 \cap \mathcal{E}_2$. In this case, we make an additional assumption on the matrices $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$.

Assumption 1. *There exists an invertible matrix* $\boldsymbol{P}$ *such that, for* $i = 1, 2$, $\boldsymbol{A}_i^\top \boldsymbol{A}_i = \boldsymbol{P}^\top \boldsymbol{D}_i \boldsymbol{P}$ *for some diagonal matrix* $\boldsymbol{D}_i = \mathrm{diag}(\boldsymbol{d}_i)$.

When matrices $\boldsymbol{A}_1^\top \boldsymbol{A}_1$ and $\boldsymbol{A}_2^\top \boldsymbol{A}_2$ satisfy Assumption 1, we say that they are "diagonalized simultaneously by a congruence relationship" (Uhlig 1973). For instance, Assumption 1 is satisfied whenever one of the matrices $\boldsymbol{A}_i^\top \boldsymbol{A}_i$ is nonsingular (Uhlig 1973, theorem 0.2). After a proper change of variable, $\boldsymbol{w} \leftarrow \boldsymbol{P}\boldsymbol{w}$ and $\boldsymbol{y} \leftarrow \boldsymbol{P}\boldsymbol{y}$, we can assume, without further loss of generality, that the matrices $\boldsymbol{A}_i$ are diagonal, that is, that we have

$$\mathcal{E}_i = \{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{D}_i^{1/2}(\boldsymbol{x} - \boldsymbol{x}_i)\| \le 1\},$$

where $\boldsymbol{D}_i^{1/2} = \mathrm{diag}(\sqrt{d_1}, \dots \sqrt{d_n})$. Let us denote $\boldsymbol{b}_i := \boldsymbol{D}_i \boldsymbol{x}_i$ and $c_i := \boldsymbol{x}_i^\top \boldsymbol{D}_i \boldsymbol{x}_i$. The objective of this section is to propose an efficient approach, based on second-order cone relaxation and bisection search, to obtain a lower bound on $\mathrm{sym}(\mathcal{E}_1 \cap \mathcal{E}_2)$ together with an approximate Minkowski center. We also provide conditions (that can be numerically verified) under which our proposed approximation is tight.

Remark 6. Assumption 1 is a much weaker assumption than simultaneous diagonalizability, that is, $\boldsymbol{A}_i^\top \boldsymbol{A}_i = \boldsymbol{P}^{-1} \boldsymbol{D}_i \boldsymbol{P}$, $i = 1, 2$. If $\boldsymbol{A}_i^\top \boldsymbol{A}_i$, $i = 1, 2$, are simultaneously diagonalizable, then Assumption 1 is satisfied. The reverse implication is not true.

5.1. Second-Order Cone Approximation

We start by reformulating the optimization problem defining Minkowski centers of $\mathcal{E}_1 \cap \mathcal{E}_2$.

Lemma 5. *For* $\mathcal{C} = \mathcal{E}_1 \cap \mathcal{E}_2$, *Problem (2) is equivalent to*

$$\begin{aligned}
\max_{\substack{\boldsymbol{w}, \boldsymbol{\xi}, \lambda \ge 0\\ \boldsymbol{\eta}^\star}} \quad & \lambda \\
\text{s.t.} \quad & \boldsymbol{d}_i^\top \boldsymbol{\xi} - 2\boldsymbol{b}_i^\top \boldsymbol{w} + (1+\lambda)c_i \le (1+\lambda), \quad \forall i \in \{1,2\},\\
& w_j^2 \le (1+\lambda)\xi_j, \quad \forall j \in [n],\\
& \|\boldsymbol{D}_i^{1/2}(\boldsymbol{w} - \boldsymbol{x}_i)\|_2^2 + \eta_i^\star(\boldsymbol{w}, \boldsymbol{\lambda}) \le 1,\ \forall i \in \{1,2\},
\end{aligned}$$

where each $\eta_i^\star(\boldsymbol{w}, \lambda)$, $i = 1, 2$, *is the objective value of a nonconvex quadratic optimization problem:*

$$\begin{aligned}
\eta_i^\star(\boldsymbol{w}, \lambda) = \max_{\boldsymbol{y}, \boldsymbol{z}} \quad & \lambda^2 \boldsymbol{d}_i^\top \boldsymbol{z} - 2\lambda(\boldsymbol{w} - \boldsymbol{x}_i)^\top \boldsymbol{D}_i \boldsymbol{y}\\
\text{s.t.} \quad & \boldsymbol{d}_k^\top \boldsymbol{z} - 2\boldsymbol{b}_k^\top \boldsymbol{y} \le 1 - c_k,\ \ \forall k \in \{1,2\},\\
& y_j^2 = z_j,\ \ \forall j \in [n].
\end{aligned} \tag{9}$$

The proof of Lemma 5 relies on simple algebraic manipulations on Problem (2) and is hence deferred to Online Appendix EC.5.

The maximization problem defining $\eta_i^\star$ is not convex due to the quadratic equality constraints $z_j = y_j^2$. Instead, we now propose a valid convex upper bound on $\eta_i^\star$, under constraint qualification conditions.

Assumption 2. *There exists* $\boldsymbol{x} \in \mathbb{R}^n$ *such that, for all* $i \in \{1,2\}$, $\|\boldsymbol{D}_i^{1/2}(\boldsymbol{x} - \boldsymbol{x}_i)\| < 1$.

In other words, we assume that $\mathcal{E}_1 \cap \mathcal{E}_2$ has a nonempty relative interior.

Lemma 6. *Under Assumption 2, for each* $i \in \{1,2\}$, *we have* $\eta_i^\star(\boldsymbol{w}, \lambda) \le \eta_i(\boldsymbol{w}, \lambda)$ *with*

$$\begin{aligned}
\eta_i(\boldsymbol{w}, \lambda) = \min_{\boldsymbol{u} \in \mathbb{R}_+^n, \boldsymbol{v} \in \mathbb{R}_+^2, \boldsymbol{\theta} \in \mathbb{R}_+^n} \quad & v_1(1 - c_1) + v_2(1 - c_2) + \boldsymbol{e}^\top \boldsymbol{\theta}\\
\text{s.t.} \quad & \lambda^2 \boldsymbol{d}_i - v_1 \boldsymbol{d}_1 + v_2 \boldsymbol{d}_2 + \boldsymbol{u} \le \boldsymbol{0},\\
& \left(v_1 b_{1,j} + v_2 b_{2,j} - \lambda d_{i,j}(w_j - x_{i,j})\right)^2 \le u_j \theta_j,\\
& \qquad \forall j \in [n].
\end{aligned} \tag{10}$$

Proof. Fix $i \in \{1,2\}$. Relaxing the constraint $y_j^2 = z_j$ into the second-order cone constraints $y_j^2 \le z_j$ leads to $\eta_i^\star(\boldsymbol{w}, \lambda) \le \eta_i(\boldsymbol{w}, \lambda)$ with

$$\begin{aligned}
\eta_i(\boldsymbol{w}, \lambda) = \max_{\boldsymbol{y}, \boldsymbol{z}} \ & \lambda^2 \boldsymbol{d}_i^\top \boldsymbol{z} - 2\lambda(\boldsymbol{w} - \boldsymbol{x}_i)^\top \boldsymbol{D}_i \boldsymbol{y}\\
\text{s.t.} \quad & \boldsymbol{d}_k^\top \boldsymbol{z} - 2\boldsymbol{b}_k^\top \boldsymbol{y} \le 1 - c_k,\ \ \forall k \in \{1,2\} \quad [\boldsymbol{v}]\\
& y_j^2 \le z_j,\ \ \forall j \in [n]. \quad [\boldsymbol{u}]
\end{aligned} \tag{11}$$

By introducing dual variables $(\boldsymbol{v}, \boldsymbol{u})$ for the constraints in (11), we have that

$$\eta_i(\boldsymbol{w}, \lambda) = \max_{\boldsymbol{y}, \boldsymbol{z} \ge \boldsymbol{0}} \ \min_{\boldsymbol{u} \in \mathbb{R}_+^n, \boldsymbol{v} \in \mathbb{R}_+^2} \ \mathcal{L}(\boldsymbol{y}, \boldsymbol{z}; \boldsymbol{u}, \boldsymbol{v}),$$

where $\mathcal{L}$ is the Lagrangian of the problem and is defined as

$$\begin{aligned}
\mathcal{L}(\boldsymbol{y}, \boldsymbol{z}; \boldsymbol{u}, \boldsymbol{v}) = {} & \lambda^2 \boldsymbol{d}_i^\top \boldsymbol{z} - 2\lambda(\boldsymbol{w} - \boldsymbol{x}_i)^\top \boldsymbol{D}_i \boldsymbol{y}\\
& + \sum_{k=1}^{2} v_k(1 - c_k - \boldsymbol{d}_k^\top \boldsymbol{z} + 2\boldsymbol{b}_k^\top \boldsymbol{y}) + \sum_{j \in [L]} u_j(z_j - y_j^2).
\end{aligned}$$

Assumption 2 implies that there exists a strictly feasible solution to (11). Hence, strong duality holds and we can invert the order of the maximization and minimization. For a fixed $(\boldsymbol{u}, \boldsymbol{v})$, by partially maximizing with respect to $\boldsymbol{z}$, we obtain

$$\max_{\boldsymbol{z} \ge \boldsymbol{0}} \left(\lambda^2 \boldsymbol{d}_i - \sum_{k=1}^{2} v_k \boldsymbol{d}_k + \boldsymbol{u}\right)^\top \boldsymbol{z} = \begin{cases} 0 & \text{if } \lambda^2 \boldsymbol{d}_i - v_1 \boldsymbol{d}_1 - v_2 \boldsymbol{d}_2 + \boldsymbol{u} \le \boldsymbol{0},\\ +\infty & \text{otherwise.} \end{cases}$$

Observe that for any $a, u \in \mathbb{R}$,

$$\max_y \{ay - uy^2\} = \begin{cases} \dfrac{a^2}{4u} & \text{if } u > 0 \\ +\infty & \text{otherwise} \end{cases} = \min_\theta \{\theta \text{ s.t. } a^2 \le 4\theta u\}.$$

Therefore, maximizing with respect to y_j,

$$\max_{y_j} 2\boldsymbol{e}_j^\top \left(-\lambda \boldsymbol{D}_i(\boldsymbol{w} - \boldsymbol{x}_i) + \sum_{k=1}^{2} v_k \boldsymbol{b}_k \right) y_j - u_j y_j^2,$$

is equivalent to minimizing θ_j subject to the constraint detailed in the final formulation. □

Remark 7. Our approach could be generalized to matrices not satisfying Assumption 1. In this case, however, (9) would involve the additional variables $\boldsymbol{Z} : Z_{i,j} = y_i y_j$ and its convex relaxation (11) would be a semidefinite optimization problem (instead of second-order cone) over $(\boldsymbol{y}, \boldsymbol{Z}) : \boldsymbol{Z} \succeq \boldsymbol{y}\boldsymbol{y}^\top$. Hence, Assumption 1 substantially improves computational tractability without great loss of generality in the case of two matrices.

5.2. Final Formulation and Numerical Algorithm

Overall, an approximate Minkowski center for $\mathcal{E}_1 \cap \mathcal{E}_2$ can be obtained by solving

$$\begin{aligned} \max_{\substack{\boldsymbol{w}, \boldsymbol{\xi}, \lambda \\ (\eta_i, v_{1,i}, v_{2,i}, \boldsymbol{u}_i, \boldsymbol{\theta}_i)_{i=1,2}}} \quad & \lambda \\ \text{s.t.} \quad & \boldsymbol{d}_i^\top \boldsymbol{\xi} - 2\boldsymbol{b}_i^\top \boldsymbol{w} + (1+\lambda)c_i \le (1+\lambda), \ \forall i \in \{1,2\}, \\ & w_j^2 \le (1+\lambda)\xi_j, \ \forall j \in [n], \\ & \|\boldsymbol{D}_i(\boldsymbol{w} - \boldsymbol{x}_i)\|_2^2 + \eta_i \le 1, \ \forall i \in \{1,2\}, \\ & v_{1,i}(1 - c_1) + v_{2,i}(1 - c_2) + \boldsymbol{e}^\top \boldsymbol{\theta}_i \le \eta_i, \ \forall i \in \{1,2\}, \\ & \lambda^2 \boldsymbol{d}_i - v_{1,i}\boldsymbol{d}_1 - v_{2,i}\boldsymbol{d}_2 + \boldsymbol{u}_i \le \boldsymbol{0}, \ \forall i \in \{1,2\}, \\ & \left(v_{1,i} b_{1,j} + v_{2,i} b_{2,j} - \lambda d_{i,j}(w_j - x_{i,j}) \right)^2 \le u_j \theta_j, \\ & \qquad \forall i \in \{1,2\}, j \in [n], \\ & \boldsymbol{\xi}, \boldsymbol{u}_i, \boldsymbol{\theta}_i \ge \boldsymbol{0}, \\ & \lambda, v_{1,i}, v_{2,i} \ge 0. \end{aligned} \tag{12}$$

In this formulation, the variables η_i satisfy $\eta_i \ge \eta_i^\star(\boldsymbol{w}, \lambda)$ so any solution $(\boldsymbol{w}, \lambda)$ feasible for (12) is feasible for the original problem and solving (12) provides a lower bound on $\operatorname{sym}(\mathcal{E}_1 \cap \mathcal{E}_2)$. Solving (12) is challenging, however, because of the bilinear product of decision variables $\lambda d_{i,j}(w_j - x_{i,j})$ in the constraints. To do so efficiently, we propose to conduct a bisection search over λ. Indeed, $\lambda \in [0,1]$ and, for a fixed λ, (12) is a second-order cone optimization problem. Consequently, we can obtain an ϵ-approximation of the objective value of (12) after solving $\log_2(\varepsilon)$ second-order cone optimization problems.

5.3. Tightness

In this section, we fix $i \in \{1,2\}$ and analyze the tightness of the relaxation $\eta_i(\boldsymbol{w}, \lambda)$. First, we provide a (numerically verifiable) condition for our relaxation to be tight.

Proposition 10. *Fix $i \in \{1,2\}$. Let $(\boldsymbol{y}^\star, \boldsymbol{z}^\star, \boldsymbol{u}^\star, \boldsymbol{v}^\star)$ be a primal-dual optimal pair of (11)-(10). If $v_1^\star v_2^\star = 0$, then $\eta_i^\star(\boldsymbol{w}, \lambda) = \eta_i(\boldsymbol{w}, \lambda)$.*

Proof. The result is a special case of Ben-Tal and den Hertog (2014, theorem 7) after noting that assumption 5 in Ben-Tal and den Hertog (2014) is automatically satisfied in our case and that their assumption 6 is equivalent to the condition $v_1^\star v_2^\star = 0$. □

Second, we show that $\eta_i(\boldsymbol{w}, \lambda)$ provides a constant factor approximation on $\eta_i^\star(\boldsymbol{w}, \lambda)$ under the additional assumption that $\boldsymbol{0}$ lies in the relative interior of $\mathcal{E}_1 \cap \mathcal{E}_2$.

Proposition 11. *Fix $i \in \{1,2\}$. Further assume that Assumption 2 is satisfied for $\boldsymbol{x} = \boldsymbol{0}$. Then,*

$$\eta_i^\star(\boldsymbol{w}, \lambda) \ge \left(\frac{1 - \gamma}{\sqrt{2} + \gamma} \right)^2 \eta_i(\boldsymbol{w}, \lambda),$$

where $\gamma = \max_k \|\boldsymbol{D}_k^{1/2} \boldsymbol{x}_k\| = \max_k \sqrt{c_k} < 1$.

The proof of Proposition 11 relies on a similar construction as in Xia et al. (2021, theorem 8). However, Xia et al. (2021) consider the special case of spheres, that is, $d_{k,j} = 1$ for all $k \in \{1,2\}, j \in [p]$. We extend their proof technique to the nonisotropic case (see details in Online Appendix EC.5) after making the following observation.

Lemma 7. *There exists an optimal solution of (11), $(\boldsymbol{y}^\star, \boldsymbol{z}^\star)$, such that, for any $j \in [p]$,*

$$(y_j^\star)^2 < z_j^\star \Rightarrow d_{i,j} > 0.$$

Proof. Let $(\boldsymbol{y}^\star, \boldsymbol{z}^\star)$ be an optimal solution of (11). Define $\mathcal{J} := \{j \in [p] \,|\, (y_j^\star)^2 < z_j^\star\}$. We assume there exists $j \in \mathcal{J}$ such that $d_{i,j} = 0$. Let us define $\bar{\boldsymbol{y}} = \boldsymbol{y}^\star$ and

$$\bar{z}_{j'} = \begin{cases} z_{j'}^\star & \text{if } j' \ne j, \\ (y_j^\star)^2 & \text{if } j' = j. \end{cases}$$

Then, $(\bar{\boldsymbol{y}}, \bar{\boldsymbol{z}})$ satisfies

$$\begin{aligned} \lambda^2 \boldsymbol{d}_i^\top \bar{\boldsymbol{z}} - 2(\boldsymbol{w} - \boldsymbol{x}_i)^\top \boldsymbol{D}_i \bar{\boldsymbol{y}} + &= -2(\boldsymbol{w} - \boldsymbol{x}_i)^\top \boldsymbol{D}_i \boldsymbol{y}^\star + \lambda^2 \boldsymbol{d}_i^\top \boldsymbol{z}^\star, \\ \boldsymbol{d}_k^\top \bar{\boldsymbol{z}} - 2\boldsymbol{b}_k^\top \bar{\boldsymbol{y}} &\le \boldsymbol{d}_k^\top \boldsymbol{z}^\star - 2\boldsymbol{b}_k^\top \boldsymbol{y}^\star \le 1 - c_k, \ \forall k \in \{1,2\}. \end{aligned}$$

In other words, $(\bar{\boldsymbol{y}}, \bar{\boldsymbol{z}})$ is feasible and optimal for (11) and $\{j \in [p] \,|\, (\bar{y}_j)^2 < \bar{z}_j\} = \mathcal{J} \setminus \{j\}$. □

5.4. Discussion: Extension to the Intersection of $m \geq 2$ Ellipsoids

The approach we outlined in this section could be extended to the intersection of $m \geq 2$ ellipsoids. However, the conditions for Assumption 1 are more stringent in this case and overly restrictive (Grimus and Ecker 1986). Consequently, as mentioned in Remark 7, our approach in the case of m ellipsoids would entail relaxing each nonconvex Problem (9) into a semidefinite optimization problem, similar to the approach of Eldar et al. (2008) for the Chebyshev center of an intersection of ellipsoids. Eventually, the resulting formulation would be a semidefinite optimization problem with m $(n \times n)$ semidefinite matrices to optimize over, analogous to the one described in Ben-Tal et al. (2009, chapter 7.2.1). Alternatively, one could follow the approach developed in Bertsimas et al. (2022) to derive safe approximation in the case of m ellipsoids of the form $\mathcal{E}_i = \{\boldsymbol{y} : \|\boldsymbol{A}_i(\boldsymbol{y} - \boldsymbol{x}_i)\| \leq 1\}$. The resulting safe approximation would involve an additional uncertain parameter, $\boldsymbol{V} \in \mathbb{R}^{n \times n}$, with bounded singular values. Again, the resulting robust counterpart is a semidefinite optimization problem that can be approximated by a second-order cone problem by bounding the matrix 2-norm by the Frobenius norm. Their approach could be also applied to derive approximate Minkowski center of arbitrary convex sets.

6. Conclusion

This paper provides a robust optimization formulation for the Minkowski centers of convex sets. Building up on this formulation, we propose tractable reformulations and efficient approximation techniques to numerically compute the Minkowski centers of a variety of sets (polyhedra, convex hulls, projections of polyhedra, intersections of ellipsoids). Theoretical benefits of Minkowski centers are numerous and well documented: They are geometrically defined and do not depend on the analytic description of the set (unlike the analytic center). Moreover, they naturally adapt to the dimension of the convex set and do not require the set to be fully dimensional (unlike centers of extremal ellipsoids such as Chebyshev centers). In addition, we illustrate their computational appeal by analyzing the algorithmic convergence of hit-and-run and cutting-plane method examples. Although the actual gains ultimately depend on the particular algorithm and instance at hand, we believe our work sheds new and practical light on Minkowski centers and exposes their potential benefits as a computational tool.

References

Bandi C, Bertsimas D (2012) Tractable stochastic analysis in high dimensions via robust optimization. *Math. Programming* 134(1):23–70.

Barnes ER, Moretti AC (2005) Some results on centers of polytopes. *Optim. Methods Software* 20(1):9–24.

Bélisle C, Boneh A, Caron RJ (1998) Convergence properties of hit–and–run samplers. *Stochastic Models* 14(4):767–800.

Bélisle CJP, Romeijn HE, Smith RL (1993) Hit-and-run algorithms for generating multivariate distributions. *Math. Oper. Res.* 18(2):255–266.

Belloni A, Freund RM (2007) On the symmetry function of a convex set. *Math. Programming* 111(1–2):57–93.

Ben-Ameur W, Ouorou A, Wang G, Żotkiewicz M (2018) Multipolar robust optimization. *EURO J. Comput. Optim.* 6(4):395–434.

Ben-Tal A, den Hertog D (2014) Hidden conic quadratic representation of some nonconvex quadratic optimization problems. *Math. Programming* 143(1):1–29.

Ben-Tal A, den Hertog D, Vial JP (2015) Deriving robust counterparts of nonlinear uncertain inequalities. *Math. Programming* 149(1):265–299.

Ben-Tal A, El Ghaoui L, Nemirovski A (2009) *Robust Optimization* (Princeton University Press, Princeton, NJ).

Ben-Tal A, Goryashko A, Guslitzer E, Nemirovski A (2004) Adjustable robust solutions of uncertain linear programs. *Math. Programming* 99(2):351–376.

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

Bertsimas D, Sim M (2004) The price of robustness. *Oper. Res.* 52(1):35–53.

Bertsimas D, Brown DB, Caramanis C (2011a) Theory and applications of robust optimization. *SIAM Rev.* 53(3):464–501.

Bertsimas D, Dunning I, Lubin M (2016) Reformulation vs. cutting-planes for robust optimization. *Comput. Management Sci.* 13(2):195–217.

Bertsimas D, Goyal V, Lu BY (2015) A tight characterization of the performance of static solutions in two-stage adjustable robust linear optimization. *Math. Programming* 150(2):281–319.

Bertsimas D, Goyal V, Sun XA (2011b) A geometric characterization of the power of finite adaptability in multistage stochastic and adaptive optimization. *Math. Oper. Res.* 36(1):24–54.

Bertsimas D, Iancu DA, Parrilo PA (2010) Optimality of affine policies in multistage robust optimization. *Math. Oper. Res.* 35(2):363–394.

Bertsimas D, Hertog DD, Pauphilet J, Zhen J (2022) Robust convex optimization: A new perspective that unifies and extends. *Math. Programming* 1–42. https://link.springer.com/article/10.1007/s10107-022-01881-w.

Bertsimas D, Litvinov E, Sun XA, Zhao J, Zheng T (2012) Adaptive robust optimization for the security constrained unit commitment problem. *IEEE Trans. Power Systems* 28(1):52–63.

Bland RG, Goldfarb D, Todd MJ (1981) The ellipsoid method: A survey. *Oper. Res.* 29(6):1039–1091.

Boyd SP, Vandenberghe L (2004) *Convex Optimization* (Cambridge University Press, Cambridge, UK).

Boyd SP, Vandenberghe L, Skaf J (2008) *Analytic Center Cutting-Plane Method.* Lecture Notes from Stanford University (Stanford University, Stanford, CA).

Caron RJ, Greenberg HJ, Holder AG (2002) Analytic centers and repelling inequalities. *Eur. J. Oper. Res.* 143(2):268–290.

Chen MH, Schmeiser B (1993) Performance of the Gibbs, hit-and-run, and metropolis samplers. *J. Comput. Graphics Statist.* 2:251–272.

Conforti M, Cornuéjols G, Zambelli G (2010) Extended formulations in combinatorial optimization. *4OR* 8(1):1–48.

Díaz J, Cuevas A, Grande F (2006) Testing multivariate uniformity: The distance-to-boundary method. *Canadian J. Statist.* 34(4):693–707.

Eldar YC, Beck A, Teboulle M (2008) A minimax Chebyshev estimator for bounded error estimation. *IEEE Trans. Signal Processing* 56(4):1388–1397.

Elhedhli S, Goffin JL, Vial JP (2009) *Nondifferentiable Optimization: Cutting Plane Methods* (Springer, Boston).

Gay DM (1985) Electronic mail distribution of linear programming test problems. *Math. Programming Soc. COAL Newsletter* 13:10–12.

Gounaris CE, Repoussis PP, Tarantilis CD, Wiesemann W, Floudas CA (2016) An adaptive memory programming framework for the robust capacitated vehicle routing problem. *Transportation Sci.* 50(4):1239–1260.

Grimus W, Ecker G (1986) On the simultaneous diagonalizability of matrices. Technical report, Vienna University, Vienna.

Güler O, Gürtuna F (2012) Symmetry of convex sets and its applications to the extremal ellipsoids of convex bodies. *Optim. Methods Software* 27(4–5):735–759.

Guslitser E (2002) Uncertainty-immunized solutions in linear programming. MS thesis, Technion, Israeli Institute of Technology, IE&M Faculty, Haifa, Israel.

Hadjiyiannis MJ, Goulart PJ, Kuhn D (2011) A scenario approach for estimating the suboptimality of linear decision rules in two-stage robust optimization. *Proc. 50th IEEE Conf. on Decision and Control and Eur. Control Conf.* (IEEE, Piscataway, NJ), 7386–7391.

Hammer PC (1951) The centroid of a convex body. *Proc. Amer. Math. Soc.* 2(4):522–525.

Housni OE, Goyal V (2021) On the optimality of affine policies for budgeted uncertainty sets. *Math. Oper. Res.* 46(2):674–711.

Huard P (1967) Resolution of mathematical programming with nonlinear constraints by the method of centers. Abadie J, ed. *Nonlinear Programming* (Springer, Basel), 207–219.

Iancu DA, Trichakis N (2014) Pareto efficiency in robust optimization. *Management Sci.* 60(1):130–147.

Jarre F (1989) On the method of analytic centers for solving smooth convex programs. *Optim. - Fifth French-German Conf. Caster Novel 1988,* Lecture Notes in Mathematics 1405 (Springer Verlag, Berlin), 69–85.

John F (1948) Extreme problems with inequalities as subsidiary conditions, studies and essays. *Traces and Emergence of Nonlinear Programming* (Springer, Basel), 187–204.

Kanungo T, Mount DM, Netanyahu NS, Piatko CD, Silverman R, Wu AY (2002) An efficient k-means clustering algorithm: Analysis and implementation. *IEEE Trans. Pattern Anal. Machine Intelligence* 24(7):881–892.

Karmarkar N (1984) A new polynomial-time algorithm for linear programming. *Proc. 16th Annual ACM Sympos. on Theory of Comput.* (Association for Computing Machinery, New York), 302–311.

Kellerer H, Pferschy U, Pisinger D (2004) *Knapsack Problems* (Springer, Berlin).

Klee V (1963) Convexity. Proc. 7th Sympos. in Pure Mathematics of the Amer. Math. Soc., vol. 7 (American Mathematical Society, Providence, RI).

Lovász L (1999) Hit-and-run mixes fast. *Math. Programming* 86(3): 443–461.

Minkowski H (1911) Allegemeine lehzätze über konvexe polyeder. *Gesellschaft Abhandlungen Leipzog-Berlin* 1:103–121.

Motzkin TS (1936) *Beiträge zur Theorie der Linearen Ungleichungen* (Azriel Press, New York).

Rademacher LA (2007) Approximating the centroid is hard. *Proc. 23rd Annual Sympos. on Comput. Geometry* (Association for Computing Machinery, New York), 302–305.

Renegar J (1988) A polynomial-time algorithm, based on Newton's method, for linear programming. *Math. Programming* 40(1): 59–93.

Roos C, Terlaky T, Vial JP (2005) *Interior Point Methods for Linear Optimization* (Springer Science & Business Media, Boston).

Schwartz JT, Sharir M (1988) A survey of motion planning and related geometric algorithms. *Artificial Intelligence* 37(1–3):157–169.

Smith RL (1984) Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. *Oper. Res.* 32(6):1296–1308.

Sonnevend G (1986) An "analytical center" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. *System Modelling and Optimization* (Springer, Berlin), 866–875.

Tarasov SP, Khachiyan LG, Erlikh II (1988) The method of inscribed ellipsoids. *Soviet Math. Doklady* 37(1):226–230.

Todd MJ (1982) On minimum volume ellipsoids containing part of a given ellipsoid. *Math. Oper. Res.* 7(2):253–261.

Uhlig F (1973) Definite and semidefinite matrices in a real symmetric matrix pencil. *Pacific J. Math.* 49(2):561–568.

Vaidya PM (1996) A new algorithm for minimizing convex functions over convex sets. *Math. Programming* 73(3):291–341.

Xia Y, Yang M, Wang S (2021) Chebyshev center of the intersection of balls: Complexity, relaxation and approximation. *Math. Programming* 187(1–2):287–315.

Yudin DB, Nemirovskii AS (1976) Informational complexity and efficient methods for the solution of convex extremal problems. *Matekon* 13(2):22–45.

Zeng B, Zhao L (2013) Solving two-stage robust optimization problems using a column-and-constraint generation method. *Oper. Res. Lett.* 41(5):457–461.

Zhen J, den Hertog D (2018) Computing the maximum volume inscribed ellipsoid of a polytopic projection. *INFORMS J. Comput.* 30(1):31–42.

Zhen J, den Hertog D, Sim M (2018) Adjustable robust optimization via Fourier-Motzkin elimination. *Oper. Res.* 66(4):1086–1100.

Dick den Hertog is a professor of operations research at the University of Amsterdam and an expert in mathematical optimization, with a focus on robust optimization. He is also active in applying optimization for a better society. He is the cofounder and science-to-impact codirector of Analytics for a Better World and received the INFORMS Franz Edelman Award in 2013 and 2021.

Jean Pauphilet is an assistant professor in the Management Science and Operations group at London Business School. His research sits at the intersection of discrete optimization, robust optimization, and machine learning, with applications to healthcare and sustainable operations. He is the recipient of several research awards, including the INFORMS George Nicholson Prize.

Mohamed Yahya Soali is a PhD candidate in the Industrial Engineering and Operations Research Department at Columbia University, advised by Prof. Vineet Goyal. His research interests include robust optimization and revenue management.