Constrained Optimal Control with Harmonic Control Lyapunov Barrier Functions applied to Reach-Avoid Problems

Amartya Mukherjee[1], Ruikun Zhou[1], and Jun Liu[1]

***Abstract*— This paper introduces harmonic control Lyapunov barrier functions (harmonic CBLF) that aid in constrained control problems such as reach-avoid problems. Harmonic CBLFs exploit the maximum principle that harmonic functions satisfy to encode the properties of control Lyapunov barrier functions (CBLFs). As a result, they can be initiated at the start of an experiment rather than trained based on sample trajectories. The control inputs are selected to maximize the inner product of the system dynamics with the steepest descent direction of the harmonic CBLF. Numerical results are presented with three different systems under two different environments. Harmonic CBLFs show a significantly low risk of entering unsafe regions and a high probability of entering the goal region.**

I. Introduction

The objective of constrained optimal control is to design a control system that optimizes a notion of performance while satisfying a set of constraints. These constraints are usually designed to prevent catastrophic events, ensure safety, and take into multiple objectives with bounds. Furthermore, with experiments conducted in real life, resetting the environment if the agent goes outside a domain can be expensive [11].

Reach-avoid problems are a subset of constrained optimal control problems where a trajectory aims to reach a goal state while avoiding a set of unsafe states. In recent years, there has been a growing interest in reach-avoid problems among the control theory and reinforcement learning community. Such methods involve learning control Lyapunov functions to certify reachability and control barrier functions that satisfy avoidability [9], [4].

In reinforcement learning, agents learn optimal policies through exploration of the environment with unknown dynamics [18]. Constraints are imposed using penalty methods [10] and lagrangian methods [1]. In these methods, two value functions are computed to estimate cumulative rewards and cumulative costs.

The problem with using two certificates or two value functions is that it leads to conflicts in the control policy whether to satisfy reachability or avoidability at a given point in time. To unify the conflicts, [6] combined reachability and safety into a single certificate called the control Lyapunov barrier function (CBLF). They also propose the Lyapunov barrier actor-critic to find a controller that satisfies both certificates.

Harmonic functions are solutions to the Laplace equation, which is a second-order linear elliptic partial differential equation (PDE). They satisfy the maximum principle in a compact set, thus making it easier to impose CLBF constraints. They do not have any critical points other than saddle points in the interior of their domain, which makes it easy to derive optimal control strategies.

Harmonic functions have been used in the control literature to derive potential fields in past works. [12] used harmonic potential functions with panel methods to build potential fields over a task space. [17] used harmonic functions with complex variables and conformal mappings to achieve moving obstacle avoidance. [5] combines harmonic functions using superposition to derive potential fields.

Safety constraints are grouped into three categories: (a) level I: constraint satisfaction encouraged, (b) level II: constraint satisfaction with high probability, and (c) level III: constraint satisfaction guaranteed [3]. In this paper, we will focus on safety level III problems, where an agent must reach a goal region while avoiding unsafe regions. For example, the unsafe regions could be regions on a floor that have holes or pillars. We achieve this by introducing harmonic control Lyapunov barrier functions (harmonic-CBLFs) that use Laplace's equation to encode the properties of control Lyapunov barrier functions (CBLFs).

Furthermore, Laplace's equation can be solved using numerical methods such as finite element methods (FEMs) and do not be trained using neural networks based on sample trajectories as done in prior work, thus significantly reducing the computational cost of performing experiments. To the best of our knowledge, our work is the first to unify harmonic functions with CBLFs and derive optimal controllers using gradient-based methods. Our experiments show a significantly low risk of trajectories entering unsafe regions.

II. Preliminaries

In this section, we will introduce known definitions and results of Control Lyapunov Barrier Functions (CBLF) and harmonic functions.

Let S be a compact subspace of $\mathbb{R}^n$ denoting the space of all admissible states in a control problem. Let $S_{goal} \subset S$ be an open set denoting the space of states that the controller aims to reach, and let S_{unsafe} be a compact subset of S denoting the space of states the controller aims to avoid. Define $S_{safe} = S \backslash \{\overline{S_{goal}} \cup S_{unsafe}\}$.

A. Control Lyapunov Barrier Functions

Definition 1: (Control Lyapunov Barrier Function) [6] A function $V : S \to \mathbb{R}$ is a Control Lyapunov Barrier Function (CLBF) if, for some constant $c, \lambda > 0$:

[1]Amartya Mukherjee, Ruikun Zhou, and Jun Liu are with the Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1 a29mukhe@uwaterloo.ca

1) $V(s) = 0$ for all $s \in S_{goal}$
2) $V(s) > 0$ for all $s \in S \backslash S_{goal}$
3) $V(s) \geq c$ for all $s \in S_{unsafe}$
4) $V(s) < c$ for all $s \in S \backslash S_{unsafe}$
5) There exists a controller π such that $E_{s'}[V(s') - V(s) + \lambda V(s)] \leq 0$ for all $s' \sim P(s'|s, \pi(a|s))$

In this paper, we will let $c = 1$.

B. Harmonic Functions

Definition 2: (Harmonic Function) [8] A harmonic function u is a C^2-function that satisfies the Laplace equation $\nabla^2 u = \nabla \cdot \nabla u = 0$

In the following theorems, we will introduce properties of harmonic functions that are considered important work in the field of elliptic PDEs, differential geometry, and complex analysis.

Theorem 1: (Mean Value Theorem) [16] If u is harmonic on a domain S, then u satisfies the mean value property in S. Furthermore, if $x \in S$ and $r > 0$ are such that $\overline{B_r(x)} \subset S$, then

$$u(x) = \frac{\int_{B_r(x)} u(y)dy}{\int_{B_r(x)} dy} \tag{1}$$

This theorem is essential for proving theorems 2 and 3 that provide bounds on harmonic functions.

Theorem 2: (Strong Maximum Principle) [8] If u is harmonic on a domain S and u has its maximum in $S \backslash \partial S$, then u is constant.

Theorem 3: (Weak Maximum Principle) [8] If $u \in C^2(S) \cup C^1(\partial S)$ is harmonic in a bounded domain S, then

$$\max_{\overline{S}} u = \max_{\partial S} u \tag{2}$$

This shows that theorems 2 and 3 can be exploited to derive a CBLF that achieves its maximum in unsafe regions and minimum in goal regions.

III. Proposed method

In this work, we will explore the intersection between CBLF and harmonic functions. By exploiting the maximum principle, we can derive a function $V : S \to \mathbb{R}$ that satisfies properties 1–4 for CLBF by imposing boundary conditions on the CBLF.

Definition 3: (Harmonic CBLF) A function $V \in C^2(S) \cup C^1(\partial S)$ is a harmonic CBLF if it satisfies the following conditions.

1) $\nabla^2 V(s) = 0$ for all $s \in S_{safe}$
2) $V(s) = 0$ for all $s \in \overline{S_{goal}}$
3) $V(s) = c$ for all $s \in \partial S \cup S_{unsafe}$

In other words, $V(s)$ is a solution to the boundary value problem given above.

Let us introduce a condition that must be satisfied for the CBLF. The following assumption is important as it implies that for any $s \in S_{safe}$, there exists a continuous path that connects it to a point in ∂S_{goal} without touching a point in S_{unsafe}

Assumption 1: For all points $x_0, x_1 \in S_{safe}$, there exists a continuous path $x(t) \in C[0, T]$ in S_{safe} such that $x(0) = x_0$ and $x(T) = x_1$.

The following theorem provides a sufficient condition to verify the 5th property of CBLF.

Theorem 4: Let V be a harmonic CBLF, and let the dynamics of the deterministic system be $\dot{x} = f(x, u)$. Under assumptions 1, the constant λ from property 5 of CBLF can be computed as:

$$\lambda = \sup_{x \in S_{safe}} \inf_{u \in U} \frac{\langle f(x, u), -\nabla V(x) \rangle}{V(x)} \Delta t, \tag{3}$$

where U is the set of control inputs and Δt is the time step size used for numerical simulation.

Proof: The proof of this is straightforward. The expression $V(x(t + \Delta t)) - V(x(t)) + \lambda V(x(t)) \leq 0$ for small Δt can be written as

$$\frac{\partial}{\partial t} V(x(t)) \Delta t + \lambda V(x(t)) \leq 0 \tag{4}$$

Using the following identity by chain rule:

$$\frac{\partial}{\partial t} V(x(t)) = \langle f(x, u), \nabla V(x) \rangle, \tag{5}$$

we can derive equation 3 above. ■

Furthermore, a sufficient condition for a system to avoid S_{unsafe} is:

$$\sup_{x \in S_{safe}} \inf_{u \in U} \langle f(x, u), \nabla V(x) \rangle \leq 0 \tag{6}$$

Since S_{unsafe} is a closed set, this means $V(x) < c$ for all $x \in S_{safe}$. And since $\frac{\partial}{\partial t} V(x(t)) \leq 0$, this means for all $t \in [0, \infty)$, $x(t)$ will never approach a point where $V(x) = c$, meaning, it will never approach a point in S_{unsafe}.

[19] shows that all critical points in $V(x)$ are saddle points. [7] added random noise to parameters while performing gradient descent and provided guarantees that the algorithm would converge to a local minimum. [13] shows that gradient descent with random parameter initialization asymptotically avoids saddle points. So in this paper, we compare deterministic control inputs with stochastic control inputs to see which method shows better convergence to S_{goal}.

At every point $x \in S_{safe}$, the control input is selected as $u = \inf_{u \in U} \langle f(x, u), \nabla V(x) \rangle + z$ where z is a small noise sampled from a normal distribution $N(0, \sigma^2)$ to mitigate local optima.

The following theorem introduces bounds that must be imposed on z so that the trajectory avoids unsafe regions.

Theorem 5: The trajectory governed by the ODE $\dot{x} = f(x, u)$ will avoid S_{unsafe} under a noisy controller $u + z$ if the noise z satisfies

$$\langle \nabla_u f(x, u) \nabla V(x), z \rangle \leq \frac{c - V(x)}{\Delta t} - \langle \nabla V(x), f(x, u) \rangle, \tag{7}$$

where Δt is the step size used in numerical simulation, and $\nabla_u f(x, u)$ is the Jacobian of f with respect to the control inputs.

Proof: Consider the first-order expansion of $x(t + \Delta t)$

$$x(t + \Delta t) = x + f(x, u)\Delta t + O(\Delta t^2) \tag{8}$$

Add some noise z to the control input

$$x(t+\Delta t) = x + f(x, u+z)\Delta t + O(\Delta t^2) \tag{9}$$

Applying the first-order expansion of $f(x, u+z)$

$$x(t+\Delta t) = x + f(x,u)\Delta t + f_u(x,u)^T z \Delta t + O(\Delta t^2) \tag{10}$$

Consider the first-order expansion of $V(x(t+\Delta t))$

$$\begin{aligned} V(x(t+\Delta t)) =& V(x) + \langle \nabla V, f(x,u) + \nabla_u f(x,u)^T z \rangle \Delta t \\ &+ O(\Delta t^2) \end{aligned} \tag{11}$$

We need to satisfy $V(x(t+\Delta t)) \leq c$

$$\langle \nabla V(x), f(x,u) + \nabla_u f(x,u)^T z \rangle \leq \frac{c - V(x)}{\Delta t} \tag{12}$$

Expanding and simplifying this expression gives the bounds in equation 13.

Furthermore, using the Cauchy-Schwarz inequality, a sufficient way to bound z is

$$||\nabla_u f(x,u) \nabla V(x)||_\infty ||z||_\infty \leq \frac{c - V(x)}{\Delta t} - \langle \nabla V(x), f(x,u) \rangle \tag{13}$$

■

Numerical solutions in section IV show that in harmonic CBLFs, the distinction between safe and unsafe regions is unclear. To make the distinction clearer, we replace the CBLF as a solution to Laplace's equation with a solution to Poisson's equation $\nabla^2 V = -6$ in this paper, meaning V is a superharmonic function. However, this method also poses a risk that V has local minima in the interior of its domain [2]. This is undesirable as local minima are harder to escape compared to saddle points. We will compare harmonic CBLFs with superharmonic CBLFs with numerical results.

IV. Results

In this section, we will explore two different reach-avoid environments, one with four small unsafe regions and one with two big unsafe regions. We will solve these environments with three different systems: Roomba, DiffDrive, and CarLikeRobot [15]. The dynamics of each of these systems are provided in Appendix I along with control inputs that minimize the expression in equation 6 for any $x \in S_{safe}$.

For every environment and system, we will compute the harmonic ($\nabla^2 V = 0$) and superharmonic ($\nabla^2 V = -6$) CBLF numerically using finite element methods. For both of the CBLFs, we compute the trajectories of the system with 1,000 different randomly initialized initial conditions and count the number of time steps (of size $\Delta t = 0.1$) it took for the system to reach S_{goal}. We will try this for deterministic ($\sigma = 0$) and stochastic ($\sigma = 0.1$) controllers and report the mean time taken to reach the goal (μ_T), the standard deviation of the time taken to reach the goal (σ_T), the number of times the system ends in an unsafe region (not included in μ_T, σ_T), and the number of times the system doesn't reach S_{goal} after 1,000 time steps (not included in μ_T, σ_T).

A. Problem I

In this problem, we explore an environment that contains a goal region near the origin and four unsafe regions in the interior of the domain.

$$S = [-1,1] \times [-1,1] \tag{14}$$
$$S_{goal} = [-0.1, 0.1] \times [-0.1, 0.1] \tag{15}$$
$$S_{unsafe} = \partial S \cup C_1 \cup C_2 \cup C_3 \cup C_4, \tag{16}$$

with the subdomains of the unsafe region given by

$$\partial S = [-1,1] \times \{-1,1\} \cup \{-1,1\} \times [-1,1] \tag{17}$$
$$C_1 = [-0.5, -0.3] \times [-0.5, -0.3] \tag{18}$$
$$C_2 = [-0.5, -0.3] \times [0.3, 0.5] \tag{19}$$
$$C_3 = [0.3, 0.5] \times [-0.5, -0.3] \tag{20}$$
$$C_4 = [0.3, 0.5] \times [0.3, 0.5] \tag{21}$$

Initial conditions are sampled as:

$$x(0) \sim U[-0.9, -0.6] \cup [0.6, 0.9], \tag{22}$$
$$y(0) \sim U[-0.9, -0.6] \cup [0.6, 0.9] \tag{23}$$
$$\theta(0) \sim U[0, 2\pi] \tag{24}$$

We derive a harmonic CLBF using finite element methods using DOLFIN [14]. We used piecewise linear trial functions and a triangular mesh. The domain has been divided into 5,000 triangular meshes of equal area. In this paper, we will let $c = 1$. This numerical solution is plotted in 2D view in figure 1 and in 3D view in figure 5.

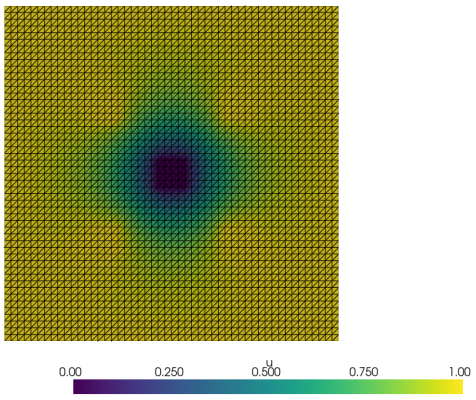

Fig. 1. Problem I with $f = 0$

Although in theory, the solution should take values less than 1.0 outside the unsafe region, the distinction is highly unclear in the numerical solution. To clarify the distinction, we replace the Laplace equation with the Poisson equation $\nabla^2 V = f$. We plotted the solution with $f = -6$ in 2D view in figure 2 and in 3D view in figure 6.

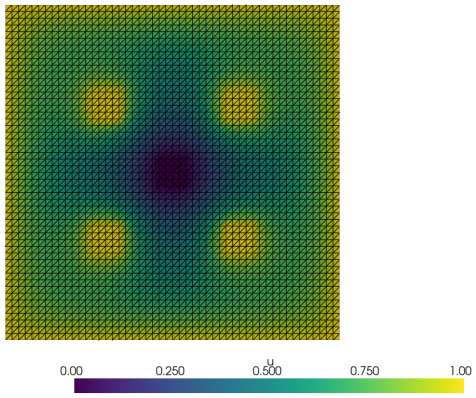

Fig. 2. Problem I with $f = -6$

This solution shows a better distinction between safe and unsafe regions, but it shows more points where $V(x)$ has a local optimum. Local optima are undesirable as $\langle f(x,u), \nabla V(x)\rangle = 0$ in those points so $V(x)$ is unlikely to show a significant decrease.

System	$\nabla^2 V$	σ	μ_T	σ_T	unsafe	no reach
Roomba	0	0	97.278	65.607	0	105
Roomba	-6	0	58.224	50.388	0	459
Roomba	**0**	**0.1**	**114.706**	**81.527**	**0**	**0**
Roomba	-6	0.1	139.606	180.247	0	185
DiffDrive	**0**	**0**	**301.403**	**216.495**	**0**	**95**
DiffDrive	-6	0	127.045	115.348	0	492
DiffDrive	0	0.1	296.800	200.713	810	0
DiffDrive	-6	0.1	186.717	157.670	709	37
CarLikeRobot	0	0	452.125	323.942	0	193
CarLikeRobot	**-6**	**0**	**277.398**	**253.148**	**0**	**47**
CarLikeRobot	0	0.1	21.200	69.889	0	930
CarLikeRobot	-6	0.1	26.887	88.174	0	920

TABLE I
RESULTS FOR PROBLEM I

The numerical results are posted in table I. Roomba achieves the best results under a harmonic CBLF with a stochastic control policy as all the trajectories converged to S_{goal}. DiffDrive achieves the best results under a harmonic CBLF with a deterministic control policy. And CarLikeRobot achieves the best results under a superharmonic CBLF with a deterministic policy.

B. Problem II

In this problem, we explore an environment that contains a goal region near the origin and two unsafe regions in the interior of the domain.

$$S = [-1,1] \times [-1,1] \tag{25}$$
$$S_{goal} = [-0.1,0.1] \times [-0.1,0.1] \tag{26}$$
$$S_{unsafe} = \partial S \cup C_5 \cup C_6, \tag{27}$$

with the subdomains of the unsafe region given by

$$\partial S = [-1,1] \times \{-1,1\} \cup \{-1,1\} \times [-1,1] \tag{28}$$
$$C_5 = [-0.5,-0.3] \times [-0.5,0.5] \tag{29}$$
$$C_6 = [0.3,0.5] \times [-0.5,0.5] \tag{30}$$

Initial conditions are sampled as:

$$x(0) \sim U[-0.9,-0.6] \cup [0.6,0.9], \tag{31}$$
$$y(0) \sim U[-0.3,0.3] \tag{32}$$
$$\theta(0) \sim U[0,2\pi] \tag{33}$$

The numerical solution to the harmonic CBLF for this problem is plotted in 2D view in figure 3 and in 3D view in figure 7.

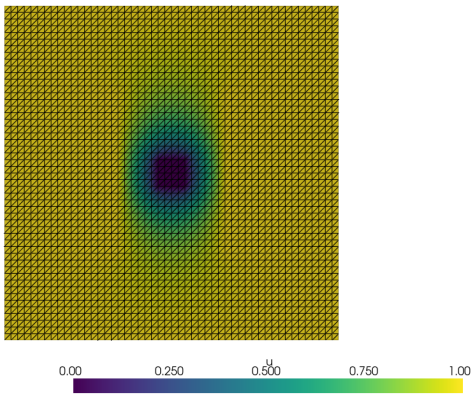

Fig. 3. Problem II with $f = 0$

To make the distinction between safe and unsafe regions clearer, we plotted the solution to the Poisson equation with $f = -6$ in 2D view in figure 4 and in 3D view in figure 8.

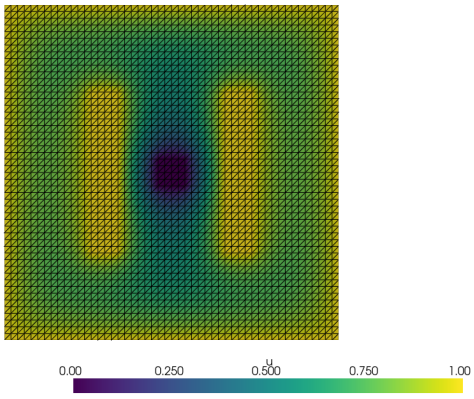

Fig. 4. Problem II with $f = -6$

System	$\nabla^2 V$	σ	μ_T	σ_T	unsafe	no reach
Roomba	0	0	85.600	54.603	12	0
Roomba	**-6**	**0**	**202.568**	**199.603**	**0**	**91**
Roomba	0	0.1	86.846	55.789	9	78
Roomba	-6	0.1	48.033	21.363	0	880
DiffDrive	**0**	**0**	**258.502**	**168.726**	**0**	**55**
DiffDrive	-6	0	96.077	28.211	0	730
DiffDrive	0	0.1	227.905	134.445	831	0
DiffDrive	-6	0.1	260.545	295.732	982	7
CarLikeRobot	**0**	**0**	**380.014**	**230.751**	**30**	**124**
CarLikeRobot	-6	0	415.246	257.179	0	570
CarLikeRobot	0	0.1	264.435	205.930	45	942
CarLikeRobot	-6	0.1	342.060	262.481	0	987

TABLE II
RESULTS FOR PROBLEM II

The numerical results are posted in table II. Roomba achieves the best results under a superharmonic CBLF with a deterministic control policy. DiffDrive achieves the best results under a harmonic CBLF with a deterministic control policy. And CarLikeRobot achieves the best results under a harmonic CBLF with a deterministic policy, although it has 30 trajectories converging to S_{unsafe} and 124 trajectories not converging to S_{goal}. Furthermore, this table shows that deterministic control policies achieve better results than stochastic control policies and have significantly more trajectories that do not converge to S_{goal}.

Tables I and II show that deterministic policies under superharmonic CBLFs always avoid S_{unsafe} as expected since the Poisson equation makes the distinction between safe and unsafe regions clearer. But it comes with the drawback that there is a high risk that the trajectory will converge to a local minimum and not reach S_{goal}. Deterministic policies generally outperform stochastic policies as the noise added to control inputs comes with the risk that it drives the trajectory away from the goal or towards an unsafe region. Despite using control inputs that solve the minimization problem in equation 6, in many episodes, the trajectory does not converge to S_{goal}. This is likely due to numerical errors associated with estimating ∇V or with solving the system dynamics.

V. CONCLUSIONS

In this paper, we introduced harmonic CBLFs that exploit the maximum principle that harmonic functions satisfy to encode the properties of CBLFs. This paper is the first to unify harmonic functions with CBLFs for an application to control theory. We select control inputs that maximize the inner product of the system dynamics with the steepest descent direction of the harmonic CBLF. This has been applied to reach-avoid problems and demonstrated a low risk of entering unsafe regions while converging to the goal region.

APPENDIX I: DYNAMICS OF EACH SYSTEM

A. Roomba

The dynamics of the Roomba are as follows:

$$\dot{x} = v\cos\theta \tag{34}$$

$$\dot{y} = v\sin\theta \tag{35}$$

$$\dot{\theta} = \omega, \tag{36}$$

where $v \in [-1, 1]$ and $\omega \in [-1, 1]$ are the control inputs. The minimization problem to equation 6 is:

$$\inf_{v\in[-1,1],\omega\in[-1,1]} (v\cos\theta, v\sin\theta)(V_x(x,y), V_y(x,y))^T \tag{37}$$

The infimum is achieved when v maximizes the negative of the magnitude of the inner product, and w is set to the direction that maximizes the magnitude of the inner product.

$$v = -\text{sign}(V_x(x,y)\cos\theta + V_y(x,y)\sin\theta) \tag{38}$$

$$w = \text{sign}(-V_x(x,y)\cos\theta + V_y(x,y)\sin\theta) \tag{39}$$

These control input pairs are used in computing the results in tables I and II.

B. DiffDrive

The dynamics of the diff-drive robot are as follows:

$$\dot{x} = (u_L + u_R)\frac{r}{2}\cos\theta \tag{40}$$

$$\dot{y} = (u_L + u_R)\frac{r}{2}\sin\theta \tag{41}$$

$$\dot{\theta} = (u_R - u_L)\frac{r}{2d}, \tag{42}$$

where $u_L \in [-1, 1]$ and $u_R \in [-1, 1]$ are the control inputs subject to the constraint $|u_L| + |u_R| \leq 1$. $r = 0.1$ and $d = 0.1$.

The minimization problem to equation 6 is:

$$\inf_{|u_L|+|u_R|\leq 1} ((u_L + u_R)\frac{r}{2}\cos\theta, (u_L + u_R)\frac{r}{2}\sin\theta) \\ (V_x(x,y), V_y(x,y))^T \tag{43}$$

The infimum is achieved when $u_L + u_R$ maximizes the negative of the magnitude of the inner product, and $u_R - u_L$ is set to the direction that maximizes the magnitude of the inner product.

$$u_R = [\text{sign}(-V_x(x,y)\cos\theta + V_y(x,y)\sin\theta) \\ - \text{sign}(V_x(x,y)\cos\theta + V_y(x,y)\sin\theta)]/2 \tag{44}$$

$$u_L = [-\text{sign}(-V_x(x,y)\cos\theta + V_y(x,y)\sin\theta) \\ - \text{sign}(V_x(x,y)\cos\theta + V_y(x,y)\sin\theta)]/2 \tag{45}$$

C. CarLikeRobot

The dynamics of the car-like robot are as follows:

$$\dot{x} = v\cos\theta \tag{46}$$
$$\dot{y} = v\sin\theta \tag{47}$$
$$\dot{\theta} = v\tan(\psi)/l \tag{48}$$
$$\dot{\psi} = w, \tag{49}$$

where $v \in [-1, 1]$ and $w \in [-1, 1]$ are the control inputs subject to the constraint $|v| \leq |w|$. $l = 0.1$. $\psi(0)$ is initialized to 0. We set $l = 0.1$.

The minimization problem to equation 6 is:

$$\inf_{|v|\leq|w|\leq 1}(v\cos\theta, v\sin\theta)(V_x(x,y), V_y(x,y))^T \tag{50}$$

The infimum is achieved when v maximizes the negative of the magnitude of the inner product, and θ is set to the direction that maximizes the magnitude of the inner product.

$$v = -\text{sign}(V_x(x,y)\cos\theta + V_y(x,y)\sin\theta) \tag{51}$$

$$w = v\text{sign}(-V_x(x,y)\cos\theta + V_y(x,y)\sin\theta) - \text{sign}(\psi) \tag{52}$$

APPENDIX II: 3D PLOTS OF HARMONIC CBLFS

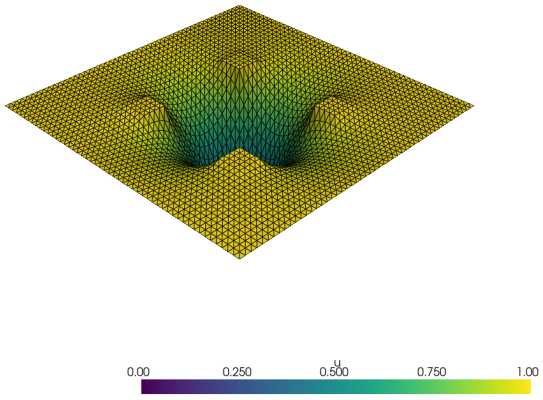

Fig. 5. Problem I with $f = 0$

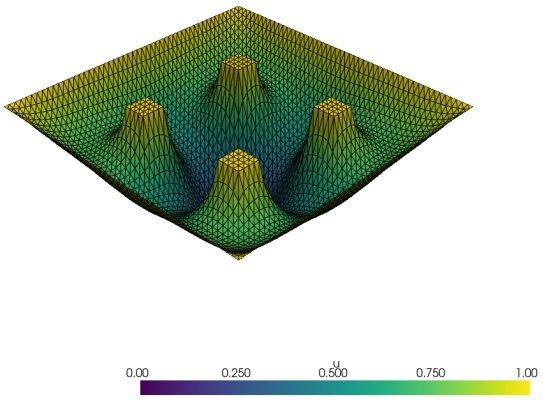

Fig. 6. Problem I with $f = -6$

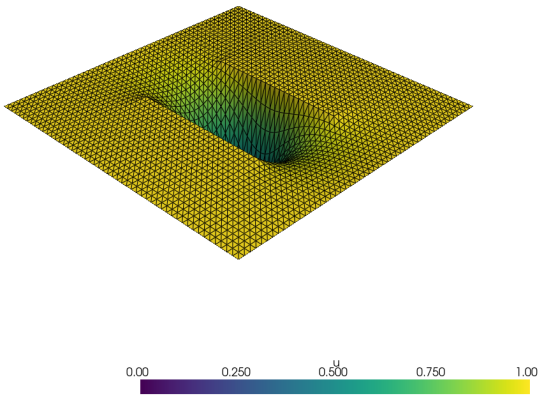

Fig. 7. Problem II with $f = 0$

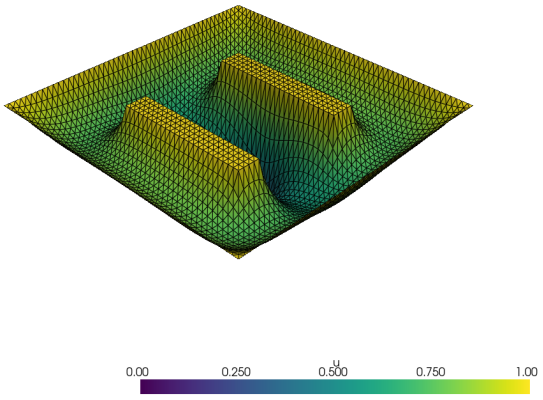

Fig. 8. Problem II with $f = -6$

ACKNOWLEDGMENT

The authors would like to thank Professor Yash Pant from the Department of Electrical Engineering, University of Waterloo for providing us with feedback for this paper.

REFERENCES

[1] Dario Amodei Alex Ray, Joshua Achiam. Benchmarking safe exploration in deep reinforcement learning. *OpenAI*, 2019.

[2] Sheldon Axler, Paul Bourdon, and Ramey Wade. *Harmonic function theory*, volume 137. Springer Science & Business Media, 2013.

[3] Lukas Brunke, Melissa Greeff, Adam W. Hall, Zhaocong Yuan, Siqi Zhou, Jacopo Panerati, and Angela P. Schoellig. Safe learning in robotics: From learning-based control to safe reinforcement learning. *Annual Review of Control, Robotics, and Autonomous Systems*, 5(1):411–444, 2022.

[4] Jason Choi, Fernando Castaneda, Claire J Tomlin, and Koushil Sreenath. Reinforcement learning for safety-critical control under model uncertainty, using control lyapunov functions and control barrier functions. *arXiv preprint arXiv:2004.07584*, 2020.

[5] Christopher I Connolly, J Brian Burns, and Rich Weiss. Path planning using laplace's equation. In *Proceedings., IEEE International Conference on Robotics and Automation*, pages 2102–2106. IEEE, 1990.

[6] Desong Du, Shaohang Han, Naiming Qi, Haitham Bou Ammar, Jun Wang, and Wei Pan. Reinforcement learning for safe robot control using control lyapunov barrier functions, 2023.

[7] Rong Ge, Furong Huang, Chi Jin, and Yang Yuan. Escaping from saddle points — online stochastic gradient for tensor decomposition, 2015.

[8] David Gilbarg and Neil S. Trudinger. *Elliptic Partial Differential Equations of Second Order, Second Edition*. Springer-Verlag, 1983.

[9] Thomas Gurriet, Andrew Singletary, Jacob Reher, Laurent Ciarletta, Eric Feron, and Aaron Ames. Towards a framework for realizable safety critical control through active set invariance. In *2018 ACM/IEEE 9th International Conference on Cyber-Physical Systems (ICCPS)*, pages 98–106. IEEE, 2018.

[10] Nicolas Heess, Dhruva Tb, Srinivasan Sriram, Jay Lemmon, Josh Merel, Greg Wayne, Yuval Tassa, Tom Erez, Ziyu Wang, SM Eslami, et al. Emergence of locomotion behaviours in rich environments. *arXiv preprint arXiv:1707.02286*, 2017.

[11] Ching-An Cheng Hoai-An Nguyen. Provable reset-free reinforcement learning by no-regret reduction. *International Conference on Machine Learning*, 2023.

[12] Jin-Oh Kim and Pradeep Khosla. Real-time obstacle avoidance using harmonic potential functions. *IEEE Transactions on Robotics and Automation*, June 1992.

[13] Jason D. Lee, Max Simchowitz, Michael I. Jordan, and Benjamin Recht. Gradient descent converges to minimizers, 2016.

[14] A. Logg and G. N. Wells. Dolfin: Automated finite element computing. *ACM Transactions on Mathematical Software 37*, 2010.

[15] Kevin M. Lynch and Frank C. Park. *Modern Robotics: Mechanics, Planning, and Control*. Cambridge University Press, 2017.

[16] Murray H. Protter and Hans F. Weinberger. *Maximum Principles in Differential Equations*. Springer, 1984.

[17] S. Kawamura S. Akishita and K. Hayashi. Laplace potential for moving obstacle avoidance and approach of a mobile robot. In *I990 Japan-USA Symposium on Flexible Automation, A Pacific Rim Conference*, 1990.

[18] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[19] A Yanushauskas. The zeros of the gradient and the hessian of an harmonic function. *Siberian Mathematical Journal*, 10(3):497–501, 1969.