Chapter 1
The method of cyclic intrepid projections: convergence analysis and numerical experiments

Heinz H. Bauschke, Francesco Iorio, and Valentin R. Koch

Abstract The convex feasibility problem asks to find a point in the intersection of a collection of nonempty closed convex sets. This problem is of basic importance in mathematics and the physical sciences, and projection (or splitting) methods solve it by employing the projection operators associated with the individual sets to generate a sequence which converges to a solution. Motivated by an application in road design, we present the method of cyclic intrepid projections (CycIP) and provide a rigorous convergence analysis. We also report on very promising numerical experiments in which CycIP is compared to a commerical state-of-the-art optimization solver.

Key words: Convex set, feasibility problem, halfspace, intrepid projection, linear inequalities, projection, road design.
AMS 2010 Subject Classification: Primary 65K05, 90C25; Secondary 90C05.

1.1 Introduction

Throughout this paper, we assume that

$$X \text{ is a real Hilbert space} \tag{1.1}$$

Heinz H. Bauschke
Mathematics, University of British Columbia, Kelowna, BC V1V 1V7, Canada, e-mail: heinz.bauschke@ubc.ca

Francesco Iorio
Autodesk Research, 210 King Street East, Suite 600, Toronto, ON M5A 1J7, Canada, e-mail: francesco.iorio@autodesk.com

Valentin R. Koch
Information Modeling & Platform Products Group (IPG), Autodesk, Inc., e-mail: valentin.koch@autodesk.com

with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$. (We also write $\|\cdot\|_2$ instead of $\|\cdot\|$ if we wish to emphasize this norm compared to other norms. We assume basic notation and results from convex analysis and fixed point theory; see, e.g., [4, 6, 15, 16, 21].)

Let $(C_i)_{i\in I}$ be a finite family of closed convex subsets of X such that

$$C := \bigcap_{i\in I} C_i \neq \varnothing. \tag{1.2}$$

We aim to find a point in C given that the individual constraint sets C_i are simple in the sense that their associated projections[1] are easy to compute. To solve the widespread *convex feasibility problem* "find $x \in C$", we employ *projection methods*. These splitting-type methods use the individual projections P_{C_i} in order to generate a sequence that converges to a point in C. For further information, we refer the reader to, e.g., [2, 4, 8, 9, 11, 15, 16, 19].

In previous work on a feasibility problem arising in road design [5], the *method of cyclic intrepid projections (CycIP)* was found to be an excellent overall algorithm. Unfortunately, CycIP was applied heuristically without an underlying convergence result.

The goal of this paper is two-fold. First, we present a checkable condition sufficient for convergence and provide a rigorous convergence proof. In fact, our main result applies to very general feasibility problems satisfying an interiority assumption. Second, we numerically compare CycIP to a commercial LP solver for test problems that are both convex and nonconvex to evaluate competitiveness of CycIP.

The remainder of the paper is organized as follows. In Section 1.2 we provide basic properties of projection operators. Useful results on Fejér monotone sequences are recalled in Section 1.3. Our main convergence results are presented in Section 1.4. In Section 1.5, we review the feasibility problem arising in road design and obtain a rigorous convergence result for CycIP. We report on numerical experiments in Section 1.6 and offer concluding remarks in Section 1.7.

We end this section with notation. The closed ball centered at $y \in X$ of radius r is $B(y;r) := \{z \in X \mid \|y-z\| \leq r\}$. Finally, we write $\mathbb{R}_+$ and $\mathbb{R}_{++}$ for the nonnegative real numbers and strictly positive reals, respectively.

1.2 Relaxed and intrepid projectors

In this section, we introduce the key operators used in the projection methods studied later.

Fact 1 (relaxed projector). Let C be a nonempty closed convex subset of X, and let $\lambda \in \,]0,2[$. Set $R := (1-\lambda)\,\mathrm{Id} + \lambda P_C$, let $x \in X$, and let $c \in C$. Then

---

[1] Given a nonempty subset S of X and $x \in X$, we write $d_S(x) := \inf_{s\in S}\|x-s\|$ for the *distance* from x to S. If S is also closed and convex, then the infimum defining $d_S(x)$ is attained at a *unique* vector called the *projection* of x onto S and denoted by $P_S(x)$ or $P_S x$.

$$\|x-c\|^2 - \|Rx-c\|^2 \geq \frac{2-\lambda}{\lambda}\|x-Rx\|^2 = (2-\lambda)\lambda d_C^2(x). \tag{1.3}$$

Proof. Combine [2, Lemma 2.4.(iv)] with [4, Proposition 4.8]. □

In fact, the relaxed projector is an example of a so-called *averaged map*; see, e.g., [1, 4, 13] for more on this useful notion.

Definition 2 (enlargement). Given a nonempty closed convex subset Z of X, and $\alpha \in \mathbb{R}_+ := \big\{\xi \in \mathbb{R} \mid \xi \geq 0\big\}$, we write

$$Z_{[\alpha]} := \big\{x \in X \mid d_Z(x) \leq \alpha\big\} = Z + B(0;\alpha) \tag{1.4}$$

and call $Z_{[\alpha]}$ the α*-enlargement* of Z.

Note that $Z_{[0]} = Z$, that $Z_{[\alpha]}$ is a nonempty closed convex subset of X, and that if $\alpha < \beta$, then $Z_{[\alpha]} \subseteq Z_{[\beta]}$. We mention in passing that the *depth*[2] of each $z \in Z$ (with respect to $Z_{[\alpha]}$), i.e., $d_{X \smallsetminus Z_{[\alpha]}}(z)$, is at least α.

Fact 3. (See, e.g., [4, Proposition 28.10].) Let C be a nonempty closed convex subset of X, and let $\beta \in \mathbb{R}_+$. Set $D := C_{[\beta]}$. Then

$$(\forall x \in X) \quad P_D x = \begin{cases} x, & \text{if } d_C(x) \leq \beta; \\ P_C x + \beta \dfrac{x - P_C x}{d_C(x)}, & \text{otherwise.} \end{cases} \tag{1.5}$$

Definition 4 (intrepid projector). Let Z be a nonempty closed convex subset of X, let $\beta \in \mathbb{R}_+$, and set $C := Z_{[\beta]}$. The corresponding *intrepid projector* $Q := Q_C$ *onto* C *(with respect to* Z *and* β*)* is defined by

$$\begin{aligned} Q\colon X \to X\colon x \mapsto x + \Big(1 - \tfrac{P_{[\beta,2\beta]} d_Z(x)}{\beta}\Big)\big(x - P_Z x\big) \\ = \begin{cases} P_Z x, & \text{if } d_Z(x) \geq 2\beta; \\ x, & \text{if } d_Z(x) \leq \beta; \\ x + \Big(1 - \tfrac{d_Z(x)}{\beta}\Big)\big(x - P_Z x\big), & \text{otherwise.} \end{cases} \end{aligned} \tag{1.6}$$

We refer to these three steps as the *projection step*, the *identity step*, and the *reflection step*, respectively.

Example 5 (intrepid projector onto a hyperslab à la Herman). Suppose that $a \in X \smallsetminus \{0\}$, let $\alpha \in \mathbb{R}$, let $\beta \in \mathbb{R}_+$, and set $Z := \big\{x \in X \mid \langle a, x\rangle = \alpha\big\}$. Then Z is a hyperplane and $Z_{[\beta]}$ is a hyperslab. Moreover, the associated intrepid projector onto $Z_{[\beta]}$ is precisely the operator considered by Herman in [18].

Proposition 6 (basic properties of the intrepid projector). Let Z be a nonempty closed convex subset of X, let $\beta \in \mathbb{R}_+$, set $C := Z_{[\beta]}$, and denote the corresponding

---

[2] This function is considered, e.g., in [7, Exercise 8.5].

intrepid projector onto C (with respect to Z and β) by Q. Now let $\alpha \in [0,\beta]$, and let $y \in Z_{[\alpha]}$, and let $x \in X$. Then

$$Qx \in [x, P_Z x] \cap C \tag{1.7}$$

and exactly one of the following holds:

(i). $d_Z(x) \le \beta$, $x = Qx \in C$, and $\|x-y\|^2 - \|Qx-y\|^2 = 0$.
(ii). $d_Z(x) \ge 2\beta$ and

$$\begin{aligned}\|x-y\|^2 - \|Qx-y\|^2 &\ge 2(\beta-\alpha)\|x-Qx\| = 2(\beta-\alpha)d_Z(x)\\ &= 2(\beta-\alpha)\big(\beta + d_C(x)\big) \ge 4\beta(\beta-\alpha).\end{aligned} \tag{1.8}$$

(iii). $\beta < d_Z(x) < 2\beta$ and

$$\begin{aligned}\|x-y\|^2 - \|Qx-y\|^2 \ge 2(\beta-\alpha)\|x-Qx\| &= \frac{2(\beta-\alpha)}{\beta} d_Z(x)\big(d_Z(x)-\beta\big)\\ &= \frac{2(\beta-\alpha)}{\beta} d_C(x)\big(\beta + d_C(x)\big).\end{aligned} \tag{1.9}$$

Consequently, in *every* case, we have

$$\|x-y\|^2 - \|Qx-y\|^2 \ge 2(\beta-\alpha)\|x-Qx\| \ge 2(\beta-\alpha)d_C(x). \tag{1.10}$$

Proof. Set $\delta := d_Z(x)$, $p := P_Z x$, and write $y = z + \alpha b$, where $z \in Z$, $b \in X$ and $\|b\| \le 1$. Note that if $\delta > \beta$, then $d_C(x) = \delta - \beta$ using Fact 3.

(i): This follows immediately from the definition of Q.

(ii): Using Cauchy–Schwarz in (1.11a), and the projection theorem (see, e.g., [4, Theorem 3.14]) in (1.11b), we obtain

$$\begin{aligned}\|x-y\|^2 - \|Qx-y\|^2 &= \|x-(z+\alpha b)\|^2 - \|p-(z+\alpha b)\|^2\\ &= \|x\|^2 - \|p\|^2 - 2\langle x, z+\alpha b\rangle + 2\langle p, z+\alpha b\rangle\\ &\ge \|x\|^2 - \|p\|^2 + 2\langle p-x, z\rangle - 2\alpha\|p-x\|\|b\| &\text{(1.11a)}\\ &\ge \|x\|^2 - \|p\|^2 + 2\langle p-x, z-p\rangle + 2\langle p-x, p\rangle - 2\alpha\delta\\ &\ge \|x\|^2 - \|p\|^2 + 2\langle p-x, p\rangle - 2\alpha\delta &\text{(1.11b)}\\ &= \|x\|^2 + \|p\|^2 - 2\langle x, p\rangle - 2\alpha\delta = \|x-p\|^2 - 2\alpha\delta\\ &= \delta^2 - 2\alpha\delta = \delta\big(\delta-2\alpha\big) \ge 2\delta(\beta-\alpha) \ge 4\beta(\beta-\alpha).\end{aligned}$$

(iii): Set $\eta := (\delta-\beta)/\beta \in \,]0,1[$. Then $Qx = (1-\eta)x + \eta p$ and hence

$$\|x-Qx\| = \eta\|x-p\| = \eta\delta = \frac{\delta-\beta}{\beta}\delta. \tag{1.12}$$

Using, e.g., [4, Corollary 2.14] in (1.13a), and Cauchy–Schwarz and [4, Theorem 3.14] in (1.13b), we obtain

$$
\begin{aligned}
\|x-y\|^2-\|Qx-y\|^2 &= \|x\|^2-\|Qx\|^2-2\langle x,y\rangle+2\langle Qx,y\rangle \\
&= \|x\|^2-\|(1-\eta)x+\eta p\|^2+2\langle(1-\eta)x+\eta p-x,y\rangle \\
&= \|x\|^2-(1-\eta)\|x\|^2-\eta\|p\|^2 \qquad (1.13a)\\
&\qquad +\eta(1-\eta)\|x-p\|^2+2\eta\langle p-x,z+\alpha b\rangle \\
&= \eta\big(\|x\|^2-\|p\|^2\big)+(1-\eta)\eta\|x-p\|^2 \\
&\qquad +2\eta\big(\langle p-x,z-p\rangle+\langle p-x,p\rangle+\alpha\langle p-x,b\rangle\big) \\
&\geq \eta\big(\|x\|^2-\|p\|^2+(1-\eta)\|x-p\|^2+2\langle p-x,p\rangle\big) \qquad (1.13b)\\
&\qquad -2\alpha\eta\|x-p\| \\
&= \eta\big(\|x-p\|^2+(1-\eta)\|x-p\|^2-2\alpha\|x-p\|\big) \\
&= \delta\eta\big((2-\eta)\delta-2\alpha\big) \\
&= \frac{\delta}{\beta}(\delta-\beta)\Big(\big(2-(\delta-\beta)\beta^{-1}\big)\delta-2\alpha\Big) \\
&= \frac{\delta}{\beta}(\delta-\beta)\big(-\beta^{-1}\delta^2+3\delta-2\alpha\big).
\end{aligned}
$$

Now the quadratic $q\colon[\beta,2\beta]\to\mathbb{R}\colon\xi\mapsto-\beta^{-1}\xi^2+3\xi-2\alpha$ has a maximizer at $\xi=(3/2)\beta$ and it satisfies $q(\beta)=q(2\beta)=2(\beta-\alpha)\geq 0$. It follows that $\min q\big([\beta,2\beta]\big)=2(\beta-\alpha)$. Therefore, (1.13) and (1.12)

$$\|x-y\|^2-\|Qx-y\|^2\geq\frac{\delta}{\beta}(\delta-\beta)2(\beta-\alpha)=2(\beta-\alpha)\|x-Qx\|. \qquad (1.14)$$

The proof of the “Consequently” part follows easily. □

Proposition 6 implies that the intrepid projector is *quasi nonexpansive*; see, e.g., [3, 8, 22, 23, 24] for further results utilizing this notion.

1.3 Fejér monotonicity

We now review the definition and basic results on Fejér monotone sequences. These will be useful in establishing our convergence results.

Definition 7. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in X, and let C be a nonempty closed convex subset of X. Then $(x_k)_{n\in\mathbb{N}}$ is *Fejér monotone with respect to C* if

$$(\forall k\in\mathbb{N})(\forall c\in C)\quad \|x_{k+1}-c\|\leq\|x_k-c\|. \qquad (1.15)$$

Fact 8. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in X that is Fejér monotone with respect to some nonempty closed convex subset C of X. Then the following hold:

(i). If $\operatorname{int} C \neq \varnothing$, then $(x_k)_{k\in\mathbb{N}}$ converges strongly to some point in X.
(ii). If each weak cluster point of $(x_k)_{k\in\mathbb{N}}$ lies in C, then $(x_k)_{k\in\mathbb{N}}$ converges weakly to some point in C.
(iii). If $d_C(x_k) \to 0$, then $(x_k)_{k\in\mathbb{N}}$ converges strongly to some point in C.

Proof. See, e.g., [2], [4, Chapter 5], or [12]. □

1.4 The method of cyclic intrepid projections

We now assume that each C_i is a closed convex subset of X, with

$$C := \bigcap_{i\in I} C_i \neq \varnothing. \tag{1.16}$$

The index set is split into two sets, corresponding to enlargements and regular sets:

$$I_0 := \big\{ i \in I \mid C_i = (Z_i)_{[\beta_i]},\ \text{where } \beta_i > 0 \big\} \quad \text{and} \quad I_1 := I \smallsetminus I_0. \tag{1.17}$$

Assume an index selector map

$$\mathrm{i}\colon \mathbb{N} \to I, \tag{1.18}$$

where $(\forall i \in I)\ \mathrm{i}^{-1}(i)$ is an infinite subset of $\mathbb{N}$. We say that the *control is quasicyclic with quasiperiod* $M \in \{1,2,\ldots\}$ if $(\forall k \in \mathbb{N})\ I = \{\mathrm{i}(k), \mathrm{i}(k+1), \ldots, \mathrm{i}(k+M-1)\}$.

Let $(\lambda_i)_{i\in I_1}$ be a family in $]0,2[$. We define a family of operators

$$(T_i)_{i\in I} \tag{1.19}$$

from X to X as follows. If $i \in I_0$, then T_i is the intrepid projector onto C_i (with respect to Z_i and β_i); if $i \in I_1$, then T_i is the relaxed projector onto C_i with relaxation parameter λ_i.

Algorithm 9 (method of cyclic intrepid projections). Given a starting point $x_0 \in X$, the *method of intrepid projections* proceeds via

$$(\forall k \in \mathbb{N}) \quad x_{k+1} := T_{\mathrm{i}(k)} x_k. \tag{1.20}$$

We begin our analysis with a simple yet useful observation.

Lemma 10. The sequence $(x_k)_{k\in\mathbb{N}}$ generated by Algorithm 9 is Fejér monotone with respect to C.

Proof. Combine Fact 1 with Proposition 6. □

We now deepen our convergence analysis. We start with the purely intrepid case.

Theorem 11 (intrepid projections only). Suppose that $I_1 = \varnothing$ and that $\operatorname{int} C \neq \varnothing$. Then $(x_k)_{k\in\mathbb{N}}$ converges strongly to some point in C.

Proof. Combining Lemma 10 with Fact 8(i), we deduce that $(x_k)_{k\in\mathbb{N}}$ converges strongly to some point $\bar{x} \in X$. Let $i \in I$. Then there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_k)_{k\in\mathbb{N}}$ such that $(\forall k \in \mathbb{N})\ \mathrm{i}(n_k) = i$. By Proposition 6, the subsequence $(x_{n_k+1})_{k\in\mathbb{N}} = (T_i x_{n_k})_{k\in\mathbb{N}}$ lies in C_i. Since C_i is closed, we deduce that $\bar{x} \in C_i$. □

The proof of the following result follows that of Herman [18] who considered more restrictive controls. (See also [10].)

Corollary 12 (parallelotope). Suppose that X is finite-dimensional, that $I_1 = \varnothing$, that each Z_i is a hyperplane, and that $\operatorname{int} C \neq \varnothing$. Then $(x_k)_{k\in\mathbb{N}}$ converges to some point in the parallelotope C *in finitely many steps*.

Proof. By Theorem 11, $(x_k)_{k\in\mathbb{N}}$ converges to some point $\bar{x} \in C$. If $\{x_k \mid k \in \mathbb{N}\} \cap \operatorname{int} C \neq \varnothing$, then $(x_k)_{k\in\mathbb{N}}$ is eventually constant. Assume to the contrary that $(x_l)_{k\in\mathbb{N}}$ is not eventually constant. Then $\bar{x} \notin \{x_k \mid k \in \mathbb{N}\} \cup \operatorname{int} C$. Since each C_i is a hyperslab, $\operatorname{bdry} C_i$ is the union of two disjoint hyperplanes parallel to Z_i. We collect these finitely many hyperplanes in a set H. The finite collection of these hyperplanes containing $\bar{x}$, which we denote by $H(\bar{x})$, is nonempty. Moreover, $(x_k)_{k\in\mathbb{N}}$ cannot have arisen with infinitely many projection steps as these only occur at a minimum distance from the sets. Therefore, infinitely many reflection steps have been executed. Hence there exists $K_1 \in \mathbb{N}$ such that iteration index k onwards, we only execute identity or reflection steps. Now let $\varepsilon > 0$ be sufficiently small such that $B(\bar{x};\varepsilon)$ makes an empty intersection with every hyperplane drawn from $H \smallsetminus H(\bar{x})$. Since $x_k \to \bar{x}$, there exists $K_2 \in \mathbb{N}$ such that $(\forall k \geq N_2)\ x_k \in B(\bar{x};\varepsilon)$. Since $\bar{x} \in C$ and $(x_k)_{k\in\mathbb{N}}$ is Fejér monotone with respect to C, it follows that $(\forall k \in \mathbb{N})\ \|x_{k+1} - \bar{x}\| \leq \|x_k - \bar{x}\|$. Hence the aforementioned reflection steps from K_2 onwards must be all with respect to hyperplanes taken from $H(\bar{x})$. Set $K := \max\{K_1, K_2\}$. It follows altogether that $(\forall k \geq K)\ 0 < \|x_{k+1} - \bar{x}\| = \|x_k - \bar{x}\|$. But this is absurd since $x_k \to \bar{x}$. □

Remark 13. Both Theorem 11 and Corollary 12 fail if $\operatorname{int} C = \varnothing$: indeed, consider two hyperslabs C_1 and C_2 in $\mathbb{R}^2$ such that C is a line (necessarily parallel to Z_1 and Z_2). If we start the iteration sufficiently close to this line, but not on this line, then $(x_n)_{n\in\mathbb{N}}$ will oscillate between two point outside C.

Furthermore, finite convergence may fail in Corollary 12 without the interiority assumption: indeed, consider a hyperslab in $\mathbb{R}^2$ which is intersected by a line at an angle strictly between 0 and $\pi/4$.

We now present our fundamental convergence result.

Theorem 14 (main result). Suppose that $\bigcap_{i\in I_1} C_i \cap \bigcap_{i\in I_0} \operatorname{int} C_i \neq \varnothing$ and that the control is quasicyclic. Then the sequence $(x_k)_{k\in\mathbb{N}}$ generated by Algorithm 9 converges weakly to some point in C. The convergence is strong provided one of the following conditions holds:

(i). X is finite-dimensional.
(ii). I_1 is either empty or a singleton.

Proof. By Lemma 10, $(x_k)_{k\in\mathbb{N}}$ is Fejér monotone with respect to C. Take $y \in \bigcap_{i\in I_1} C_i \cap \bigcap_{i\in I_0} \operatorname{int} C_i$. Writing $\|x_0 - y\|^2 = \sum_{k\in\mathbb{N}} \|x_k - y\|^2 - \|x_{k+1} - y\|^2$, and recalling Fact 1 and Proposition 6, we deduce that $x_k - x_{k+1} \to 0$ and that $d_{C_{\mathrm{i}(n)}}(x_k) \to 0$. The quasicyclicality of the control now yields

$$\max\left\{d_{C_i}(x_k) \mid i \in I\right\} \to 0. \tag{1.21}$$

Therefore, every weak cluster point of $(x_k)_{k\in\mathbb{N}}$ lies in C. By Fact 8(ii), there exists $\bar{x} \in X$ such that

$$x_k \rightharpoonup \bar{x} \in C \tag{1.22}$$

as announced.

Let us turn to strong convergence. Item (i) is obvious since strong and weak convergence coincide in finite-dimensional Hilbert space.

Now consider (ii). If $I_1 = \varnothing$, then strong convergence follows from Theorem 11. Thus assume that I_1 is a singleton. By, e.g., [2, Theorem 5.14],

$$d_C(x_k) \to 0. \tag{1.23}$$

Hence, using Fact 8(iii), we conclude that $x_k \to \bar{x}$. □

Remark 15. Our sufficient conditions for strong convergence are sharp: indeed, Hundal's example [20] shows that strong convergence may fail if (i) X is infinite-dimensional and (ii) I_1 contains more than one element.

1.5 CycIP and the road design problem

From now on, we assume that

$$X = \mathbb{R}^n, \tag{1.24}$$

and that we are given n breakpoints

$$t = (t_1, \ldots, t_n) \in X \quad \text{such that } t_1 < \cdots < t_n. \tag{1.25}$$

The problem is to

$$\text{find } x = (x_1, \ldots, x_n) \in X \tag{1.26}$$

such that all of the following constraints are satisfied:

- **interpolation constraints:** For a subset J of $\{1, \ldots, n\}$, we have $x_j = y_j$, where $y \in \mathbb{R}^J$ is given.

- **slope constraints:** each slope $s_j := (x_{j+1} - x_j)/(t_{j+1} - t_j)$ satisfies $|s_j| \leq \sigma_j$, where $j \in \{1, \ldots, n-1\}$ and $\sigma \in \mathbb{R}^{n-1}_{++}$ is given.
- **curvature constraints:** $\gamma_j \geq s_{j+1} - s_j \geq \delta_j$, for every $j \in \{1, \ldots, n-2\}$, and for given γ and δ in $\mathbb{R}^{n-2}$.

This problem is of fundamental interest in road design; see [5] for further details.

By grouping the constraints appropriately, this feasibility problem can be reformulated as the following convex feasibility problem involving six sets:

$$\text{find } x \in C = \bigcap_{i \in I} C_i = C_1 \cap \cdots \cap C_6, \tag{1.27}$$

where $I := \{1, \ldots, 6\}$; see [5, Section 2] for details. These sets have additional structure: C_1 is an affine subspace incorporating the interpolation constraints, C_2 and C_3 are both intersections of hyperslabs with normal vectors having disjoint support modeling the slope constraints, and the curvature constraints are similarly incorporated through C_4, C_5, and C_6. All these sets have explicit and easy-to-implement (regular and intrepid) projection formulas. Since only the set C_1 has no interior, we set $T_1 = P_{C_1}$. For every $i \in \{2, \ldots, 6\}$, we set Q_{C_i}. (If we set $T_i = P_{C_i}$, we get the classical method of cyclic projections.) This gives rise to the algorithm, which we call the method of **cyclic intrepid projections (CycIP)**.

We thus obtain the following consequence of our main result (Theorem 14):

Corollary 16 (strict feasibility). Suppose that $C_1 \cap \bigcap_{2 \leq i \leq 6} \operatorname{int} C_i \neq \varnothing$, i.e., there exists a *strictly feasible* solution to (1.27), i.e., it satisfies the interpolation constraints, and it satisfies the slope and curvature constraint inequalities *strictly*. Then the sequence generated by CycIP converges to a solution of (1.27).

In [5], which contains a comprehensive comparison of various algorithms for solving (1.27), CycIP was found to be the best overall algorithm. However, due to the interpolation constraint set C_1, which has *empty interior*, the convergence of CycIP is not guaranteed by Theorem 11 or convergence results derived earlier. Corollary 16 is the first *rigorous justification* of CycIP in the setting of road design.

Remark 17 (nonconvex minimum-slope constraints). In [5] we also considered a variant of the slope constraints with an imposed minimal strictly positive slope. This is a setting of significant interest in road design as zero slopes are not favoured because of, e.g., drainage problems. The accordingly modified sets C_2 and C_3 are in that case *nonconvex*; however, explicit formulas for (regular and intrepid) projections are still available. The application of CycIP must then be regarded as a heuristic as there is no accompanying body of convergence results.

In the following section, we will investigate the numerical performance of CycIP and compare it to a linear programming solver.

1.6 Numerical results

We generate 87 random test problems[3] as in [5]. The size of each problem, n, satisfies $341 \le n \le 2735$. These problems are significantly larger than those of [5] because we wish to compare execution time rather than number of iterations.

Consider the following two measures of infeasibility:

$$(\forall x \in X) \quad d_2(x) := \sqrt{\sum_{i=1}^{6} d_{C_i}^2(x)} \quad \text{and} \quad d_\infty(x) := \max_{i \in I} \|x - P_{C_i}x\|_\infty, \tag{1.28}$$

where $\|x\|_\infty$ is the max-norm[4] of x. Note that $d_2(x) = d_\infty(x) = 0$ if and only if $x \in C$. Set $\varepsilon := 5 \cdot 10^{-4}$, and let $(x_k)_{k\in\mathbb{N}}$ be a sequence generated by CycIP. We employ either $d_2(x_k) < \varepsilon$ or $d_\infty(x_k) < \varepsilon$ as stopping criterion.

Let $\mathscr{P}$ be the set of test problems, and let $\mathscr{A}$ be the set of algorithms. Let $(x_k^{(a,p)})_{k\in\mathbb{N}}$ be the sequence generated by algorithm $a \in \mathscr{A}$ applied to the problem $p \in \mathscr{P}$. To compare the performance of the algorithms, we use *performance profiles*[5]: for every $a \in \mathscr{A}$ and for every $p \in \mathscr{P}$, we set

$$r_{a,p} := \frac{\tau_{a,p}}{\min\left\{\tau_{a',p} \mid a' \in \mathscr{A}\right\}} \ge 1, \tag{1.29}$$

where $\tau_{a,p} \in \{1, 2, \ldots, \tau_{\max}\}$ is the time that a requires to solve p and $\tau_{\max}$ is the maximum time allotted for all algorithms. If $r_{a,p} = 1$, then a uses the least amount of time to solve problem p. If $r_{a,p} > 1$, then a requires $r_{a,p}$ times more time for p than the algorithm that uses the least amount of time for p. For each algorithm $a \in \mathscr{A}$, we plot the function

$$\rho_a \colon \mathbb{R}_+ \to [0,1] \colon \kappa \mapsto \frac{\operatorname{card}\left\{p \in \mathscr{P} \mid \log_2(r_{a,p}) \le \kappa\right\}}{\operatorname{card} \mathscr{P}}, \tag{1.30}$$

where "card" denotes the cardinality of a set. Thus, $\rho_a(\kappa)$ is the percentage of problems that algorithm a solves within factor 2^κ of the best algorithms. Therefore, an algorithm $a \in \mathscr{A}$ is "fast" if $\rho_a(\kappa)$ is large for κ small; and a is "robust" if $\rho_a(\kappa)$ is large for κ large.

To compare CycIP with a linear programming solver, we model (1.27) as the constraints of a *Linear Program (LP)*. As objective function, we use $x \mapsto \|x - x_0\|_1$, where $\|x\|_1$ denotes the 1-norm[6] of x. As LP solver, we use Gurobi 5.5.0, a state-of-the-art mathematical programming solver [17]. CycIP was implemented with the

---

[3] In [5], the authors compared CycIP with a Swiss Army Knife. The Wenger Swiss Army Knife version XXL, listed in the Guinness Book of World Records as the world's most multi-functional penknife, contains 87 tools.

[4] Recall that if $x = (\xi_1, \ldots, \xi_n) \in X$, then $\|x\|_\infty = \max\{|\xi_1|, \ldots, |\xi_n|\}$.

[5] For further information on performance profiles, we refer the reader to [14].

[6] Recall that if $x = (\xi_1, \ldots, \xi_n) \in X$, then $\|x\|_1 = |\xi_1| + \cdots + |\xi_n|$.

C++ programming language. We run the experiments on a Linux computer with a 2.4 GHz Intel® Xeon® E5620 CPU and 24 GB of RAM. As time measurement, we use *wall-clock time*[7]. We limit the solving time to $\tau_{\max} := 150$ seconds for each problem and algorithm.

Figure 1.1 shows the performance profile for the convex case. Here, CycIP uses a cyclic control with period 6 and the randomized **rCycIP** variant has quasicyclic control satisfying $(\forall k \in \mathbb{N})\ \{\mathrm{i}(6k), \mathrm{i}(6k+1), \ldots, \mathrm{i}(6k+5)\} = \{1, 2, \ldots, 6\}$, i.e., for every $k \in \mathbb{N}$, $(\mathrm{i}(6k), \mathrm{i}(6k+1), \ldots, \mathrm{i}(6k+5))$ is a randomly generated permutation of $(1, 2, \ldots, 6)$. Depending on whether d_2 or d_∞ was used as the infeasibility measure, we write CycIP_2 and CycIP_∞, respectively, and similarly for rCycIP.

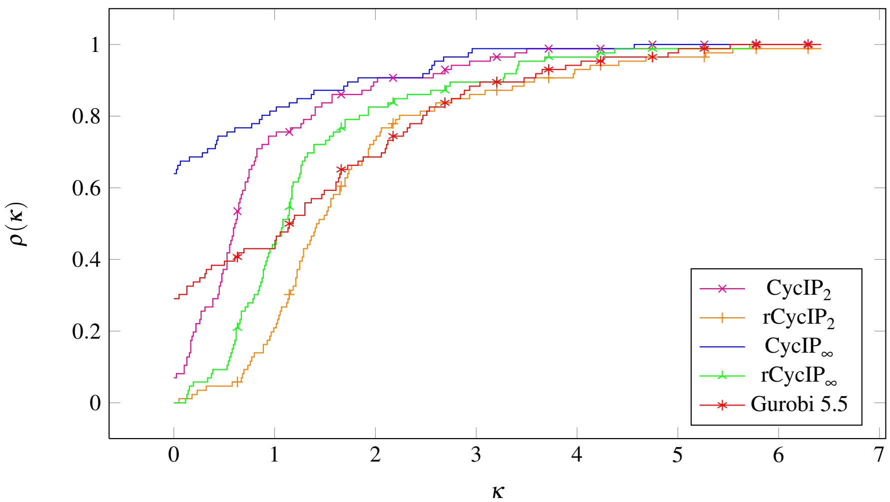

Fig. 1.1 Performance profile for convex problems.

For the nonconvex case, shown in Figure 1.2, we included a minimum slope constraint as mentioned in Remark 17. For the LP solver, the resulting model becomes a *Mixed Integer Linear Program*.

[7] To allow for a more fair comparison, we included in wall-clock time only the time required for running the solver's software itself (and not the time for loading the problem data or for setting up the solver's parameters).

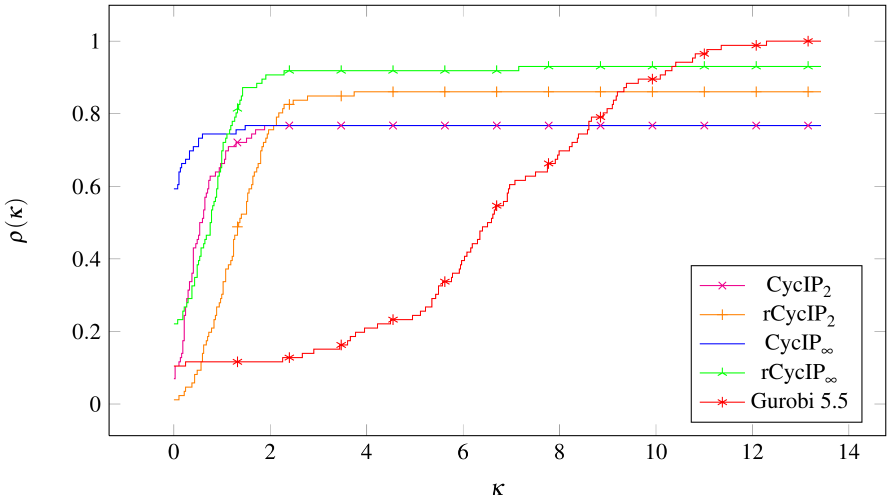

Fig. 1.2 Performance profile for nonconvex problems.

We infer from the figures that for convex problems, CycIP_∞ solves the test problems quickly and robustly. For nonconvex problems, CycIP_∞ is still fast, but less robust than the slower randomized variant rCycP_∞. Gurobi is the slowest — but also the most robust — algorithm.

1.7 Conclusion

In this work, we proved that the method of cyclic intrepid projections converges to a feasible solution under quasicyclic control and an interiority assumption. Specialized to a problem arising in road design, this leads to the first rigorous proof of convergence of CycIP. Numerical results show that CycIP is competitive compared to a commercial optimization solver, especially in terms of speed. Randomization strategies increase robustness in case of nonconvex problems for which there is no underlying convergence theory. Future work will focus on obtaining theoretical convergence results and on experimenting with other algorithms to increase robustness in the nonconvex setting.

Acknowledgements The authors thank the referee for very careful reading and constructive comments, Dr. Ramon Lawrence for the opportunity to run the numerical experiments on his server, and Scott Fazackerley and Wade Klaver for technical help. HHB also thanks Dr. Masato Wakayama and the Institute of Mathematics for Industry, Kyushu University, Fukuoka, Japan for their hospitality — some of this research benefited from the extremely stimulating environment during the "Math-for-Industry 2013" forum. HHB was partially supported by the Natural Sciences and Engineering Research Council of Canada (Discovery Grant and Accelerator Supplement) and by the Canada Research Chair Program.

References

1. Baillon, J.B., Bruck, R.E., Reich, S.: On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces. Houston J. Math. **4**, 1-9 (1978)
2. Bauschke, H.H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems. SIAM Rev. **38**, 367–426 (1996)
3. Bauschke, H.H., Combettes, P.L.: A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces. Math. Oper. Res. **26**, 248–264 (2001)
4. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, New York (2011)
5. Bauschke, H.H., Koch, V.R.: Projection methods: Swiss Army knives for solving feasibility and best approximation problems with halfspaces. Contemporary Mathematics. In press.
6. Borwein, J.M., Vanderwerff, J.D.: Convex Functions. Cambridge University Press, Cambridge (2010)
7. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
8. Cegielski, A.: Iterative Methods for Fixed Point Problems in Hilbert Spaces. Springer, New York (2012)
9. Censor, Y., Zenios, S.A.: Parallel Optimization. Oxford University Press (1997)
10. Chen, W., Herman, G.T.: Efficient controls for finitely convergent sequential al gorithms. ACM Trans. Math. Software **37** 14 (2010)
11. Combettes, P.L.: Hilbertian convex feasibility problems: convergence of projection methods. Appl. Math. Optim. **35** 311-330 (1997)
12. Combettes, P.L.: Quasi-Fejérian analysis of some optimization algorithms. In D. Butnariu, Y. Censor, S. Reich (editors), Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications, Elsevier, New York, 115-152 (2001)
13. Combettes, P.L.: Solving monotone inclusions via compositions of nonexpansive averaged operators. Optimization **53**, 475-504 (2004)
14. Dolan, E.D., Moré, J.J, Benchmarking optimization software with performance profiles. Math. Program. (Ser. A). **91**, 201-213 (2002)
15. Goebel, K., Kirk, W.A.: Topics in Metric Fixed Point Theory. Cambridge University Press, Cambridge (1990)
16. Goebel, K., Reich, S.: Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings. Marcel Dekker, New York (1984)
17. Gurobi Optimization, Inc.: Gurobi Optimizer Reference Manual. http://www.gurobi.com, (2013).
18. Herman, G.T.: A relaxation method for reconstructing objects from noisy x-rays. Math. Programming **8**, 1-19 (1975)
19. Herman, G.T.: Fundamentals of Computerized Tomography: Image Reconstruction from Projections. Second edition. Springer, New York (2009)
20. Hundal, H.: An alternating projection that does not converge in norm. Nonlinear Anal. **57**, 35-61 (2004)
21. Rockafellar, R.T., Wets, R.J-B: Variational Analysis. Springer-Verlag, New York (1998)
22. Slavakis, K., Yamada, I.: The adaptive projected subgradient method constrained by families of quasi-nonexpansive mappings and its applicaton to online learning. SIAM J. Optim. **23**, 126-152 (2013)
23. Yamada, I., Ogura, N.: Hybrid steepest descent method for variational inequality problem over the fixed point set of certain quasi-nonexpansive mappings. Numer. Funct. Anal. Optim. **25**, 619-655 (2004).
24. Yamada, I., Yukawa, M., Yamagishi, M.: Minimizing the Moreau envelope of nonsmooth convex functions over the fixed point set of certain quasi-nonexpansive mappings. In H.H. Bauschke, R.S. Burachik, P.L. Combettes, V. Elser, D.R. Luke, and H. Wolkowicz (editors), Fixed-Point Algorithms for Inverse Problems in Science and Engineering, Springer Optimization and Its Applications **49**, Springer, New York, 345-390 (2011)