# A Safe Data-Driven Optimization Approach for Robot Navigation in Dynamic Environments

Dadi Hrannar Davidsson, Lasse Due Hornshøj, Søren Haugaard, Adrian Iribar, Oghuz Madinali, Sanjula De Silva, Rahul Misra, Henrik Schiøler, and Shahab Heshmati-Alamdari, *Member IEEE*

***Abstract*— This paper presents a novel, data-driven motion planning strategy for autonomous mobile robots navigating in dynamic environments with human interactivity. The proposed approach utilizes a receding-horizon optimization framework that integrates predictive models of the robot motion with unknown-form safety constraints encapsulating human movement uncertainties, and complex dynamics of human-human and human-robot interactions. The functional form of constraints is unknown instead, we obtain only measurements and gradients of the constraint i.e. 1st order online optimization. Specifically, data-driven log barrier functions enforce safety constraints by penalizing closeness to constraint boundaries. The proposed strategy enables reliable, efficient, and safe robot navigation in high-density environments, making it particularly suitable for applications in pedestrian and other interactive spaces. Finally, realistic simulation studies validate the effectiveness of the proposed framework in balancing safety with operational efficiency.**

## I. Introduction

The deployment of mobile robots in various fields has expanded significantly in recent years, with applications spanning from multi-robot logistics and service tasks to healthcare and urban mobility [1], [2]. However, integrating autonomous mobile robots into crowded pedestrian areas poses unique challenges, especially regarding safe and smooth operation in dynamic environments [3]. A key obstacle is robust and precise motion planning that proactively and responsively accounts for human dynamics to ensure safety and fluidity [4]. In crowded environments, uncertain human intentions require robust trajectory estimation with explicit uncertainty modeling integrated into planning algorithms to ensure safety and reliability in such dynamic environments [5]. Many navigation methods treat pedestrians as independent agents, non-interactive entities, causing uncertainty to escalate [6] and making robots susceptible to the Freezing Robot Problem (FRP) [7] (Fig. 1). Additionally, implicit pedestrian interactions involving frequent speed and direction adjustments complicate motion dynamics, further increasing FRP risk [8]. While several models attempt human prediction [9], most neglect human-human interactions and lack quantifiable safety guarantees [5]. In addition, a substantial body of work addresses multi-agent navigation without relying on predicted human trajectories. When prediction is used, two main approaches are identified [4]: (i) decoupled methods, which predict human motion independently of planning (often leading to FRP) [6], and (ii) coupled methods, which model interactions during joint trajectory planning [10]. Coupled or interactive approaches can be explicit, using structured models like geometry or game theory [11], or implicit, relying on data-driven techniques such as Inverse Reinforcement Learning [12] and Deep Learning [13]. However, deep learning models, in addition to lacking formal safety guarantees, often lack goal-directed planning, which limits their use in safe navigation [5]. In this work, a coupled and implicit motion planning strategy is proposed for autonomous robot navigation in dynamic environments involving human interaction, such as crowded pedestrian areas, with formal safety guarantees. A surrogate optimization problem is constructed using data samples that include both the value and gradient of an unknown constraint function. The cost function comprises a known term, the Euclidean distance to a predefined goal, and a data-driven log barrier function. We assume access to both the value and the gradient of the unknown constraint function at any given query point. Thus, our approach falls under the category of first-order online optimization methods [14], which have demonstrated efficient convergence properties for non-smooth convex optimization problems [15]. An adaptive step size ensures safety with respect to the unknown constraint [16]. Despite the appeal of such data-driven optimization methods, to the best of the authors' knowledge, this is the first application of such methods to systems with dynamics, a hallmark of control problems. We address this gap by combining first-order online optimization methods with a receding horizon scheme and designing a projected gradient descent method that enforces dynamical constraints at each iteration at each iteration of the online optimization. Moreover, we prove recursive feasibility and cost convergence.

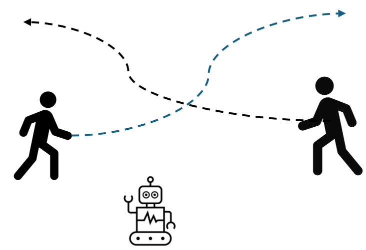

Fig. 1: Illustration of the 'freezing robot' problem in dynamic environments with human interaction. Effective navigation in shared spaces relies on understanding intra-human dynamics and human-robot interactions for safe path planning and collision avoidance.

Dadi, Lasse, Søren, Oghuz, Adrian, Gayath, Rahul, Henrik and Shahab are with the Section of Automation & Control, Department of Electronic Systems, Aalborg University, Denmark, Email: {ddavid20, lhorns20, shauga20, airiba23, omadin23, gdesil23}@student.aau.dk, {rmi, henrik, shhe}@es.aau.dk.

Simulation results demonstrate the method's ability to handle unpredictable crowd behavior, even without an explicit model of the constraint function.

## II. Problem formulation

Consider a mobile robot with linear dynamics:

$$\dot{x}(t) = Ax(t) + Bu(t), \tag{1}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^m$ the control input, and matrices $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ define the system evolution. The control objective is to find an optimal sequence of inputs minimizing a cost function $c : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}$, subject to constraints $h : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$ represented by inequalities $h(x,u) \le 0$. Practically, $h(x,u)$ captures safety constraints in environments with dynamic human agents acting as moving obstacles. A key challenge is that the functional form of $h(x,u)$ is unknown; we only access instantaneous measurements at specific states and controls. This approach reflects the complexity inherent in human-human and human-robot interactions, which are unpredictable and difficult to explicitly model. By utilizing data-driven measurements, such as positions obtained from sensors like Lidar, the framework dynamically adapts to these evolving interactions, ensuring safety without explicit modeling of human movement patterns. Discretizing dynamics (1) via Euler approximation with step size $\Delta$ leads to the following discrete-time optimal control formulation:

$$\min_{u_k} \sum_{k=1}^{\tau-1} c_k\left(x_k, u_k\right) + c_\tau\left(x_\tau\right) \tag{2a}$$

$$\text{s.t. } x_{k+1} = x_k + \Delta\left(Ax_k + Bu_k\right), \tag{2b}$$

$$h_k(x_k, u_k) \le 0, \tag{2c}$$

where $\tau$ is the terminal time, $c_\tau(x_\tau)$ denotes the terminal cost, and $k$ is the discrete time step. We assume the costs $c_k(x_k,u_k)$ and the unknown constraint $h_k(x_k,u_k)$ are twice differentiable with respect to the decision variable $u_k$. Additionally, we assume the initial state $x_0$ lies within the interior of the feasible region defined by the constraint (2c).

## III. Receding Horizon Control scheme

We propose solving the discrete-time optimal control problem (2) via a receding horizon control approach over a finite horizon $N$. Define the state sequence starting from time step $k+1$ as follows:

$$\bar{x} = [x(k+1 \mid k)^T, x(k+2 \mid k)^T \ldots, x(k+N \mid k)^T]^T.$$

Similarly, the approximately optimal control input $u_k$ is chosen as the first element from the control sequence

$$\bar{u} = [u(k \mid k)^T, u(k+1 \mid k)^T, \ldots, u(k+N-1 \mid k)^T]^T,$$

where $\bar{u}$ represents the sequence of optimal decision variables for the following optimization problem:

$$\min_{\bar{x},\bar{u}} \quad V(\bar{x},\bar{u}) = \min_{\bar{u}} \sum_{j=1}^{N-1} c\left(x_{k+j|k}, u_{k+j|k}\right) + c_N\left(x_{k+N|k}\right) \tag{3a}$$

$$\text{s.t.} \quad x_{k+j|k} = x_{k+j-1|k} + \Delta\left(Ax_{k+j-1|k} + Bu_{k+j-1|k}\right), \quad \forall j = 1,\ldots,N-1, \quad x(k|k) = x \tag{3b}$$

$$h_{k+j-1|k}(x_{k+j-1|k}, u_{k+j-1|k}) \le 0, \quad \forall j = 1,\ldots,N-1, \tag{3c}$$

$$x_{k+N|k} \in X_f, \quad u_k \in U. \tag{3d}$$

Here, $x_0$ represents the initial system configuration in (3), corresponding to the state measurement at time step $k$. An admissible control law refers to a state-feedback controller satisfying constraints (3b) and (3c). Our objective is to design a control policy that optimally drives the system to a predefined terminal set $X_f$, which, in robot navigation scenarios, consists of goal states or waypoints.

**Assumption III.1** ( [17])**.** The following standard assumptions are required,

1) The origin of the dynamical system (1) is the steady state to be stabilized and lies within the set $X_f$.
2) The stage cost function $c(x,u)$ is positive, and is zero only at $c(0,0)$.
3) There is a local control law $K(x)$ such that the terminal set $X_f$ remains invariant, and the state and input constraints, $x \in X_f$ and $K(x) \in U$ respectively are satisfied.
4) The terminal cost $c_N(x)$ acts as a Lyapunov function within the set $X_f$, fulfilling the inequality $c_N(x(k+1)) - c_N(x(k)) \le -c(x,K(x))$ for every $x_k \in X_f$ subject to the local control law $K(x)$.
5) The initial state $x_0$ for (3) is feasible.

Since the functional form of $h$ is unknown, directly solving (3c) is infeasible. In practice, $h$ represents a safety function capturing the complex dynamics of human-human and human-robot interactions. In the next subsection, we address how to handle constraint (3c).

### A. Satisfying the unknown constraint

To solve the constrained optimization problem (3), we transform it into an equivalent unconstrained problem by enforcing the unknown constraints via a barrier function. Specifically, the barrier function $\mathrm{B}_\eta(\bar{x},\bar{u})$, estimated subsequently, is defined as:

$$\mathrm{B}_\eta(\bar{x},\bar{u}) := V(\bar{x},\bar{u}) + \eta \sum_{j=1}^{N-1} \left(-\log(-h_{k+j-1|k}(x_{k+j-1|k}, u_{k+j-1|k}))\right), \tag{4}$$

where $\eta$ is a tuning parameter that shall be tuned adaptively as we shall see later. The gradient of $B_\eta(\bar{x},\bar{u})$ is calculated as follows (using the chain rule),

$$\nabla \mathrm{B}_\eta(\bar{x},\bar{u}) := \nabla V(\bar{x},\bar{u}) + \eta \sum_{j=1}^{N-1} \frac{\nabla h_{k+j-1|k}(x_{k+j-1|k},u_{k+j-1|k})}{-h_{k+j-1|k}(x_{k+j-1|k},u_{k+j-1|k})}. \quad (5)$$

Since the functional form of $h_{k+j-1|k}(x_{k+j-1|k},u_{k+j-1|k})$ is unknown, and we solely have access to data samples or observations, it is necessary to estimate the barrier term $\mathrm{B}_\eta(\bar{x},\bar{u})$ and its gradient $\nabla \mathrm{B}_\eta(\bar{x},\bar{u})$ from the collected observations. To that end, we assume the local smoothness of the Barrier function or more specifically, for any given $a,b \in \mathbb{R}^n \times \mathbb{R}^m$,

$$\left\|\nabla \mathrm{B}_\eta(a) - \nabla \mathrm{B}_\eta(b)\right\| \leq M \|a-b\|. \quad (6)$$

The paper [16] provides a method to compute the Lipschitz constant $M$. Let $H_k$ represent measurements of the unknown constraint function $h(x_k,u_k)$ at a given query point $(x_k,u_k)$, which vary throughout the horizon in (3) as states evolve according to (3b). In practice, $H_k$ corresponds to real-time distance measurements obtained from onboard sensors (e.g., Lidar). Additionally, assuming the robot has an IMU for measuring body velocity, and a vision system to detect human velocities, we obtain the gradient $\nabla H_k(x_k,u_k)$ at each time step $k$. Thus, at each time step $k$, we have the observations:

$$O_k = [H_k(x_k,u_k), \nabla H_k(x_k,u_k)]. \quad (7)$$

Let the estimated barrier function given the observations $O_k(H_k,\nabla H_k)$ be denoted as $\hat{\mathrm{B}}_\eta(\bar{x},\bar{u})$. More specifically,

$$\hat{\mathrm{B}}_\eta(\bar{x},\bar{u}) = V(\bar{x},\bar{u}) - \eta \sum_{j=1}^{N-1} \log(-H_k(x_{k+j-1},u_{k+j-1})). \quad (8)$$

It should be noted that the term $\eta$ is adaptively tuned on the basis of each $-H_k(x_k,u_k)$ and a simple adaptive rule can be $\eta = e^{-y^2}$ where $y = \sum_{j=1}^{N-1} -H_k(x_{k+j-1},u_{k+j-1})$. This adaptive rule ensures that for feasible iterates the value of the log term is $\approx 0$ since the value of $\eta \to 0$ in (8) for $H_k(x_k,u_k) > 2$.

Recall that in $\hat{\mathrm{B}}_\eta(\bar{x},\bar{u})$, only the data-driven log barrier is estimated from observations, while the cost $V(\bar{x},\bar{u})$ is a known quadratic distance to the goal. We assume unbiased measurements, i.e., $h_k(x_k,u_k) \approx \mathbb{E}[H_k(x_k,u_k)]$, where the expectation is with respect to the measurement noise. We further assume that the measurements $H_k(x_k,u_k)$ have a bounded variance $\sigma^2$ [16]. Similar assumptions apply to gradient measurements $\nabla H_k(x_k,u_k)$. Let $\mathrm{G}(H_k,\nabla H_k)$ denote the gradient of the estimated barrier function i.e. $\nabla\hat{\mathrm{B}}_\eta(\bar{x},\bar{u})$. Given $O_k$, $\mathrm{G}(H_k,\nabla H_k)$ is estimated as:

$$\mathrm{G}(H_k,\nabla H_k) := \nabla V(\bar{x},\bar{u}) + \eta \sum_{j=1}^{N-1} \frac{\nabla H_{k+j-1|k}(x_{k+j-1|k},u_{k+j-1|k})}{-H_{k+j-1|k}(x_{k+j-1|k},u_{k+j-1|k})}, \quad (9)$$

where $\nabla V(\bar{x},\bar{u})$ is known, as it corresponds to the gradient of the Euclidean distance between the robot and the goal. Next, we address enforcing the dynamic constraint (3b) and subsequently combine both constraints in a gradient descent scheme.

### *B. Satisfying the known dynamics constraint*

For simplicity, we compactly rewrite the constraint equations in (3b) as:

$$\hat{A}([0^T, x(k+1\,|\,k)^T, \ldots, x(k+N-1\,|\,k)^T]^T + [x(k\,|\,k)^T, 0^T, \ldots, 0^T]^T) + \tilde{B}\bar{u} = [x(k+1\mid k)^T, \ldots, x(k+N\mid k)^T]^T = \bar{x}.$$

where matrices $\hat{A}$ and $\tilde{B}$ are defined appropriately. More concisely:

$$(\hat{A}\tilde{C} - I)\bar{x} + \tilde{B}\bar{u} = \tilde{A}\bar{x} + \tilde{B}\bar{u} = \bar{y} \quad (10)$$

where $0 \in \mathbb{R}^n$ is the vector of zeros elements, and

$$\bar{y} = -\hat{A}[x(k\mid k)^T, 0^T, \ldots, 0^T]^T$$
$$\tilde{C}\bar{x} = [0^T, x(k+1\mid k)^T, \ldots, x(k+N-1\mid k)^T]^T$$

We define $\mathscr{A}_{\bar{y}}$ as the set of state and input sequences satisfying the dynamical constraint (10), explicitly given by:

$$\mathscr{A}_{\bar{y}} = \{(\bar{x},\bar{u}) \mid \tilde{A}\bar{x} + \tilde{B}\bar{u} = \bar{y}\}. \quad (11)$$

The optimization problem (3) can now be equivalently expressed as the following data-driven formulation:

$$\min_{\bar{x},\bar{u}} \quad \hat{\mathrm{B}}_\eta(\bar{x},\bar{u}) \quad (12a)$$

$$\text{s.t.} \quad (\bar{x},\bar{u}) \in \mathscr{A}_{\bar{y}}. \quad (12b)$$

The known dynamic constraints (12b) are enforced using a projection operator projecting the outputs of the data-driven log barrier onto the set of feasible inputs. Since the considered dynamical system is linear, the set $\mathscr{A}_{\bar{y}}$ is convex in the decision variables $\bar{z} := (\bar{x},\bar{u})$. We define the projection operator $\Pi$ as:

$$\Pi_{\mathscr{A}_{\bar{y}}}(z) = \arg\min_{z' \in \mathscr{A}_{\bar{y}}} \left\|z' - z\right\| \quad (13)$$

Due to the convexity of the set (11), the projection (13) has a unique solution [18]. For brevity, we denote the decision variable of problem (12) as $\bar{z} = (\bar{x},\bar{u})$.

### *C. Projected Gradient Descent scheme for* (12)

We shall now focus on solving the optimization problem given by (12). Let $t = 1,\ldots,T$ denote the $t^{th}$ iterate of the algorithm solving the optimization problem (12) as shown in Fig.2. By using the projection operator (13), we unconstrain the optimization problem (12). The variable $\bar{z}_{t+1}$ is then updated with an adaptive, safety-aware step size $\gamma_t$ as follows:

$$\bar{z}_{t+1} = \bar{z}_t - \Pi_{\mathscr{A}_0} \gamma_t \mathrm{G}(H_k,\nabla H_k). \quad (14)$$

We shall now design the adaptive safe step size following the ideas of [16]. Let $\alpha_t^j := -h_{k+j|k}(x_{k+j|k}, u(j)_t)$ represent the unknown constraint, where $u(j)_t$ is the $j^{th}$ component of $\bar{z}_t$ corresponding to $u_{k+j|k}$. Let $\theta_t^j$ represent the inner product

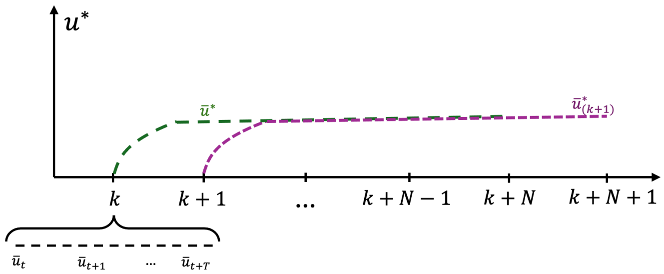


Fig. 2: The figure illustrates the relation between the time scales $t$ and $k$.

between the gradient of the unknown constraint function and the gradient step of the update equation (14),

$$\theta_t^j = \left\langle \nabla h_{k+j-1|k}(x_{k+j-1|k}, u_{k+j-1|k}), \frac{\mathrm{G}(H_k, \nabla H_k)}{\|\mathrm{G}(H_k, \nabla H_k)\|} \right\rangle. \tag{15}$$

The adaptive safe step size $\gamma_t$ is chosen as follows,

$$\gamma_t = \min\left\{ \min_{j=1,\ldots,N-1} \left\{ \frac{\underline{\alpha_t^j}}{2 \mid \theta_t^j \mid + \sqrt{\underline{\alpha_t^j} M}} \right\} \frac{1}{\|\mathrm{G}\|}, \frac{1}{M} \right\}, \tag{16}$$

where $G = \mathrm{G}(H_k, \nabla H_k)$, $\underline{\alpha_t^j}$ is the lower bound on $\alpha_t^j$ and $\mid \theta_t^j \mid$ is the absolute value of $\theta_t^j$. See paper [16] for the derivation of the same. In essence, each update moves the decision variable in the direction of $G(H_k, \nabla H_k)$, decreasing the barrier function $\hat{\mathrm{B}}_\eta(\bar{z})$ and thus enhancing safety. Since $\hat{\mathrm{B}}_\eta(\bar{z})$ sharply increases near the boundary of constraint (3c), moving toward decreasing $\hat{\mathrm{B}}_\eta(\bar{z})$ ensures constraint satisfaction. However, overly large steps may bypass local minima, inadvertently increasing $\hat{\mathrm{B}}_\eta(\bar{z})$. To prevent this, the adaptive step size $\gamma_t$ needs to be modified to be small enough to correspond to the norm of the estimated gradient $\mathrm{G}(H_k, \nabla H_k)$. Specifically, the control sequence is updated using the adaptive step size $\gamma_t$ defined in (16), which ensures the step size remains bounded by the Lipschitz constant $M$. This guarantees that each control update respects the safety constraint $h(x,u) < 0$. We now state the main theorem:

**Proposition 1.** *Consider the surrogate optimization problem* (12) *that needs to be solved sequentially for optimal control of system* (2). *The objective function $\hat{\mathrm{B}}_\eta(\bar{z})$ consists of costs $V(\bar{z})$ and the log-barrier $\eta \sum_{j=1}^{N-1} \log(-H_k(x_{k+j-1}, u_{k+j-1}))$ with respect to the measurements $H_k, \nabla H_k$. Furthermore, the considered optimization problem is subject to the constraint $\bar{z} \in \mathscr{A}_{\bar{y}}$. Then, the following statements are true:*

- *The control law $\bar{z}_{t+1}$ updated using the projected gradient descent scheme* (14) *solves the optimization problem for each time instant k.*
- *The control input is feasible at each time instant k.*
- *The control sequence guarantees convergence of the cost $\hat{\mathrm{B}}_\eta$ to its optimal value.*

*Proof.* We prove the proposition in two parts: showing that (1) projected gradient descent in the direction of $\hat{\mathrm{B}}_\eta(\bar{z})$ ensures safety while maintaining feasibility of the dynamical constraints, and (2) cost reduction and convergence of the state trajectory to the goal.

**Gradient descent based update:**

The KKT conditions for optimization problem (12) are

$$\nabla \hat{\mathrm{B}}_\eta(\bar{z}) + [\tilde{A}\tilde{B}]^T \nu^\star = 0, \tag{17a}$$
$$\tilde{A}\bar{x} + \tilde{B}\bar{u} - \bar{y} = 0, \tag{17b}$$
$$\nu_j^\star \geq 0, \quad \forall j = 1, \ldots, N-1, \tag{17c}$$
$$\nu_j^\star (\tilde{A}\bar{x} + \tilde{B}\bar{u} - \bar{y})_j = 0, \tag{17d}$$

where $\nu^\star$ the lagrange multiplier associated with the dynamic constraint (17b). It should be noted that only the two former KKT conditions i.e., (17a) and (17b) are relevant for equality constraints considering the given problem of (12). Now recall the definition of the set $\mathscr{A}_{\bar{y}}$

$$\mathscr{A}_{\bar{y}} = \{(\bar{x}, \bar{u}) \mid \tilde{A}\bar{x} + \tilde{B}\bar{u} = \bar{y}\}.$$

In addition, we define specifically:

$$\mathscr{A}_0 = \{(\bar{x}, \bar{u}) \mid \tilde{A}\bar{x} + \tilde{B}\bar{u} = 0\}. \tag{18}$$

This definition enables us to express the problem in term of projection onto $\mathscr{A}_0$, which is the set of all points satisfying the dynamic constraint. The gradient descent scheme updates the decision variable $\bar{z}$ by moving in the direction of the gradient $\mathrm{G}(H_k, \nabla H_k)$, while projecting onto the constraint set $\mathscr{A}_0$ to maintain feasibility. Therefore, We consider the following gradient descend scheme:

$$\bar{z}_{t+1} = \bar{z}_t - \Pi_{\mathscr{A}_0} \gamma_t \mathrm{G}(H_k, \nabla H_k) \tag{19}$$

where $\Pi_{\mathscr{A}_0}$ is the projection operator onto $\mathscr{A}_{\bar{y}}$ i.e. that keeps the iterates within the feasible set $\mathscr{A}_0$, where it is assumed that $\bar{z}_0 \in \mathscr{A}_{\bar{y}}$. For any equilibrium $\bar{z}_\infty$ of the recursion (19) (i.e., when the solution is no longer changing), we have

$$\Pi_{\mathscr{A}_0} \mathrm{G}(H_k, \nabla H_k) = 0 \tag{20}$$

and $\bar{z}_\infty \in \mathscr{A}_{\bar{y}}$ (feasible) from induction over $t$. To perform the projection, assume an orthonormal basis $\{b_1, ..b_{(N-1)n}\}$ for the row space of the constraint matrix $[\tilde{A}\tilde{B}] = [r_1^T, .., r_{(N-1)n}^T]^T$. This basis allows us to represent the projection of $\mathrm{G}(H_k, \nabla H_k)$ onto $\mathscr{A}_0$ as:

$$\Pi_{\mathscr{A}_0} \mathrm{G}(H_k, \nabla H_k) = \mathrm{G}(H_k, \nabla H_k) - \sum_{i=1}^{(N-1)n} d_i b_i \tag{21}$$

where $d_i$ are coefficients that capture the components of $\mathrm{G}(H_k, \nabla H_k)$ in the direction of each basis vector $b_i$. These components are determined to ensure that the projected gradient lies within the feasible space defined by $\mathscr{A}_0$. More precisely, the coefficient $d_i$ are determined to ensure:

$$[\tilde{A}\tilde{B}] \Pi_{\mathscr{A}_0} \mathrm{G}(H_k, \nabla H_k) = 0. \tag{22}$$

Since $\{b_1, ..b_{(N-1)n}\}$ spans the row space of $[\tilde{A}\tilde{B}]$, (22) is equivalent to :

$$< b_j, \Pi_{\mathscr{A}_0} \mathrm{G}(H_k, \nabla H_k) > \tag{23}$$
$$= < b_j, \mathrm{G}(H_k, \nabla H_k) > - \sum_{i=1}^{(N-1)n} d_i < b_j, b_i > \tag{24}$$
$$= < b_j, \mathrm{G}(H_k, \nabla H_k) > - d_j = 0 \quad \forall j \in \{1, .., (N-1)n\} \tag{25}$$

yielding

$$d_j = < b_j, \mathrm{G}(H_k, \nabla H_k) > \tag{26}$$

where $\langle \alpha, \beta \rangle$ denotes the inner product of vectors $\alpha$ and $\beta$. Again, since $\{b_1, ..b_{(N-1)n}\}$ spans the row space of $[\tilde{A}\tilde{B}]$ we have $b_i = \sum_{j=1}^{(N-1)n} a_{i,j} r_j$, where $\{a_{i,j}\}$ are the coefficients of the coordinate transform between $\{b_1, ..b_{(N-1)n}\}$ and $\{r_1, .., r_{(N-1)n}\}$. With (20) and (21) we can then express the gradient $\mathrm{G}(H_k, \nabla H_k)$ as linear combination of row vectors:

$$\mathrm{G}(H_k, \nabla H_k) = \sum_{i=1}^{(N-1)n} d_i \sum_{j=1}^{(N-1)n} a_{i,j} r_j = \sum_{j=1}^{(N-1)n} \sum_{i=1}^{(N-1)n} d_i a_{i,j} r_j = \sum_{j=1}^{(N-1)n} \nu_j^{\star} r_j \tag{27}$$

where $\nu_j^{\star}$ are identified as Lagrange multipliers. Thus, the KKT conditions (17) apply to the equilibrium point. To address the feasibility of the receding horizon approach, we follow a method similar to that in [19], [20]. Initially, we consider the current state $x_k$ and the optimal control sequence $[u^{\star}(1), u^{\star}(2), \ldots, u^{\star}(N-1)]$ obtained by solving (12). At the next time step, with the updated state $x_{k+1}$, we extend this control sequence by appending the terminal controller $K(x_N^{\star})$. The resulting control sequence, $[u^{\star}(1), u^{\star}(2), \ldots, u^{\star}(N-1), K(x_N^{\star})]$, remains feasible (albeit generally sub-optimal) due to the assumption specified in (III.1). Consequently, the optimization problem (12) retains feasibility at each time step.

**Convergence to the goal via cost reduction:** Goal convergence follows if the optimal value function $J^{\star}(x_k) := \hat{\mathrm{B}}_{\eta}(z^{\star})$ can be shown to be a Lyapunov function satisfying:

$$J^{\star}(x_{k+1}) - J^{\star}(x_k) < 0, \quad \forall x \neq 0.$$

At the outset, recall that due to the properties of $\eta$, the optimal $J^{\star}(x_k) \approx V(z^{\star})$ as the log barrier term will be negligible owing to the feasibility of the optimal solution. Therefore, for the optimal control sequence $[u^{\star}(1), u^{\star}(2), \ldots, u^{\star}(N-1)]$ calculated at $x_k$, the optimal value function is the following as per (3),

$$J^{\star}(x_k) = \sum_{k=1}^{N-1} c_k(x_k^{\star}, u^{\star}(k))) + c_N(x_N^{\star}),$$

whereas for the feasible (but sub-optimal) control sequence $[u^{\star}(1), u^{\star}(2), \ldots, u^{\star}(N-1), K(x_N^{\star})$ for $x_{k+1}$, we have the following sub-optimal value function,

$$J'(x_{k+1}) = \sum_{k=2}^{N} c_k(x_k^{\star}, u^{\star}(k))) + c_N(x_{N+1}),$$

where $x_{N+1}$ is obtained by applying $K(x_N^{\star})$ at state $x_N^{\star}$. Due to the sub-optimality,

$$J^{\star}(x_{k+1}) \leq J'(x_{k+1}),$$

and we substitute $J'(x_{k+1})$ with $J^{\star}(x_k) - c_1(x_1^{\star}, u^{\star}(1)) + c(x_N^{\star}, K(x_N^{\star}))$, where the last term is the cost due to the rest of the trajectory after time $N$ (or after the trajectory enters the terminal set). Assumption III.1 includes the following inequality for the terminal set, $c_N(x(k+1)) - c_N(x(k)) \leq -c(x, K(x))$ for every $x_k \in X_f$. Thus, we can neglect the last term and conclude the following,

$$J^{\star}(x_{k+1}) - J^{\star}(x_k) < -c_1(x_1^{\star}, u^{\star}(1)), \quad \forall x \neq 0.$$

■

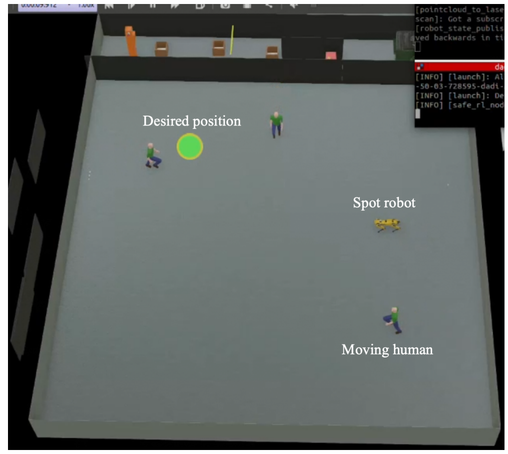


Fig. 3: Developed simulation environment in Webots involving a Spot robot in a workspace with moving human agents.

## IV. Simulation Studies

Real-time simulations were conducted to validate the proposed approach using a Webots-based environment [21] integrated with ROS. A Spot legged robot operated in a constrained workspace containing static obstacles and randomly moving human agents (see Fig. 3). The robot followed a sequence of waypoints, with human movements being unknown and non-predefined. Spot was modeled with holonomic dynamics, $\dot{x}(t) = u(t)$, and equipped with LiDAR for detecting obstacles and humans. Human agents were randomly placed with randomly assigned directions to simulate natural behavior. This setup enabled dynamic navigation and real-time adaptation to environmental changes. Simulation results for the first waypoint (Fig. 4–5) highlight the system's dynamic response and performance. The robot smoothly slows down to yield when its path overlaps with the human's trajectory and resumes once a safe distance is reached. A video demonstrating various simulation scenarios is available at: `https://youtu.be/29TS0rNJo_s`.

## V. Conclusion

In this paper, we have introduced an online learning framework for safe and efficient robot navigation in dynamic, human-interactive environments. Our approach integrates data-driven safety constraints of unknown functional forms through log barrier functions with adaptive step sizing, extending the first-order online optimization method to linear dynamical systems via a receding horizon scheme. We utilized projected gradient descent to enforce the constraints associated with robot dynamics and established proof for both recursive feasibility and cost convergence within the context of receding horizon optimization. Simulation results validate the effectiveness of our method, showing collision-free and smooth goal-directed navigation. Future work will

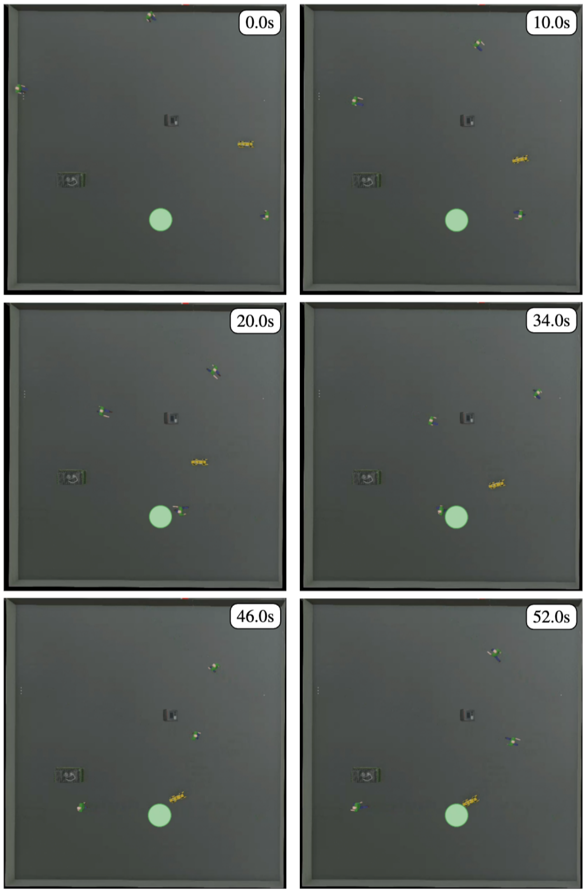


Fig. 4: The evolution of the system for waypoint 1.

be to extend our framework to nonlinear robot dynamics, enhancing robustness against uncertainties, and performing real-world experimental validations.

## ACKNOWLEDGMENT

This research was supported by the Aalborg Robotics Challenge (ARC), an initiative at Aalborg University that supports multidisciplinary research, education, and innovation.

## REFERENCES

[1] A. Nikou, S. Heshmati-Alamdari, C. K. Verginis, and D. V. Dimarogonas, "Decentralized abstractions and timed constrained planning of a general class of coupled multi-agent systems," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 990–995.

[2] P. Marmaglio, D. Consolati, C. Amici, and M. Tiboni, "Autonomous vehicles for healthcare applications: A review on mobile robotic systems and drones in hospital and clinical environments," *Electronics*, vol. 12, no. 23, p. 4791, 2023.

[3] P. Salvini, D. Paez-Granados, and A. Billard, "Safety concerns emerging from robots navigating in crowded pedestrian areas," *International Journal of Social Robotics*, vol. 14, no. 2, pp. 441–462, 2022.

[4] C. Mavrogiannis, F. Baldini, A. Wang, D. Zhao, P. Trautman, A. Steinfeld, and J. Oh, "Core challenges of social robot navigation: A survey," *ACM Transactions on Human-Robot Interaction*, vol. 12, no. 3, pp. 1–39, 2023.

[5] L. Lindemann, M. Cleaveland, G. Shim, and G. J. Pappas, "Safe planning in dynamic environments using conformal prediction," *IEEE Robotics and Automation Letters*, 2023.

[6] N. E. Du Toit and J. W. Burdick, "Robot motion planning in dynamic, uncertain environments," *IEEE Transactions on Robotics*, vol. 28, no. 1, pp. 101–115, 2011.

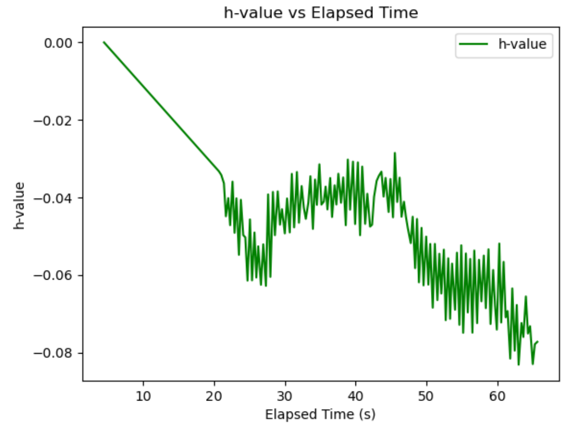


Fig. 5: The value of $h$ function remains consistently $< 1$ as the robot approaches the waypoint, indicating satisfaction of the safety constraints throughout the trajectory

[7] P. Trautman, J. Ma, R. M. Murray, and A. Krause, "Robot navigation in dense human crowds: Statistical models and experimental studies of human–robot cooperation," *The International Journal of Robotics Research*, vol. 34, no. 3, pp. 335–356, 2015.

[8] Y. Chen, F. Zhao, and Y. Lou, "Interactive model predictive control for robot navigation in dense crowds," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 4, pp. 2289–2301, 2021.

[9] S. Liu, P. Chang, Z. Huang, N. Chakraborty, K. Hong, W. Liang, D. L. McPherson, J. Geng, and K. Driggs-Campbell, "Intention aware robot crowd navigation with attention-based interaction graph," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 12 015–12 021.

[10] H. Kretzschmar, M. Spies, C. Sprunk, and W. Burgard, "Socially compliant mobile robot navigation via inverse reinforcement learning," *The International Journal of Robotics Research*, vol. 35, no. 11, pp. 1289–1307, 2016.

[11] A. Turnwald and D. Wollherr, "Human-like motion planning based on game theoretic decision making," *International Journal of Social Robotics*, vol. 11, pp. 151–170, 2019.

[12] H. Kretzschmar, M. Spies, C. Sprunk, and W. Burgard, "Socially compliant mobile robot navigation via inverse reinforcement learning," *The International Journal of Robotics Research*, vol. 35, no. 11, pp. 1289–1307, 2016.

[13] T. Fan, X. Cheng, J. Pan, P. Long, W. Liu, R. Yang, and D. Manocha, "Getting robots unfrozen and unlost in dense pedestrian crowds," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1178–1185, 2019.

[14] A. S. Nemirovskij and D. B. Yudin, *Problem complexity and method efficiency in optimization*. Wiley-Interscience, 1983.

[15] I. Usmanova, "Safe learning via constrained stochastic optimization," Ph.D. dissertation, ETH Zurich, 2022.

[16] I. Usmanova, Y. As, M. Kamgarpour, and A. Krause, "Log barriers for safe black-box optimization with application to safe reinforcement learning," *Journal of Machine Learning Research*, vol. 25, no. 171, pp. 1–54, 2024.

[17] D. L. Marruedo, T. Alamo, and E. Camacho, "Input-to-state stable mpc for constrained discrete-time nonlinear systems with bounded additive uncertainties," *41st IEEE CDC*, pp. 4619 – 4624, 2002.

[18] S. Boyd and L. Vandenberghe, "Convex optimization, cambridge univ," *Press, UK*, 2004.

[19] H. Chen and F. Allgöwer, "A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability," *Automatica*, vol. 34, no. 10, pp. 1205–1217, 1998.

[20] R. Findeisen, L. Imsland, F. Allgower, and B. A. Foss, "State and output feedback nonlinear model predictive control: An overview," *European journal of control*, vol. 9, no. 2-3, pp. 190–206, 2003.

[21] O. Michel, "Webots: Professional mobile robot simulation," *Journal of Advanced Robotics Systems*, vol. 1, no. 1, pp. 39–42, 2004. [Online]. Available: http://www.ars-journal.com/International-Journal-of-Advanced-Robotic-Systems/Volume-1/39-42.pdf