

A Differentiable Distance Metric for Robotics Through Generalized Alternating Projection

Vinicius M. Gonçalves, Shiqing Wei, Eduardo Malacarne Soeiro de Souza, Prashanth Krishnamurthy, Anthony Tzes, Farshad Khorrami

Abstract—In many robotics applications, it is necessary to compute not only the distance between the robot and the environment, but also its derivative - for example, when using control barrier functions. However, since the traditional Euclidean distance is not differentiable, meaning it is not guaranteed to be differentiable everywhere, there is a need for alternative distance metrics that possess this property. Recently, a metric with guaranteed differentiability was proposed [1]. This approach has some important drawbacks, which we address in this paper. We provide much simpler and practical expressions for the smooth projection for general convex polytopes. Additionally, as opposed to [1], we ensure that the distance vanishes as the objects overlap. We show the efficacy of the approach in experimental results. Our proposed distance metric is publicly available through the Python-based simulation package UAIBot.

Index Terms—Constrained Motion Planning, Computational Geometry, Motion Control

I. INTRODUCTION

Many obstacle avoidance techniques require the derivative of the distance between the robot and the environment. For example, the widely used Control Barrier Function (CBF) approach [2] incorporates an inequality constraint into an optimization problem to enforce obstacle avoidance. This constraint may involve the first derivative of the distance function. However, it is well known that the traditional Euclidean distance between two general objects is not guaranteed to be even once differentiable [3]¹. Moreover, in some applications, e.g., [1], higher order derivatives of the distance function are required. Thus, even in cases where the distance function is differentiable, its higher-order derivatives can be extremely large, making it unsuitable for these applications. This can lead to excessively high control inputs or inputs that vary too quickly (e.g., high accelerations in case the input is velocity). In either case, it may be infeasible to apply to a real robot [1].

Manuscript received: September 16, 2025; Revised: December 1, 2025; Accepted: January 8, 2026.

This paper was recommended for publication by Editor Cosimo Della Santina upon evaluation of the Associate Editor and Reviewers comments.

V. M. Gonçalves and E. M. S. Souza are with the Department of Electrical Engineering, Federal University of Minas Gerais (UFMG), Brazil. S. Wei, P. Krishnamurthy and F. Khorrami are with New York University, Tandon School of Engineering, U.S.A. A. Tzes is with New York University, Abu Dhabi, United Arab Emirates.

The first author gratefully acknowledges support from the Brazilian funding agencies “Conselho Nacional de Desenvolvimento Científico e Tecnológico” (CNPq), under Grant 301298/2025-4, and “Fundação de Amparo à Pesquisa do Estado de Minas Gerais” (FAPEMIG) under grant APQ-02144-18.

Digital Object Identifier (DOI): see top of this page.

¹In this paper, we use *differentiable* as in *everywhere differentiable*. In general, Euclidean distance is differentiable *almost* everywhere.

Many works have addressed this issue. Some propose specific geometric shapes for which the derivatives of the Euclidean distance are well-defined [3]–[6], while others adopt alternative distance metrics that ensure differentiability [1], [7]–[14]. A recent review of approaches tackling the differentiability problem can be found in [1].

In [1], a smooth version of the half-squared distance between two convex sets is presented². The core idea behind this proposal is the definition of a smooth version of the half-squared *point-to-set* function for a set \mathcal{S} , denoted as $E_{\mathcal{S}}(p)$, which is based on the computation of an integral. From this, by analogy with the true Euclidean half-squared distance, the authors define the *smooth projection* of the point p onto \mathcal{S} as $\Pi_{\mathcal{S}}(p) \triangleq p - \frac{\partial E_{\mathcal{S}}}{\partial p}(p)$. The authors then prove that the classical Von Neumann’s alternating projection algorithm (which projects onto both sets iteratively until convergence) guarantees convergence when we replace the true Euclidean projection with the smooth projection. The resulting converged smooth witness points, a^* and b^* , can then be used in a function to compute the smooth distance between the two sets.

This approach has two important drawbacks. First, computing the projections is difficult, requiring the computation of complex integrals even for simple shapes such as boxes. Second, it lacks an important property: the distance does not vanish when two objects overlap. While the distance decreases as they get closer, it remains a small positive value even when they intersect. Also, this residual value is difficult to determine *a priori*, as it depends on the specific nature of the overlap. As a result, it is not straightforward to simply “subtract out” this positive offset from the smooth distance.

In this paper, we build upon [1] to remove these limitations. For the first problem, we demonstrate that the integral can be eliminated by identifying the essential properties necessary for convergence. This provides much simpler and practical expressions for the smooth projection, not only for the objects considered previously but also for general convex polytopes, which would be very difficult using [1]. For the second problem, we introduce an additional property that ensures the distance vanishes when the objects overlap, which was absent in [1]. Overall, the proposed algorithm is simple to implement, it is significantly faster (as can be seen in Subsection III-A), and it does not require any specialized solvers. Beyond providing an algorithm, our approach also contributes to the theoretical understanding of smooth distance metrics. In

²As explained in [1], we use half-squared distances instead of standard distances because it is more mathematically convenient.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

particular, our work leverages a game-theoretic interpretation of the proposed distance in [1] - which was absent from the previous paper - to obtain some properties. We have integrated our implementation into the UAIBot package³, which is available for installation online. We conducted tests to evaluate the convergence properties of our algorithm and carried out experiments on a Franka Emika robot by incorporating our proposed metric into a CBF framework. All code used for comparisons and experiments, implemented in this package, is publicly available⁴. Finally, for the sake of readability, all proofs are provided in the Appendix at the end of the paper.

II. MATHEMATICAL NOTATION AND PRELIMINARIES

If u is a vector, $\|u\|$ represents the Euclidean distance, and if M is a matrix, $\|M\|$ is the spectral norm (i.e., the square root of the largest eigenvalue of $M^\top M$). $I_{n \times n}$ represents the identity matrix of order n , whereas $0_{n \times m}$ represents the zero matrix of order $n \times m$. If M and N are symmetric matrices, $M \geq N$ (resp. $M > N$) should be interpreted as $M - N$ being positive semidefinite (resp. positive definite). All vectors are column vectors. If $f : \mathbb{R}^n \mapsto \mathbb{R}$, $\frac{\partial f}{\partial p}$ (the gradient) is a column vector. A function $\rho : \mathbb{R}^n \mapsto \mathbb{R}$ that is at least twice differentiable is said to be *strictly convex*⁵ in a set \mathcal{P} if $\frac{\partial^2 \rho}{\partial p^2}(p) > 0$ for all $p \in \mathcal{P}$. A set is said to be *regular* if it is compact and has interior points.

A function $F : \mathbb{R}^n \mapsto \mathbb{R}^n$ is said to be *globally contractive* if there exist $0 < C < 1$ such that for any $u, v \in \mathbb{R}^n$, $\|F(u) - F(v)\| \leq C\|u - v\|$. If F is differentiable, it is said to be *locally contractible* if there exist $0 < c < 1$ such that $\left\| \frac{\partial F}{\partial p}(p) \right\| < c$ for all p . Being locally contractible everywhere implies being globally contractible [15].

III. MAIN RESULTS

A. Generalized Alternating Projection Algorithm

Consider the following definition.

Definition 1. Let $\mathcal{S} \subseteq \mathbb{R}^n$ be a convex regular set. For $k \geq 2$, a function $E_{\mathcal{S}} : \mathbb{R}^n \mapsto \mathbb{R}$ is said to be a k^{th} order generalized point-to-set half squared metric (**k-P2S** henceforth) if the following holds:

- 1) $E_{\mathcal{S}}(p) \geq 0$;
- 2) $I_{n \times n} > \frac{\partial^2 E_{\mathcal{S}}}{\partial p^2}(p) > 0_{n \times n}$ iff $p \notin \mathcal{S}$ and $\frac{\partial^2 E_{\mathcal{S}}}{\partial p^2}(p) = 0$ if $p \in \mathcal{S}$;
- 3) $E_{\mathcal{S}}(p)$ is k -times differentiable on p ;
- 4) $E_{\mathcal{S}}(p)$ vanishes iff $p \in \mathcal{S}$;

Furthermore, given this function $E_{\mathcal{S}}$, we define the k^{th} order generalized projection as the function $\Pi_{\mathcal{S}} : \mathbb{R}^n \mapsto \mathbb{R}^n$ given by $\Pi_{\mathcal{S}}(p) \triangleq p - \frac{\partial E_{\mathcal{S}}}{\partial p}(p)$. \square

Note that, as it could have been expected from a “traditional” projection, there is no guarantee that $\Pi_{\mathcal{S}}(p) \in \mathcal{S}$ for all p . However, as a consequence of Property (4), this is true when $p \in \mathcal{S}$. This is a fundamental result for our approach.

³See https://github.com/viniciusmgn/gap_diff_distance.

⁴See https://github.com/viniciusmgn/gap_diff_distance.

⁵We use in this paper a slightly more restrictive definition of “strictly convex” than the traditional definition, in which $f(x) = x^4$, for example, would be strictly convex.

Proposition 1. A k -P2S function has $\frac{\partial E_{\mathcal{S}}}{\partial p}(p) = 0$ iff $p \in \mathcal{S}$. Consequently, $\Pi_{\mathcal{S}}(p) = p$ iff $p \in \mathcal{S}$.

We recall that all proofs are listed in the Appendix. Our proposed “distance”, for clarity reasons, will be referred to as *metric*. Furthermore, as opposed to [1], our metric is not necessarily smooth (i.e., infinitely differentiable), but only differentiable a finite number of times. Thus, henceforth, we will denote it a *differentiable metric*.

Property (4), and its consequence in Proposition 1, is essential to guarantee that the set-to-set differentiable metric vanishes when the objects overlap, which [1] lacks. Indeed, the half squared metric given by the integral shown in [1]

$$E_{\mathcal{S}}(p) = -h^2 \ln \left(\frac{1}{\text{Vol}(\mathcal{S})} \int_{\mathcal{S}} e^{-\|p-s\|^2/(2h^2)} ds \right) \quad (1)$$

in which $\text{Vol}(\mathcal{S})$ is the n -dimensional volume of \mathcal{S} and $h > 0$ is a smoothing parameter, has the Properties (1), (2) and (3) (with $k = \infty$) but it lacks Property (4).

We proceed by establishing the generalized alternating algorithm when generalized projections, as in Definition 1, are used instead of the classical ones.

Proposition 2. (Generalized alternating algorithm) Let \mathcal{A} and \mathcal{B} be two regular convex sets in which $\mathcal{A} \cap \mathcal{B} = \emptyset$. Consider one generalized projection $\Pi_{\mathcal{A}}$ and $\Pi_{\mathcal{B}}$ for each one of them. Then, for any initial condition $a[0] \in \mathbb{R}^n$, the sequence:

$$a[k+1] = \Pi_{\mathcal{A}}(\Pi_{\mathcal{B}}(a[k])) \quad (2)$$

converges to a point a^* . Furthermore, this point is unique. \square

Note that the statement of the result is structurally the same as Proposition 6 in [1] which, in turn, is structurally the same as the one in Von Neumann’s alternating algorithm [16]. The important difference is that in each one of these cases, the projections Π are different. Furthermore, note that the case in which the objects overlap, $\mathcal{A} \cap \mathcal{B} \neq \emptyset$, is not considered by this proposition. Experimental results show that the algorithm converges even in this case, but the proof seems more complex. This is unnecessary, however. As it will be shown, we can define a^* and b^* in this case to be simply any pair of points such that $a^* = b^* \in \mathcal{A} \cap \mathcal{B}$.

Definition 2. Let $\mathcal{A} \cap \mathcal{B} = \emptyset$. We define as the pair (a^*, b^*) , called differentiable witness points, the pair formed by the (since Proposition 2 guarantees uniqueness) limit point of (2) a^* and its respective counterpart $b^* = \Pi_{\mathcal{B}}(a^*)$. If $\mathcal{A} \cap \mathcal{B} \neq \emptyset$, we simply pick (a^*, b^*) as any points in which $a^* = b^* \in \mathcal{A} \cap \mathcal{B}$. \square

Note that when $\mathcal{A} \cap \mathcal{B} \neq \emptyset$, the differentiable witness points can be found by any classical convex distance calculating algorithm, as the traditional alternating algorithm [16] or GJK [17]. We proceed by, as in [1], using the differentiable witness point to define a differentiable metric between sets. We now define our metric similarly to [1], but with the key difference that we use our differentiable witness points (Definition 2) instead of those defined in [1].

Definition 3. Given two regular convex sets \mathcal{A} and \mathcal{B} , let (a^*, b^*) be differentiable witness points. Then, define the k^{th} order generalized set-to-set half squared metric (**k-S2S** hence-

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

forth) as:

$$\Lambda_{\mathcal{A},\mathcal{B}} \triangleq E_{\mathcal{A}}(b^*) + E_{\mathcal{B}}(a^*) - \frac{\|a^* - b^*\|^2}{2}. \quad (3)$$

□

Note that, from Definitions 1, 2 and 3, when $\mathcal{A} \cap \mathcal{B} \neq \emptyset$, $\Lambda_{\mathcal{A},\mathcal{B}} = 0$ (because $E_{\mathcal{A}}(b^*) = E_{\mathcal{B}}(a^*) = 0$ and $a^* = b^*$). However, although (a^*, b^*) was defined in a “discontinuous” way (from the generalized alternating algorithm when $\mathcal{A} \cap \mathcal{B} = \emptyset$ and from the traditional witness points otherwise), it turns out that $\Lambda_{\mathcal{A},\mathcal{B}}$ is “continuous”, because in the limit case in which the objects overlap, $\Lambda_{\mathcal{A},\mathcal{B}}$ goes to 0. This is one of the key improvements in relation to the result in [1].

Proposition 3. *If $\mathcal{A} \cap \mathcal{B} = \emptyset$, $\Lambda_{\mathcal{A},\mathcal{B}}$ is positive. Also, in the limit case in which the two objects overlap, it goes to 0.* □

The differentiability of $\Lambda_{\mathcal{A},\mathcal{B}}$, when the two objects do not overlap, is guaranteed in the following result.

Proposition 4. *Let $\tau \in \mathbb{R}$. Consider two k -times differentiable rigid transformations $T_A, T_B : \mathbb{R}^n \times \mathbb{R} \mapsto \mathbb{R}^n$ in the variable τ , and extend this transformation to a set $\mathcal{P} \subseteq \mathbb{R}^n$ as $T(\mathcal{P}, \tau) = \{T(p, \tau) \mid p \in \mathcal{P}\}$. Suppose the regular convex sets \mathcal{A} and \mathcal{B} are moving according to $\mathcal{A}(\tau) = T_A(\mathcal{A}_0, \tau)$ and $\mathcal{B}(\tau) = T_B(\mathcal{B}_0, \tau)$ for two fixed regular convex sets \mathcal{A}_0 and \mathcal{B}_0 . Suppose $\Lambda_{\mathcal{A}(\tau), \mathcal{B}(\tau)}$ is built using two k -P2S. Then $\phi(\tau) = \Lambda_{\mathcal{A}(\tau), \mathcal{B}(\tau)}$ is k -times differentiable in the variable τ as long as $\mathcal{A}(\tau) \cap \mathcal{B}(\tau) = \emptyset$. □*

The result is established if $\Lambda_{\mathcal{A},\mathcal{B}}$ depends on a single variable τ . However, if it depends on multiple variables, we can apply Proposition 4 to each of these variables, provided the necessary conditions are met. With this, we are able to state our main result, which is a summary of all the results and definitions in this section:

Theorem 1. *Using the definition of generalized projection in Definition 1, we can obtain the points (a^*, b^*) (as in Definition 2) from the iterative algorithm in (2), which is guaranteed to converge (Proposition 2). Then, plugging it into (3), we obtain a metric $\Lambda_{\mathcal{A},\mathcal{B}}$ that is guaranteed to vanish when $\mathcal{A} \cap \mathcal{B} \neq \emptyset$ (Proposition 3) and that is k -times differentiable under a rigid motion (Proposition 4).*

B. Creating k -P2S functions

The construction of k -P2S functions for regular convex polytopes \mathcal{S} is outlined in this section. Property (2) in Definition 1 implies that $E_{\mathcal{S}}(p)$ should be strictly convex, i.e., $\frac{\partial^2 E_{\mathcal{S}}}{\partial p^2}(p) > 0$. Our construction begins by defining a function that satisfies all the properties in Definition 1, except for this one, which will hold with ≥ 0 instead of > 0 (thus ensuring normal convexity rather than strict convexity). We will then explain how to deform this function to preserve the other properties and achieve strict convexity. We will refer to this preliminary function as **weak k -P2S functions** and use $e_{\mathcal{S}}(p)$ instead of $E_{\mathcal{S}}(p)$ to represent them.

For this purpose, consider the following definition.

Definition 4. *For $k \geq 2$, a function $\Phi : \mathbb{R} \mapsto \mathbb{R}$ is said to be a basic k -P2S function if it has the following properties:*

- 1) $\Phi(s) \geq 0$;
- 2) $1 > \Phi''(s) > 0$ if $s \geq 0$ and $\Phi''(s) = 0$ for $s \leq 0$;
- 3) $\Phi(s)$ is k -times differentiable on s ;
- 4) $\Phi(s) = 0$ iff $s \leq 0$.

□

A base k -P2S function is “essentially” a k -P2S function (as in Definition 1) for the one-dimensional set $(-\infty, 0]$, with the difference that this set is non-compact (as the regularity condition in Definition 1 requires).

This function will serve as a base (hence the name) for constructing more complex (weak) k -P2S functions. Obtaining base k -P2S functions is not difficult. Intuitively, the second derivative of Φ must resemble a differentiable version of the Heaviside indicator function that is $k - 2$ times differentiable at $s = 0$. Hence, one can use, for example:

$$\Phi(s) = \int_0^s \int_0^\xi \left(1 - (r+1)^{-1/h}\right)^{k-1} dr d\xi = \sum_{i=0}^{k-1} \binom{k-1}{i} \frac{(-1)^{k-1-i}}{1-i/h} \left[\frac{(s+1)^{2-i/h} - 1}{2-i/h} - s \right] \quad (4)$$

for $s \geq 0$ and 0 otherwise. Here, h is a positive smoothing parameter and $k \geq 2$ is an integer. Generally, the smaller is the h and the larger is k , the smoother the metric will be. However, the convergence of the algorithm will become slower.

Consider the regular convex polytope \mathcal{S} as the intersection of half-spaces, i.e., $\mathcal{S} = \{p \in \mathbb{R}^n \mid u_i^\top p + v_i \leq 0, i = 1, \dots, n\}$. We can assume, without loss of generality, that $\|u_i\| = 1$. A function that can serve as a weak k -P2S for \mathcal{S} is thus:

$$e_{\mathcal{S}}(p) = \sum_{i=1}^n W_i \Phi(u_i^\top p + v_i) \quad (5)$$

in which W_i are positive constants. Each property in Definition 1 will be studied separately.

Property (1): This easily follows from the fact that $W_i > 0$ and Property (1) in Definition 4.

Property (2): For Property (2), noting that a weak k -P2S should satisfy $I_{n \times n} > \frac{\partial^2 e_{\mathcal{S}}}{\partial p^2} \geq 0_{n \times n}$, we observe that the Hessian of $e_{\mathcal{S}}(p)$ is:

$$\frac{\partial^2 e_{\mathcal{S}}}{\partial p^2}(p) = \sum_{i=1}^n W_i \Phi''(u_i^\top p + v_i) u_i u_i^\top$$

which is a positive semidefinite matrix because $\Phi''(u_i^\top p + v_i) > 0$ (see Property (2) of Definition 4), W_i are nonnegative scalars, and each rank-1 matrix $u_i u_i^\top$ is positive semidefinite. Since this is a positive semidefinite matrix, for it to have all its eigenvalues less than 1, it suffices that its spectral norm is less than 1. A simple bound for the spectral norm using the triangle inequality gives:

$$\left\| \frac{\partial^2 e_{\mathcal{S}}}{\partial p^2}(p) \right\| \leq \sum_{i=1}^n W_i \Phi''(u_i^\top p + v_i)$$

in which we used the fact that $\Phi''(s) \geq 0$ and, since $\|u_i\| = 1$, $\|u_i u_i^\top\| = 1$. Thus, it suffices that the right-hand side is less than 1. Each Φ'' is bounded by 1 (see Property (2) of Definition 4), so we could choose $W_i = \frac{1}{n+\epsilon}$ for a small positive number

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

ϵ , for all i . However, it is beneficial to have the largest W_i possible: small W_i 's make the function Φ small and will hinder the convergence of the iterative projection algorithm in (2). Therefore, one could select $W_i = \frac{1}{m+\epsilon}$, where m is the maximum number of inequalities $u_i^\top p + v_i$ that can be positive simultaneously for all p . The problem of finding the maximum number of linear inequalities that can be positive at the same time can be cast as a mixed linear integer program and solved offline.

Property (3): This follows directly from the fact that Φ is k -times differentiable.

Property (4): For $e_S(p) = 0$ if and only if $p \in \mathcal{S}$, we note that each term $W_i \Phi(u_i^\top p + v_i)$ is nonnegative. Thus, the sum only vanishes when each term vanishes. Since $W_i > 0$, this occurs only when all $\Phi(u_i^\top p + v_i)$ vanish. But $\Phi(s)$ vanishes only when $s \leq 0$, which implies that $u_i^\top p + v_i \leq 0$.

C. Obtaining k -P2S functions from weak k -P2S functions

To obtain a convex function $E_S(p)$ as a k -P2S function from a weak k -P2S function $e_S(p)$, we use this result:

Theorem 2. Let \mathcal{S} be a regular convex set. Let $\rho: \mathbb{R}^n \mapsto \mathbb{R}$ be a k -times differentiable strictly convex function with $\|\frac{\partial^2 \rho}{\partial p^2}\| < 1$ and that is negative when evaluated at any point of \mathcal{S} . Furthermore, that ρ has vanishing gradient only if $p \in \mathcal{S}$. Let e_S be a weak k -P2S function for the set \mathcal{S} . Then, there exists positive scalars ϵ and σ such that:

$$E_S(p) = \epsilon \rho(p) + \sqrt{\sigma^2 e_S(p)^2 + \epsilon^2 \rho(p)^2} \quad (6)$$

is a k -P2S function for the set \mathcal{S} .

One example of function ρ that satisfies the requirements is $\rho(p) = 0.5(\|p - p_c\|^2 - R^2)$, in which p_c and R are such that the ball $\{p \in \mathbb{R}^n \mid \|p - p_c\| \leq R\}$ strictly covers \mathcal{S} , which is possible because \mathcal{S} is regular and thus compact. It is strictly convex, since its Hessian is $\epsilon I_{n \times n}$, and it evaluates to a negative number inside the set \mathcal{S} , since any point in the set is inside the ball, and thus $\|p - p_c\|^2 - R^2 \leq 0$. Furthermore, the gradient of ρ vanishes only at $p = p_c \in \mathcal{S}$.

There does not exist a simple procedure to choose ϵ and σ . The issue is that choosing very small σ and ϵ creates a valid k -P2S function, but with a very small Hessian, and thus the convergence of the generalized alternating algorithm is slow. On the other hand, if they are very large, the condition $I_{n \times n} > \frac{\partial^2 E_S}{\partial p^2}(p)$ (Property (2) in Definition 1) may be violated. These parameters can be obtained through experimentation: select one and test random points p in \mathbb{R}^n to check if $I_{n \times n} > \frac{\partial^2 E_S}{\partial p^2}(p)$ holds.

Figure 1 shows level sets for $e_S(p)$ and the respective function $E_S(p)$ constructed from Theorem 2 with a function $\rho(p) = 0.5(\|p - p_c\|^2 - R^2)$. In this case, \mathcal{S} (the zero level set of both functions) is a pentagon. Note that in $e_S(p)$ the sublevel sets⁶ are not strictly convex shapes (having “flat” sides), because $e_S(p)$ is not strictly convex. The transformation in (6) “bulges” the sublevel sets (except the zero sublevel

set) and transforms them into strictly convex shapes, because $E_S(p)$ is strictly convex.

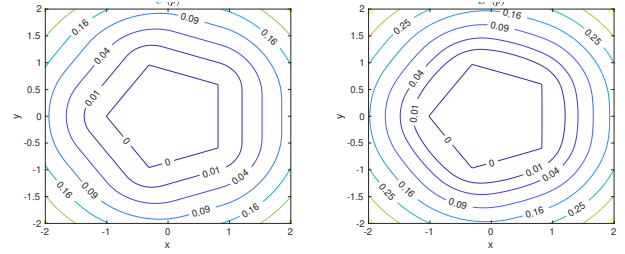


Fig. 1: Contours of the functions $e_S(p)$ (left) and the strictly convexified $E_S(p)$ (right). In this case, \mathcal{S} is a pentagon (the zero sublevel set of both functions).

IV. EXPERIMENTS

A. Experimental computational time

To test the convergence of the generalized alternating algorithm in (2), 50,000 pairs of random regular convex polytopes, each with 10 inequalities of the form $u_i^\top p + v_i \leq 0$, were randomly generated. We considered only those pairs where the true Euclidean distance between them was at least 5 cm. Next, the k -P2S functions were constructed using Φ from (4) with $k = 2$, $h = 0.1$, $e_S(p)$ as defined in (5) with $W_i = 1/6$, and finally $E_S(p)$ as in (6) with $\epsilon = 0.01$, $\sigma = 0.989$, and $\rho(p) = \|p - p_c\| - R^2$, where p_c and R are chosen such that the ball covers the polytope.

The iterative algorithm was initialized with $a[0]$ as the true closest point (in the Euclidean sense) from \mathcal{A} to \mathcal{B} , computed using GJK’s algorithm [17]. Convergence was considered achieved when $\|a[k+1] - a[k]\| < 10^{-3}$. The computational times include the time required to run GJK’s algorithm. The computer used was an Intel i7 with a clock speed of 2.30 GHz. The mean computational time was $70\mu s$, with the maximum being $200\mu s$. The mean number of iterations was 22.42, and the maximum 295.

We have also compared our approach with the Euclidean distance (computed using GJK [17]⁷) and the approach in [1] in terms of computational time. All algorithms were implemented in C++. In this comparison, we used only boxes and cylinders, as it is very complex to handle general polygons in the approach of [1]. In 40,000 trials, we used the same pair of objects for all three approaches, as well as the same starting point and the same smoothing parameter h for our approach and for [1], while changing h from 0.01 to 0.08 with steps of 0.01. In this case, we have the following results for the computational time (showing mean \pm standard deviation): **(ours):** $40 \pm 31 \mu s$, **(Euclidean):** $29 \pm 4 \mu s$ and **([1]):** $146 \pm 407 \mu s$.

B. Experiments with the robot

To showcase the benefits of using a differentiable distance metric—particularly the one proposed in this paper—we conducted an experiment in which a Franka Emika Panda robot

⁶The c sublevel set of a function $f: \mathbb{R}^n \mapsto \mathbb{R}$ is $\{p \in \mathbb{R}^n \mid f(p) \leq c\}$.

⁷Using the specific implementation in <https://gist.github.com/vurtun/29727217c269a2bf4c0ed9a1d11cb40>.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

was tasked with reaching a target pose with its end-effector while avoiding collision with obstacles, self-collision, and respecting joint limits. We implemented a CBF-based controller through a convex quadratic programming formulation, [2] using both our proposed metric and the Euclidean distance, applied to both robot-to-obstacle and self-collision avoidance. In both cases, the robot was controlled via joint velocities, $u = \dot{q}$. One CBF inequality:

$$\nabla_q \Lambda_{ij}^{obs}(q)u \geq -\eta_{obs}(\Lambda_{ij}^{obs}(q) - \lambda^{obs}) \quad (7)$$

was implemented for each pair of objects (i, j) , where i belongs to the robot and j belongs to the environment. The function Λ_{ij}^{obs} denotes the distance (either Euclidean or the square root of twice our half-squared metric), η_{obs} is a positive scalar and λ^{obs} is a scalar safety margin. We have similarly implemented CBFs for auto-collision avoidance (also using the two metrics) and joint-limits avoidance. Figure 2 illustrates the experimental setup through snapshots. The obstacle is composed of three boxes forming a corridor, with the target pose located inside. The obstacle avoidance constraint prevents, among other collisions, the “shoulder” of the robotic arm from hitting the sides of this corridor.

This environment is well-suited to highlight the differentiability issue in the control context: the Euclidean distance becomes non-differentiable in the presence of parallel objects (e.g., two parallel boxes). Since the gripper of the Franka Emika robot is box-shaped and modeled as a box, and the environment is also modeled using boxes, there are many opportunities for such parallel (or nearly parallel) configurations between the gripper and the environment, thus causing issues when we apply the controller with the Euclidean distance. Figure 3 illustrates this: two snapshots of a 3D visualization of the experiment, taken 20 ms apart and using the Euclidean distance, are shown. The figure depicts the gripper model in semitransparent red, together with the two closest points (witness points) between the gripper and the yellow obstacle, shown in green and cyan. It can be observed that, although only 20 ms apart and with the manipulator barely changing, the cyan point quickly switches from one side of the box, which the respective movement of the green point, to the other due to the near parallelism of the red and yellow boxes. This fast-switching behavior, alternating between extremes, persists from $t = 0$ to $t = 3$, just before the gripper enters the hole. Such rapid changes in the witness points are directly related to non-differentiability, since the derivative of the distance function depends on the witness points.

This parallelism causes non-differentiability of the Euclidean distance which, in turn, causes a chattering-like behavior in the control input. This causes the robot to “shake”, especially near the end-effector. This behavior is absent when using our proposed metric. Figure 4 shows the plot for the control input (joint velocity) for the last joint, in our approach and in Euclidean, in the critical phase from $t = 0$ to $t = 5$. One can see that our approach produces much smoother control inputs, while also accomplishing the task and respecting the constraints. A similar behavior appears in all the other joints. Furthermore, this benefit is obtained at a reasonable additional computational cost. At each time step, all distance



Fig. 2: Snapshots of the experimental setup. We need to achieve a target pose in the middle of this corridor.

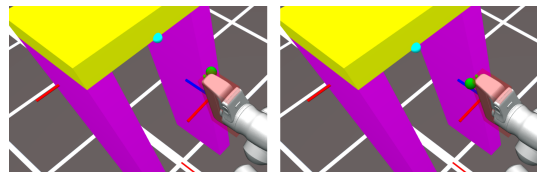


Fig. 3: Two snapshots of a 3D visualization of the experiment using the Euclidean distance, taken 20ms apart from each other. Note how the cyan point jumps from one extreme of the red box to the other.

computations required for determining the control input take, on average, $1000\mu s$ with our metric, compared to $700\mu s$ for the Euclidean distance

Footage of this experiment, along with a 3D visualization, is provided in the accompanying video.

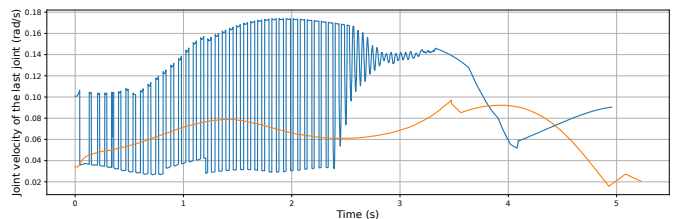


Fig. 4: Joint velocities for the last joint, with Euclidean distance (blue) and our metric (orange).

V. CONCLUSION

In this paper, we build upon the work [1] and propose a new differentiable distance metric between objects. Compared to the approach in [1], our method is significantly simpler to implement—particularly for general convex shapes represented by half-spaces, for example. Furthermore, it is significantly faster and it possesses the important feature of vanishing when the objects overlap, which was absent in [1]. We also provide formal proofs of convergence and other theoretical results, which we believe are of independent theoretical interest.

APPENDIX

Proof of Proposition 1: from Property (1) in Definition 1, $E_S(p) \geq 0$. From Property (4) in Definition 1, the global

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

minimum of this function is achieved iff $p \in \mathcal{S}$. From Property (2) in Definition 1, $E_{\mathcal{S}}(p)$ is a convex function. Since convex functions only have global minima, and the function is differentiable, the gradient vanishes if and only if p lies in this global minima, that is, if $p \in \mathcal{S}$. With this property and the definition of $\Pi_{\mathcal{S}}(p) = p - \frac{\partial E_{\mathcal{S}}}{\partial p}(p)$, it is easy to see that $\Pi_{\mathcal{S}}(p) = p$ iff $p \in \mathcal{S}$. \square

Lemma 1. *Consider the map $F(a) \triangleq \Pi_{\mathcal{A}}(\Pi_{\mathcal{B}}(a))$ coming from two k -P2S functions. Then it is locally contractible if $a \notin \mathcal{A} \cap \mathcal{B}$.*

Proof. The Jacobian of $F(a) = \Pi_{\mathcal{A}}(\Pi_{\mathcal{B}}(a))$ is:

$$\frac{\partial F}{\partial a}(a) = \frac{\partial \Pi_{\mathcal{B}}}{\partial p}(a) \frac{\partial \Pi_{\mathcal{A}}}{\partial p}(\Pi_{\mathcal{B}}(a)). \quad (8)$$

Taking the spectral norm $\|\cdot\|$ of both sides and using the submultiplicativity of the spectral norm, we obtain

$$\left\| \frac{\partial F}{\partial a}(a) \right\| \leq \left\| \frac{\partial \Pi_{\mathcal{B}}}{\partial p}(a) \right\| \left\| \frac{\partial \Pi_{\mathcal{A}}}{\partial p}(\Pi_{\mathcal{B}}(a)) \right\|. \quad (9)$$

The Jacobian of the map $\Pi_{\mathcal{A}}(p)$ is $I - \frac{\partial^2 E_{\mathcal{A}}}{\partial p^2}(p)$, which is a symmetric matrix and hence has a spectral norm equal to its largest eigenvalue in absolute value. According to Property (2) in Definition 1, this matrix has eigenvalues in the open interval $(0, 1)$ if $p \notin \mathcal{A}$ and eigenvalue 1 if $p \in \mathcal{A}$. The same is true, analogously, for $\Pi_{\mathcal{B}}$. So, as long as either $\Pi_{\mathcal{B}}(a) \notin \mathcal{A}$ or $a \notin \mathcal{B}$, which is always true if $\mathcal{A} \cap \mathcal{B} = \emptyset$, from (9) it follows that $\left\| \frac{\partial F}{\partial a}(a) \right\|$ is strictly less than 1 and the map is guaranteed to be locally contractible. \square

Proof of Proposition 2: Lemma 1 shows that if $\mathcal{A} \cap \mathcal{B} = \emptyset$, the map F is always locally contractible. This implies global contractivity ([15]). Thus, Banach's fixed point theorem [18] guarantees that the sequence in (2) converges, and that the limit point is unique. \square

Lemma 2. *Let \mathcal{S} be a regular convex set. Then the function $\Pi_{\mathcal{S}} : \mathbb{R}^n \mapsto \mathbb{R}^n$ is injective, that is, if the equation $\Pi_{\mathcal{S}}(p) = q$ has a solution p for a given q , it has to be unique.*

Proof. $\Pi_{\mathcal{S}}$ is the gradient of the function $\pi_{\mathcal{S}}(p) = \|p\|^2/2 - E_{\mathcal{S}}(p)$ which, according to Property (2) in Definition 1, is a differentiable strictly convex function, and thus it is an injective function (see [19]). \square

Lemma 3. *Let \mathcal{A}, \mathcal{B} be two regular convex sets in which $\mathcal{A} \cap \mathcal{B} = \emptyset$. If $E_{\mathcal{A}}$ and $E_{\mathcal{B}}$ are two k -P2S functions for two regular convex sets, we have:*

$$\Lambda_{\mathcal{A}, \mathcal{B}} = \min_a \max_b \left(E_{\mathcal{A}}(b) + E_{\mathcal{B}}(a) - \frac{\|a - b\|^2}{2} \right) \quad (10)$$

in which the optimizer b for a given a , $b^\circ(a)$, is such that $\Pi_{\mathcal{A}}(b^\circ) = a$, whereas the outer optimizer for a , a° is such that $\Pi_{\mathcal{B}}(a^\circ) = b^\circ$.

Proof. This is a game-theoretic interpretation of $\Lambda_{\mathcal{A}, \mathcal{B}}$, in which we have a two-player competitive game between player A (minimizer) and player B (maximizer). Let $\Lambda(a, b) \triangleq E_{\mathcal{A}}(b) + E_{\mathcal{B}}(a) - \frac{\|a - b\|^2}{2}$ and $\Lambda_{\mathcal{B}}(a) \triangleq \max_b \Lambda(a, b)$. We

start with player B , by maximizing $\Lambda(a, b)$ over b . By differentiating and setting to 0, we find that the optimal strategy $b^\circ(a)$ for player B in accordance to A 's choice a is to have $\Pi_{\mathcal{A}}(b^\circ(a)) = a$ if this equation has a solution (a is in $\Pi_{\mathcal{A}}$'s range), or set $b^\circ(a)$ to an infinite vector following the direction given by the gradient of Λ in b , namely $a - \Pi_{\mathcal{A}}(b)$, forever to earn an infinite objective function Λ (B 's gain is unlimited). Since A wants to minimize Λ , we can assume it is going to choose a such that $\Pi_{\mathcal{A}}(b^\circ(a)) = a$ has a solution to prevent this. Furthermore, this equation will have a unique solution, that can be found by inverting $\Pi_{\mathcal{A}}$ (see Lemma 2). Finally, this strategy is indeed the maximizer because the Hessian of $\Lambda(a, b)$ in b is $\frac{\partial^2 E_{\mathcal{A}}}{\partial p^2}(b^\circ) - I_{n \times n}$, which, according to Property (2) of Definition 1 is negative definite.

The game proceeds with player A 's turn. From the previous player action, $\Lambda_{\mathcal{B}}(a) = \max_b \Lambda(a, b) = \Lambda(a, b^\circ(a))$. Differentiating in a and setting to 0, and using the fact that $\Pi_{\mathcal{A}}(b^\circ(a)) = a$ (which is B 's strategy), we obtain that the optimal strategy a° to A is to have $\Pi_{\mathcal{B}}(a^\circ) = b^\circ(a^\circ)$. This also always has a solution a° . Indeed, applying $\Pi_{\mathcal{A}}$ on both sides (which is an injective mapping, see Lemma 2, and thus this is an "if and only if" step) and using the fact that $\Pi_{\mathcal{A}}(b^\circ) = a^\circ$ (from B 's strategy), we have $\Pi_{\mathcal{A}}(\Pi_{\mathcal{B}}(a^\circ)) = \Pi_{\mathcal{A}}(b^\circ(a^\circ)) = a^\circ$. This is the fixed point equation for which Proposition 2 ensures that the iterative algorithm in (2) will converge to its unique solution a^* since $\mathcal{A} \cap \mathcal{B} = \emptyset$. This shows that the optimizer for both A and B are the differentiable witness points a^* and $b^* = \Pi_{\mathcal{B}}(a^*)$, in accordance to Definition 2.

Finally, we need to show that A 's strategy is indeed a minimizer. This holds because the Hessian of $\Lambda_{\mathcal{B}}(a)$ is

$$\frac{\partial^2 E_{\mathcal{B}}}{\partial p^2}(a^\circ) - I_{n \times n} - \left(\frac{\partial^2 E_{\mathcal{A}}}{\partial p^2}(b^\circ) - I_{n \times n} \right)^{-1} \quad (11)$$

which is positive definite: $\frac{\partial^2 E_{\mathcal{B}}}{\partial p^2}(a^\circ)$ is positive definite (Property (2) of Definition 1) and $-I_{n \times n} - \left(\frac{\partial^2 E_{\mathcal{A}}}{\partial p^2}(b^\circ) - I_{n \times n} \right)^{-1}$ is as well. To see why for the latter, from Property (2) of Definition 1, the eigenvalues of $H = \frac{\partial^2 E_{\mathcal{A}}}{\partial p^2}(b^\circ)$ lie on $(0, 1)$. Let $f(h) = -1 - 1/(h-1) = h/(1-h)$. Thus, the eigenvalues of $W = f(H) = -I_{n \times n} - (H - I_{n \times n})^{-1}$ lie on the range $(0, \infty)$. \square

Lemma 4. *Let \mathcal{A}, \mathcal{B} be two regular convex sets in which $\mathcal{A} \cap \mathcal{B} = \emptyset$. Let (a_0^*, b_0^*) be any pair of points that achieves the smallest (true) Euclidean distance between them. Then:*

$$E_{\mathcal{B}}(a_0^*) \geq \Lambda_{\mathcal{A}, \mathcal{B}} \geq \min_a \Lambda(a, a) \quad (12)$$

in which we borrowed the definition of $\Lambda(a, b)$ from the proof of Lemma 3.

Proof. We apply the game-theoretic/optimization interpretation of $\Lambda_{\mathcal{A}, \mathcal{B}}$ from Lemma 3. For the leftmost inequality, since $\Lambda_{\mathcal{A}, \mathcal{B}} = \min_a \Lambda_{\mathcal{B}}(a)$, then $\Lambda_{\mathcal{A}, \mathcal{B}} \leq \Lambda_{\mathcal{B}}(a)$ for any a . Take $a = a_0^*$. In this case, we have

$$\Lambda_{\mathcal{B}}(a_0^*) = \Lambda(a_0^*, b^\circ(a_0^*)) \quad (13)$$

in which, according to Lemma 3, $b^\circ(a_0^*)$ is such that $\Pi_{\mathcal{A}}(b^\circ(a_0^*)) = a_0^*$. Note that, since $a_0^* \in \mathcal{A}$, $b^\circ(a_0^*) = a_0^*$ is a solution to this equation (See Proposition 1). Furthermore,

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

from Lemma 2, $\Pi_{\mathcal{A}}$ is injective, and thus $\Pi_{\mathcal{A}}(b^\circ(a_0^*)) = a_0^*$ implies $b^\circ(a_0^*) = a_0^*$. Hence:

$$\Lambda_{\mathcal{B}}(a_0^*) = \Lambda(a_0^*, a_0^*) = E_{\mathcal{A}}(a_0^*) + E_{\mathcal{B}}(a_0^*) - \frac{\|a_0^* - a_0^*\|^2}{2} \quad (14)$$

which is equal to $E_{\mathcal{B}}(a_0^*)$ since $E_{\mathcal{A}}(a_0^*) = 0$ (once $a_0^* \in \mathcal{A}$, see Property (4) of Definition 1).

For the rightmost inequality, note that $\Lambda_{\mathcal{B}}(a) = \max_b \Lambda(a, b) \geq \Lambda(a, b)$ for all b . Take $b(a) = a$, and thus $\Lambda_{\mathcal{B}}(a) \geq \Lambda(a, a)$ for any a . Taking the minimum in a in both sides and recalling Lemma 3, the result follows. \square

Proof of Proposition 3:

Proof of positivity: We utilize the definitions of $\Lambda(a, b)$ and $\Lambda_{\mathcal{B}}(a)$ from the proof of Lemma 3. From the rightmost inequality in Lemma 4, $\Lambda_{\mathcal{A}, \mathcal{B}} \geq \min_a \Lambda(a, a)$. But $\Lambda(a, a) = E_{\mathcal{A}}(a) + E_{\mathcal{B}}(a)$, and this is nonnegative (Property (1) of Definition 1). Furthermore, $E_{\mathcal{A}}(a) + E_{\mathcal{B}}(a)$ can be 0 only if $E_{\mathcal{A}}(a) = E_{\mathcal{B}}(a) = 0$ (Property (1) together with Property (4) of Definition 1). This implies that a point a such that $a \in \mathcal{A}$ and $a \in \mathcal{B}$ should exist, that is, $\mathcal{A} \cap \mathcal{B} \neq \emptyset$.

Proof of going to zero when $\mathcal{A} \cap \mathcal{B} \neq \emptyset$: in that case, we start by using the leftmost inequality in Lemma 4. As the two objects tend to overlap, the point a_0^* , that is always inside \mathcal{A} , tends to get closer to \mathcal{B} as well (since, by definition, a_0^* is the closest point in \mathcal{A} to \mathcal{B}). But when $a_0^* \in \mathcal{B}$, we have that $E_{\mathcal{B}}(a_0^*) = 0$ (Proposition 1). Finally, $E_{\mathcal{B}}$ is continuous (since it is k times differentiable, $k \geq 2$ from Property (3) in Definition 1). Hence, from both inequalities in Lemma 4, $\Lambda_{\mathcal{A}, \mathcal{B}}$ is squeezed from the left (due to this argument) and from the right (see the ‘‘Proof of positivity’’ above) to 0. And the result follows. \square

Proof of Proposition 4:

Part I: we start by showing that if τ is such that $\mathcal{A}(\tau) \cap \mathcal{B}(\tau) = \emptyset$ and $a^*(\tau)$ is a limit point of (2), that is, $a^*(\tau) = \Pi_{\mathcal{A}(\tau)}(\Pi_{\mathcal{B}(\tau)}(a^*(\tau)))$, then $a^*(\tau)$ is k -times differentiable. The proof is by induction.

Let $F(p, \tau) = \Pi_{\mathcal{A}(\tau)}(\Pi_{\mathcal{B}(\tau)}(p))$ and thus the fixed point equation is $a^*(\tau) = F(a^*(\tau), \tau)$. Let $M(\tau) = I_{n \times n} - \frac{\partial F}{\partial p}(a^*(\tau), \tau)$. Differentiate $a^*(\tau) = F(a^*(\tau), \tau)$ k -times. We obtain an equation of the form:

$$M(\tau) \frac{d^k a^*}{d\tau^k}(\tau) = G\left(a^*(\tau), \frac{da^*}{d\tau}(\tau), \dots, \frac{d^{k-1} a^*}{d\tau^{k-1}}(\tau), \tau\right) \quad (15)$$

in which the function $G : \mathbb{R}^n \times \mathbb{R}^n \times \dots \times \mathbb{R}^n \times \mathbb{R} \mapsto \mathbb{R}^n$ is continuous in all its arguments, because the functions $E_{\mathcal{A}(\tau)}(p)$, $B_{\mathcal{A}(\tau)}(p)$ are k times differentiable on p (Property (3) in Definition 1) and also in τ because the transformations $T_{\mathcal{A}}(p, \tau)$, $T_{\mathcal{B}}(p, \tau)$ were assumed to be k times differentiable.

We proceed by noting that it is established, in the proof of Lemma 1, that $\|\frac{\partial F}{\partial p}(a^*(\tau), \tau)\| < 1$ as long as $\mathcal{A} \cap \mathcal{B} = \emptyset$. Hence $M(\tau)$ is invertible and consequently:

$$\frac{d^k a^*}{d\tau^k}(\tau) = M(\tau)^{-1} G\left(a^*(\tau), \frac{da^*}{d\tau}(\tau), \dots, \frac{d^{k-1} a^*}{d\tau^{k-1}}(\tau), \tau\right). \quad (16)$$

which shows that $\frac{d^k a^*}{d\tau^k}(\tau)$ is continuous as long as $\frac{d^m a^*}{d\tau^m}(\tau)$ is continuous for $m = 0, \dots, k-1$. The base case of the induction, the continuity of $a^*(\tau)$ on τ , comes from the fact that the map

$F(p, \tau)$ is contractive as long as $\mathcal{A}(\tau) \cap \mathcal{B}(\tau) = \emptyset$ (see [20]).

Part II: we need to prove that $b^*(\tau) = \Pi_{\mathcal{B}(\tau)}(a^*(\tau))$ is also k times differentiable. The proof is analogous to the one in Part I, by differentiating this equation k times in τ .

Part III: for the final part, we note that the k^{th} derivative $\Lambda_{\mathcal{A}(\tau), \mathcal{B}(\tau)}$ involves the k^{th} derivatives of the functions $E_{\mathcal{A}(\tau)}(p)$ and $E_{\mathcal{B}(\tau)}(p)$ on p and τ . These are continuous (from Property (3) in Definition 1 for p and also in τ because the transformations $T_{\mathcal{A}}(p, \tau)$, $T_{\mathcal{B}}(p, \tau)$ were assumed to be k times differentiable). Finally, it also depends on the k^{th} derivatives of $a^*(\tau)$ and $b^*(\tau)$ on τ , which were established to be continuous in Part I and Part II, respectively. \square

Lemma 5. *Let $A, B : \mathbb{R}^n \mapsto \mathbb{R}$ be two functions that are at least twice differentiable, in which A is strictly convex everywhere and B is convex everywhere. Furthermore, suppose that $B \geq 0$ and that A and B do not vanish simultaneously. Then:*

$$C(p) = A(p) + \sqrt{A(p)^2 + B(p)^2} \quad (17)$$

is strictly convex in the set $\mathcal{P} = \{p \in \mathbb{R}^n \mid B(p) > 0\}$.

Proof. It can be checked by applying the derivation rules that:

$$\frac{\partial^2 C}{\partial p^2} = \frac{C \frac{\partial^2 A}{\partial p^2} + B \frac{\partial^2 B}{\partial p^2}}{\sqrt{A^2 + B^2}} + \frac{\left(A \frac{\partial B}{\partial p} - B \frac{\partial A}{\partial p}\right) \left(A \frac{\partial B}{\partial p} - B \frac{\partial A}{\partial p}\right)^\top}{(\sqrt{A^2 + B^2})^3}. \quad (18)$$

Note that the division by $\sqrt{A^2 + B^2}$ never causes problems because A and B cannot simultaneously vanish. Furthermore, $C > 0$ when $p \in \mathcal{P}$. Thus, when $p \in \mathcal{P}$, the first term in (18) is a positive definite matrix since $B \geq 0$, $\frac{\partial^2 A}{\partial p^2} \geq 0$, $C > 0$ and $\frac{\partial^2 B}{\partial p^2} > 0$. In addition, the second term is also positive semidefinite (a matrix of the form αuu^\top for any vector $u \in \mathbb{R}^n$ and a nonnegative α is always positive semidefinite). Since the sum of a positive definite and a positive semidefinite matrix is always a positive definite matrix, the desired result follows. \square

Lemma 6. *If $U : \mathbb{R}^n \mapsto \mathbb{R}$ is convex, twice differentiable with $\|\frac{\partial^2 U}{\partial p^2}\| \leq \mu$, and $U_{\min} = \min_q U(q)$ is finite, then:*

$$\left\| \frac{\partial U}{\partial p}(p) \right\|^2 \leq 2\mu(U(p) - U_{\min}) \quad \forall p. \quad (19)$$

Proof. This is a simple application of Descent Lemma [21]: if $U : \mathbb{R}^n \mapsto \mathbb{R}$ is convex, twice differentiable with $\|\frac{\partial^2 U}{\partial p^2}\| \leq \mu$, then, for any q, p , $U(q) \leq U(p) + (q-p)^\top \nabla U(p) + \frac{\mu}{2} \|p-q\|^2$. After minimizing both sides in q and reorganizing, the result follows. \square

Lemma 7. *Let A and B be convex twice differentiable functions such that $\|\frac{\partial^2 A}{\partial p^2}\| \leq \alpha$ and $\|\frac{\partial^2 B}{\partial p^2}\| \leq \beta$ for constant $\alpha, \beta > 0$. Furthermore, let $A_{\min} = \min_p A(p)$, $B_{\min} = \min_p B(p)$ and $V_{\min} = \min_p \sqrt{A(p)^2 + B(p)^2}$. Then, the Hessian in (18) satisfies:*

$$\left\| \frac{\partial^2 C}{\partial p^2} \right\| \leq \alpha + \left(3 + \frac{\sqrt{A_{\min}^2 + B_{\min}^2}}{V_{\min}} \right) \sqrt{\alpha^2 + \beta^2}. \quad (20)$$

Proof. Let $V = [A \ B]^\top$, $\nu = [\alpha \ \beta]^\top$, $M = [-\frac{\partial B}{\partial p} \ \frac{\partial A}{\partial p}]$ and $V_{\min} = [A_{\min} \ B_{\min}]^\top$. We begin by noting that the rightmost element at the right-hand side of (18) can be written as

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

$V^T M M V$. Applying the norm in both sides of (18), imposing the bounds $\|\frac{\partial^2 A}{\partial p}\| \leq \alpha$ and $\|\frac{\partial^2 B}{\partial p}\| \leq \beta$ and applying Cauchy-Schwarz's inequality/ submultiplicativity of the spectral norm, we can bound $\|V^T M M V\| \leq \|M\|^2 \|V\|^2$ and thus:

$$\left\| \frac{\partial^2 C}{\partial p^2} \right\| \leq \frac{C}{\|V\|} \alpha + \frac{B}{\|V\|} \beta + \frac{\|M\|^2}{\|V\|}. \quad (21)$$

Since $C = A + \|V\|$, the previous equality can be written as:

$$\left\| \frac{\partial^2 C}{\partial p^2} \right\| \leq \alpha + \nu^T \frac{V}{\|V\|} + \frac{\|M\|^2}{\|V\|}. \quad (22)$$

Again applying Cauchy-Schwarz to $\nu^T \frac{V}{\|V\|}$, we find that it is smaller or equal than than $\|\nu\| = \sqrt{\alpha^2 + \beta^2}$, and thus:

$$\left\| \frac{\partial^2 C}{\partial p^2} \right\| \leq \alpha + \sqrt{\alpha^2 + \beta^2} + \frac{\|M\|^2}{\|V\|}. \quad (23)$$

We will bound $\frac{\|M\|^2}{\|V\|}$. We note that $\|M\| \leq \|M\|_F$, in which $\|M\|_F$ is the Frobenius norm [22], thus

$$\|M\|^2 \leq \left\| \frac{\partial A}{\partial p} \right\|^2 + \left\| \frac{\partial B}{\partial p} \right\|^2. \quad (24)$$

From Lemma 6:

$$\begin{aligned} \frac{\|M\|^2}{\|V\|} &\leq 2\alpha \frac{A}{\|V\|} + 2\beta \frac{B}{\|V\|} - 2\alpha \frac{A_{\min}}{\|V\|} - 2\beta \frac{B_{\min}}{\|V\|} = \\ &2 \frac{V^T \nu}{\|V\|} - 2 \frac{V_{\min}^T \nu}{\|V\|} \end{aligned} \quad (25)$$

Applying Cauchy-Schwarz in $V^T \nu$ and $V_{\min}^T \nu$:

$$\frac{\|M\|^2}{\|V\|} \leq \left(2 + \frac{\|V_{\min}\|}{\|V\|} \right) \|\nu\| \leq \left(2 + \frac{\|V_{\min}\|}{V_{\min}} \right) \|\nu\| \quad (26)$$

which concludes the proof. \square

Proof of Theorem 2:

We will prove each one of the properties in Definition 1.

Property (1): comes easily from the expression in (6) and the fact that $\alpha > 0$ and $e_S(p) \geq 0$.

Property (2): from Lemma 5 using $A = \varepsilon\rho$ and $B = \sigma e_S$, it is established that $\frac{\partial^2 E_S}{\partial p^2}(p) > 0$.

To show that σ and ε can be chosen so $I_{n \times n} > \frac{\partial^2 E_S}{\partial p^2}(p)$, we use Lemma 7 again using $A = \varepsilon\rho$ and $B = \sigma e_S$. In this case, $\alpha = \varepsilon$ and $\beta = \sigma$ can be taken. It is clear that ε and σ can be taken sufficiently small to make the right-hand side of (20) as small as necessary, eventually below 1. Note that in this case, A_{\min} is a finite negative number, $B_{\min} = 0$ and V_{\min} is a nonzero finite number (since $\varepsilon\rho$ and σe_S cannot vanish simultaneously once ρ needs to be zero when e_S vanishes).

Property (3): can be checked by inspection, provided that e_S and ρ are k times differentiable, and e_S and ρ cannot vanish simultaneously.

Property (4): from the expression in (6), it is clear that it is zero only if $e_S(p) = 0$. Furthermore, since ρ is negative when $e_S(p) = 0$, it indeed vanishes when $p \in \mathcal{S}$. \square

REFERENCES

[1] V. M. Gonçalves, A. Tzes, F. Khorrami, and P. Fraisse, "Smooth distances for second order kinematic robot control," *IEEE Transactions on Robotics*, vol. 40, pp. 2950 – 2966, 2024.

[2] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: Theory and applications," in *Proceedings of the 18th European Control Conference (ECC)*, (Naples, Italy), IEEE, June 2019. Conference Date: 25–28 June 2019.

[3] A. Escande, S. Miossec, M. Benallegue, and A. Kheddar, "A strictly convex hull for computing proximity distances with continuous gradients," *IEEE Transactions on Robotics*, vol. 30, no. 3, pp. 666–678, 2014.

[4] A. Escande, S. Miossec, and A. Kheddar, "Continuous gradient proximity distance for humanoids free-collision optimized-postures," in *2007 7th IEEE-RAS International Conference on Humanoid Robots*, pp. 188–195, 2007.

[5] A. Dietrich, T. Wimböck, A. Albu-Schäffer, and G. Hirzinger, "Integration of reactive, torque-based self-collision avoidance into a task hierarchy," *IEEE Transactions on Robotics*, vol. 22, pp. 1278–1293, 11 2012.

[6] K. Tracy, T. A. Howell, and Z. Manchester, "Diffpills: Differentiable collision detection for capsules and padded polygons." <https://arxiv.org/abs/2207.00202>, 2022.

[7] X. Zhu, H. Ding, and S. Tso, "A pseudodistance function and its applications," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 2, pp. 344–352, 2004.

[8] J. Xu, Z. Liu, C. Yang, L. Li, and Y. Pei, "A pseudo-distance algorithm for collision detection of manipulators using convex-plane-polygons-based representation," *Robotics and Computer-Integrated Manufacturing*, vol. 66, p. 101993, 12 2020.

[9] A. Schmeißer, R. Wegener, D. Hietel, and H. Hagen, "Smooth convolution-based distance functions," *Graphical Models*, vol. 82, pp. 67–76, 2015.

[10] M. Sanchez, O. Fryazinov, P. Fayolle, and A. Pasko, "Convolution filtering of continuous signed distance fields for polygonal meshes," *Computer Graphics Forum*, vol. 34, p. 277–288, 09 2015.

[11] K. Tracy, T. A. Howell, and Z. Manchester, "Differentiable collision detection for a set of convex primitives," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3663–3670, 2023.

[12] H. J. T. Suh, T. Pang, and R. Tedrake, "Bundled gradients through contact via randomized smoothing," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4000–4007, 2022.

[13] V. M. Gonçalves, P. Krishnamurthy, A. Tzes, and F. Khorrami, "Bssm: Gpu-accelerated point-cloud distance metric for motion planning," *IEEE Robotics and Automation Letters*, vol. 9, no. 11, pp. 10319–10326, 2024.

[14] S. Wei, B. Dai, R. Khorrambakht, P. Krishnamurthy, and F. Khorrami, "Diffocclusion: Differentiable optimization based control barrier functions for occlusion-free visual servoing," *IEEE Robotics and Automation Letters*, vol. 9, no. 4, pp. 3235–3242, 2024.

[15] T. V. Petersdorff, "Notes on contractible maps." <https://terpconnect.umd.edu/~petersd/666/fixpoint.pdf>.

[16] H. Bauschke and J. J. Borwein, "On the convergence of von Neumann's alternating projection algorithm for two sets," *Set-Valued Analysis*, vol. 1, pp. 185–212, 06 1993.

[17] E. G. Gilbert, D. W. Johnson, and S. S. Keerthi, "A fast procedure for computing the distance between complex objects in three-dimensional space," *IEEE Journal on Robotics and Automation*, vol. 4, no. 2, pp. 193–203, 1988.

[18] K. Ciesielski, "On Stefan Banach and some of his results," *Banach Journal of Mathematical Analysis*, vol. 1, no. 1, pp. 1 – 10, 2007.

[19] A. Griewank, H. T. Jongen, and M. K. Kwong, "The equivalence of strict convexity and injectivity of the gradient in bounded level sets," *Mathematical Programming*, vol. 51, pp. 273–278, jul 1991.

[20] M. Beldziński, M. Galewski, and I. Kossowski, "Dependence on parameters for nonlinear equations—abstract principles and applications," *Mathematical Methods in the Applied Sciences*, vol. 45, no. 3, pp. 1668–1686, 2022.

[21] D. P. Bertsekas, *Nonlinear Programming*. Belmont, Massachusetts: Athena Scientific, 2nd ed., 1999.

[22] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.