

Autonomous Ground Navigation in Highly Constrained Spaces: Lessons learned from The 2nd BARN Challenge at ICRA 2023

Competition Organizers: Xuesu Xiao¹, Zifan Xu², Garrett Warnell^{2,3}, and Peter Stone^{2,4},
Winning Teams: Ferran Gebelli Guinjoan⁵, Rômulo T. Rodrigues⁶, Herman Bruyninckx^{6,7},
Hanjaya Mandala⁸, Guilherme Christmann⁸,
Jose Luis Blanco-Claraco⁹, and Shravan Somashekara Rai¹⁰

Abstract—The 2nd BARN (Benchmark Autonomous Robot Navigation) Challenge took place at the 2023 IEEE International Conference on Robotics and Automation (ICRA 2023) in London, UK and continued to evaluate the performance of state-of-the-art autonomous ground navigation systems in highly constrained environments. Compared to The 1st BARN Challenge at ICRA 2022 in Philadelphia, the competition has grown significantly in size, doubling the numbers of participants in both the simulation qualifier and physical finals: Ten teams from all over the world participated in the qualifying simulation competition, six of which were invited to compete with each other in three physical obstacle courses at the conference center in London, and three teams won the challenge by navigating a Clearpath Jackal robot from a predefined start to a goal with the shortest amount of time without colliding with any obstacle. The competition results, compared to last year, suggest that the teams are making progress toward more robust and efficient ground navigation systems that work out-of-the-box in many obstacle environments. However, a significant amount of fine-tuning is still needed onsite to cater to different difficult navigation scenarios. Furthermore, challenges still remain for many teams when facing extremely cluttered obstacles and increasing navigation speed. In this article, we discuss the challenge, the approaches used by the three winning teams, and lessons learned to direct future research.

I. THE 2ND BARN CHALLENGE OVERVIEW

The 2nd BARN (Benchmark Autonomous Robot Navigation) Challenge [1] took place as a conference competition at the 2023 IEEE International Conference on Robotics and Automation (ICRA 2023) in London, UK. As a continuation of The 1st BARN Challenge at ICRA 2022 in Philadelphia, the 2nd challenge aimed to evaluate the capability of state-of-the-art navigation systems to move robots through static, highly-constrained obstacle courses, an *ostensibly* simple problem even for many experienced robotics researchers, but in fact, as the results from the 1st competition suggested, a problem far away from being solved [2].

Each team needed to develop an entire navigation software stack for a standardized and provided mobile robot, i.e., a Clearpath Jackal [3] with a 2D 270°-field-of-view Hokuyo LiDAR for perception and a differential drive system with 2m/s maximal speed for actuation. The developed navigation software stack needed to autonomously drive the robot from

a given starting location through a dense obstacle field and to a given goal without any collisions with obstacles or any human interventions. The team whose system could best accomplish this task within the least amount of time would win the competition. The 2nd BARN Challenge had two phases: a qualifying phase evaluated in simulation, and a final phase evaluated in three physical obstacle courses. The qualifying phase took place before the ICRA 2023 conference using the BARN dataset [4] (with the recent addition of DynaBARN [5]), which is composed of 300 obstacle courses in Gazebo simulation randomly generated by cellular automata. The top six teams from the simulation phase were then invited to compete in three different physical obstacle courses set up by the organizers at ICRA 2023 in the ExCeL London conference center.

In this article, we report on the simulation qualifier and physical finals of The 2nd BARN Challenge at ICRA 2023, present the approaches used by the top three teams, discuss lessons learned from the challenge compared against The 1st BARN Challenge at ICRA 2022, and point out future research directions to solve the problem of autonomous ground navigation in highly constrained spaces.

II. SIMULATION QUALIFIER

The simulation qualifier of The 2nd BARN Challenge started on January 1st, 2023. The qualifier used the BARN dataset [4], which consists of 300 5m × 5m obstacle environments randomly generated by cellular automata (see examples in Fig. 1), each with a predefined start and goal. These obstacle environments range from relatively open spaces, where the robot barely needs to turn, to highly dense fields, where the robot needs to squeeze between obstacles with minimal clearance. The BARN environments are open to the public, and were intended to be used by the participating teams to develop their navigation stack. Another 50 unseen environments, which are not available to the public, were generated to evaluate the teams' systems. A random BARN environment generator was also provided to the teams so that they could generate their own unseen test environments.¹

In addition to the 300 BARN environments, six baseline approaches were also provided for the participants' reference,

¹George Mason University ²The University of Texas at Austin ³Army Research Laboratory ⁴Sony AI ⁵Flanders Make ⁶Flanders Make@KU Leuven ⁷Eindhoven University of Technology ⁸Inventec Corporation ⁹University of Almeria ¹⁰Teladoc Health Inc.

¹<https://github.com/dperille/jackal-map-creation>

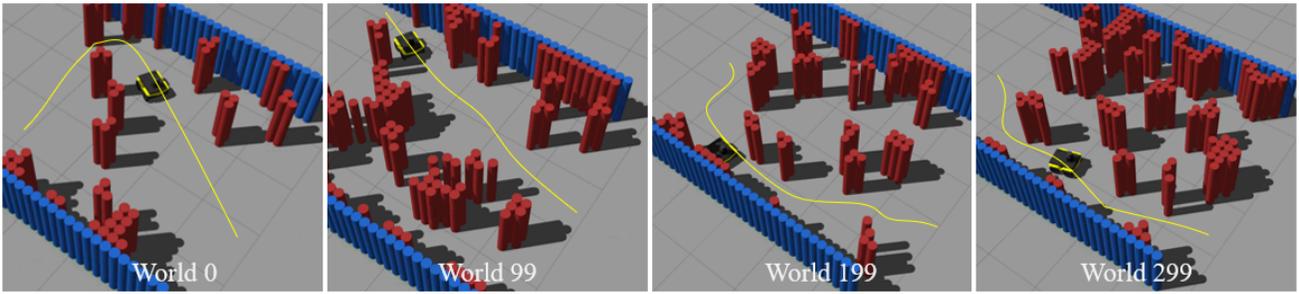


Fig. 1: Four example BARN environments in the Gazebo simulator (ordered by ascending relative difficulty level).

ranging from classical sampling-based [6] and optimization-based navigation systems [7], to end-to-end machine learning methods [8], [9], and hybrid approaches [10]. All baselines were implementations of different local planners used in conjunction with Dijkstra’s search as the global planner in the ROS `move_base` navigation stack [11]. To facilitate participation, a training pipeline capable of running the standardized Jackal robot in the Gazebo simulator with ROS Melodic (in Ubuntu 18.04), with the option of being containerized in Docker or Singularity containers for fast and standardized setup and evaluation, was also provided.²

A. Rules

Each participating team was required to submit their developed navigation system as a (collection of) launchable ROS node(s). The challenge utilized a standardized evaluation pipeline³ to run each team’s navigation system and compute a standardized performance metric that considers navigation success rate (collision or not reaching the goal counts as failure), actual traversal time, and environment difficulty (measured by optimal traversal time). Specially, the score s for navigating each environment i was computed as

$$s_i = 1_i^{\text{success}} \times \frac{OT_i}{\text{clip}(AT_i, 4OT_i, 8OT_i)},$$

where the indicator function 1_{success} evaluates to 1 if the robot reaches the navigation goal without any collisions, and evaluates to 0 otherwise. AT denotes the actual traversal time, while OT denotes the optimal traversal time, as an indicator of the environment difficulty and measured by the shortest traversal time assuming the robot always travels at its maximal speed (2m/s):

$$OT_i = \frac{\text{Path Length}_i}{\text{Maximal Speed}}.$$

The Path Length is provided by the BARN dataset based on Dijkstra’s search from the given start to goal. The clip function clips AT within $4OT$ and $8OT$ in order to assure navigating extremely quickly or slowly in easy or difficult environments respectively won’t disproportionately scale the score. Note that the hyper-parameters 4 and 8 for OT are manually selected before The 1st BARN Challenge,

and the organizers found out that the performance of the submitted navigation systems in The 2nd BARN Challenge has reached the upper bound of this specific metric, 0.25. So the organizers plan to change these hyper-parameters next year to increase the upper bound. The overall score of each team is the score averaged over all 50 unseen test BARN environments, with 10 trials in each environment. Higher scores indicate better navigation performance. The six baselines score between 0.1627 and 0.2334 [1].

B. Results

The simulation qualifier started on January 1st, 2023 and lasted through a soft submission deadline (April 20th, 2023) and a hard submission deadline (May 20th, 2023). Submitting by the soft deadline will guarantee an invitation to the final physical competition given good navigation performance in simulation and leave sufficient time for invited participants to make travel arrangements to London. The hard deadline is to encourage broader participation, but final physical competition eligibility will depend on the available capacity and travel arrangement made beforehand. In total, ten teams from all over the world submitted their navigation systems. The performance of each submission was evaluated by the standard evaluation pipeline. The results are shown in Tab. I with the baselines shown in the forth column as a reference.

TABLE I: Simulation Results

Rank.	Team	Score	Baseline
1	KUL+FM	0.2490	
2	INVENTEC	0.2445	
3	University of Almeria	0.2439	
4	UT AMRL	0.2424	LfLH [9]
5	Temple TRAIL	0.2290	
6	UVA AMR	0.2237	
7	RIL	0.2203	E-Band [7], e2e [8]
8	Staxel	0.2019	APPLR-DWA [10]
9	The MECO Barners	0.1829	(Fast) DWA [6]
10	Lumnonicity	NA	

Compared to the simulation competition in The 1st BARN Challenge at ICRA 2022, in which only one team (Temple TRAIL) outperformed all baselines, four teams (KUL+FM, INVENTEC, University of Almeria, and UT AMRL) achieved better score than the best baseline, Learning from Learned Hallucination (LfLH, 0.2334) [9]. The

²https://github.com/Daffan/ros_jackal

³<https://github.com/Daffan/nav-competition-icra2022>

top six teams, KUL+FM (KU Leuven and Flanders Make), INVENTEC (Inventec Corporation), University of Almeria (University of Almeria and Teladoc Health Inc.), UT AMRL (The University of Texas at Austin), Temple TRAIL (Temple University), UVA AMR (University of Virginia) were invited to the physical finals at ICRA 2023. Note that Temple TRAIL’s simulation score was decreased compared to last year, since the team has learned from last year’s experience and focused on the sim-to-real gap.

III. PHYSICAL FINALS

The physical finals took place at ICRA 2023 in the ExCel London Conference Center on May 30th and May 31st, 2023 (Fig. 2). Two physical Jackal robots with the same sensors and actuators were provided by the competition sponsor, Clearpath Robotics.

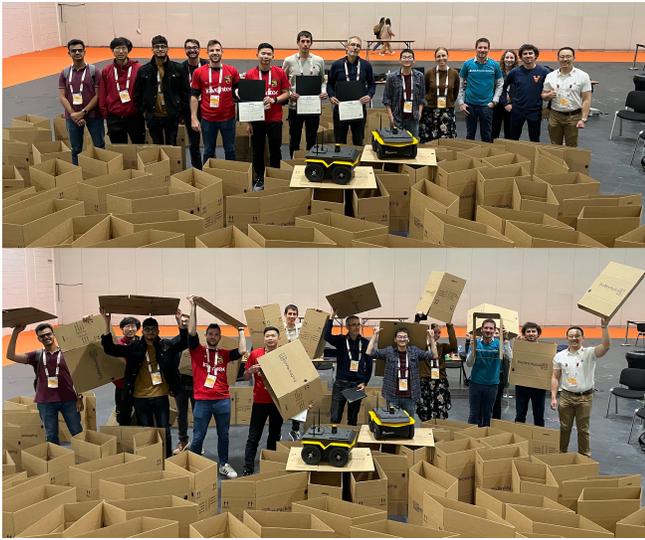


Fig. 2: Final physical competition participants, sponsor Clearpath Robotics, and organizers at The 2nd BARN Challenge in London, UK.

A. Rules

Physical obstacle courses were set up using 90 cardboard boxes in the conference center (Fig. 3). The organizers used the same guidelines to set up three obstacle courses as in The 1st BARN Challenge, i.e., all courses aimed at testing a navigation system’s local planning and therefore had an obvious passage but with minimal clearance (a few centimeters around the robot) when traversing this passage.

The organizers also used the same competition rules agreed by all the physical competition participants: Although it was still impractical to run exactly the same navigation systems submitted by the teams in the simulation qualifier due to poor performance in the real world, the organizers reduced both the set-up time and competition time from last year’s 30 minutes to 20 minutes. After the 20-minute set-up time, each team had the opportunity to run five timed trials (after notifying the organizers to be timed) within another 20-minute. The fastest three out of the five timed trials were

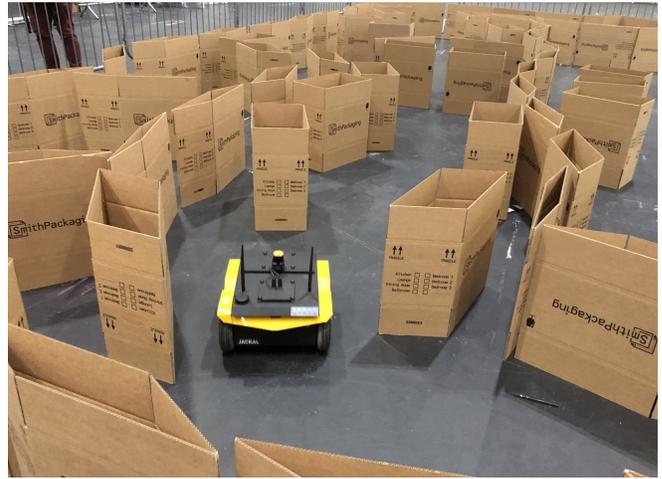


Fig. 3: One (out of three) physical obstacle courses during the finals.

counted, and the team that had the most successful trials (reaching the goal without any collision) was the winner. In the case of a tie, the team with the fastest average traversal time would be declared the winner.

B. Results

The six teams’ navigation performance is shown in Tab. II. Note that due to travel-related reasons, Temple TRAIL cannot attend in-person, so the organizers ran their system submitted to the simulation qualifier for them, which, unfortunately but also as expected, cannot finish one single trial without onsite fine-tuning. Compared to last year, many teams struggled less on the obstacle avoidance problem in the first two easier environments, and therefore were able to shift their attention to increasing speed and were mostly navigating at a much higher speed ($> 1.0\text{m/s}$). The detailed results of all five timed trials (only the top three were counted in the final score) are listed in the last three columns of Tab. II, where “X” indicates failure.

The winner, KUL+FM, is the very first team in The BARN Challenge history that has finished all nine counted physical trials without any collision. In fact, they only failed three trials in all 15 timed trials in the three obstacle courses. The 2nd place winner, INVENTEC, was able to quickly navigate all six counted trials in the first two environments, sometimes even faster than KUL+FM, but didn’t manage to finish the last most constrained obstacle course. University of Almeria also failed all three trials in the last course and one in the first one.

IV. TOP THREE TEAMS AND APPROACHES

In this section, we report the approaches used by the three winning teams.

A. KU Leuven and Flanders Make (KUL+FM)

The core algorithm of the KUL+FM team is the adaptive free-space motion tube [12]. Consider that a robot’s maneuver is defined by a curvature that the robot follows

TABLE II: Physical Results

Rank.	Team	Success/Total	Average Time	Course 1	Course 2	Course 3
1	KUL+FM	9/9	34/56/91	36/X/33/34/34	63/64/52/47/54	86/79/X/79/X
2	INVENTEC	6/9	44/64/NA	47/42/45/45/41	58/66/X/67/X	X/X/X/X/X
3	University of Almeria	5/9	119/79/NA	X/103/134/X/X	85/86/93/X/53	X/X/X/X/X
4	UVA	4/9	68/NA/103	X/X/76/57/71	X/X/X/X/X	103/X/X/X/X
5	UT AMRL	3/9	84/NA/NA	91/X/79/X/81	X/X/X/X/X	X/X/X/X/X
6	Temple TRAIL	0/9	NA/NA/NA	X/X/X/X/X	X/X/X/X/X	X/X/X/X/X

for a time horizon (T) at a constant forward velocity. The motion tube corresponds to the swept volume, that is, the area to be occupied by the vehicle when performing a maneuver. To deal with the discrete resolution of range sensors as well as measurement uncertainty, the footprint of the platform is inflated. Figure 4(a) shows the motion tube for a particular maneuver using the physical (blue polygon) and the inflated (green polygon) footprint of the vehicle. To evaluate whether a motion-tube is within the free-space, the edge of the inflated swept volume is sampled at an interval d_{sample} (Fig. 4(b)). For computational efficiency, samples are projected in the sensor space by associating them to beam indices. A candidate maneuver is said to be available if, for each sample, the corresponding measurement reported by the sensor is greater than the distance from the sensor to the sample. The reasoning for choosing the inflation and sampling interval values are discussed in the original paper [12]. Finally, within the available motion tube, one can decide the control input of the vehicle based on a high-level application goal.

In essence, instead of inflating obstacles in a local map, the adaptive motion tube inflates the robot and its corresponding trajectory. Because motion-tubes are computed with respect to the robot, there is no dependency between localization accuracy and free-space navigation. Another important characteristic of the method is the computational efficiency thanks to the projection of Cartesian samples in the sensor space, which allows computing thousands of motion tubes in low-end computers. A potential disadvantage is that the method relies on the current sensor reading, therefore it is not robust to outliers and measurement errors. Fortunately, both measurement outliers and errors are rather rare and negligible in the proximity of the sensor.

Software stack: The KUL+FM’s software stack for The BARN Challenge uses adaptive free-space motion tube for local navigation, ROS Global Planner for global planning, and Hector SLAM [13] for online mapping and tracking. Based on the most updated map, the global planner provides sub-goal for the local navigation. These sub-goals are used to assign costs to available motion tubes: the closer a maneuver takes the vehicle to a sub-goal, the smaller its cost is. The control input sent to the vehicle corresponds to the weighted average of the available tubes.

Parameter tuning: The parameter of the local navigation were kept constant during the simulation and the physical competition (2000 motion tubes with different curvatures and forward velocities, $T = 1$ s, $d_{sample} = 0.02$ m). Most of the

parameter tuning took place in the global navigation: to avoid obstacles becoming too large and causing the global planner to fail in narrow passages, the team incrementally decreased the obstacle inflation and the distance of the obstacles to be considered in the planner.

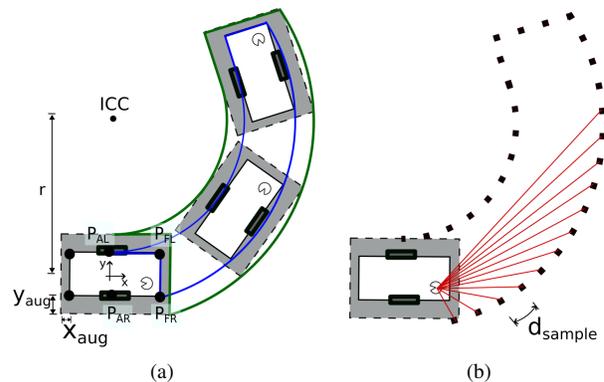


Fig. 4: (KUL+FM) (a) A motion tube corresponds to the inflated trajectory of the vehicle. (b) For computational efficiency, samples are projected in the sensor space.

B. INVENTEC

The INVENTEC team’s approach was to extend the best-performing baseline, Learning from Learned Hallucination (LlLH) [9] with improved collision check and recovery behaviors via a Finite State Machine (FSM). The main driving mode relies on a learning-based model that learns to drive a robot by collecting random trajectories and hallucinating obstacles [9], [14], [15]. However, to address poor generalization of the learned model in out-of-distribution deployment scenarios, the team introduced two alternative modules for front safety checks: Footprint Inflation (FI) used in the simulation qualifier and Model Predictive Control (MPC) used in the physical finals. Additionally, during backward movements, the team performed safety checks by extracting a Region of Interest (RoI) from the obstacle costmap in the robot memory. Details can be found in the comprehensive technical report by the INVENTEC team [16].

Navigation FSM: The FSM consists of five states: *Initial*, *Heading*, *LfLH*, *Forward*, and *Backtrack*, as shown in Fig. 5. In the *Initial* state, the navigation controller waits for the path computed by Dijkstra’s search in the `move_base` global planner with `NavFn` plugin. Then, the state is switched to *Heading*, which aligns the robot to the target path within a

tolerance of $\pm 30^\circ$. Then, the *LfLH* model produces velocity commands taking as input the current LiDAR scans and a local goal drawn from the global path 0.5m ahead of the robot. At every step a safety check is performed. If a future collision is detected, the state changes to *Backtrack* recovery behavior and an alternative slow *Forward* state if the robot is stuck during backtrack.

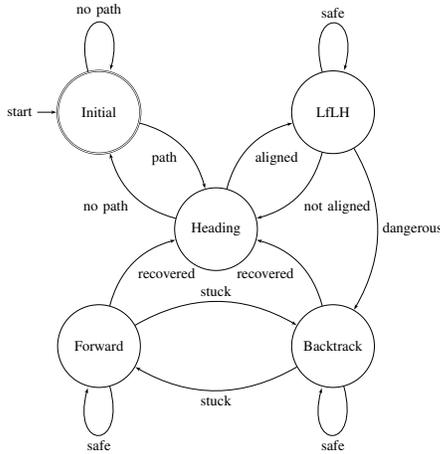


Fig. 5: (INVENTEC) Navigation Finite State Machine.

Recovery Behaviors: During the forward movement in *LfLH* the robot’s path is recorded (green line shown in Fig.6(a)). When *Backtrack* is first triggered, it samples a point 0.3 meters behind the robot along the recorded path, aligns the heading to the target point and performs a straight backward command. Moving backward means moving towards the LiDAR blind spot. Therefore, the team defined a rectangular RoI in the costmap directly behind the robot, illustrated in Fig. 6(b)). At every step, the method checks for possible collisions in the costmap RoI which contains information about past obstacles. If a potential collision is detected during the reverse movement, the state is switched to a slow-forward recovery behavior.

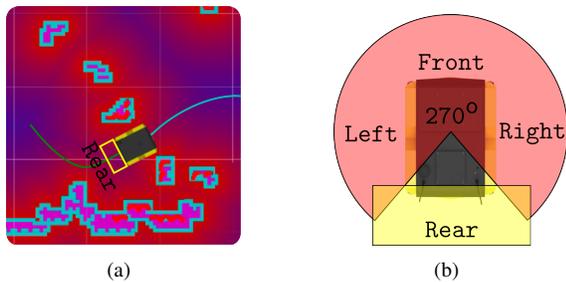


Fig. 6: (INVENTEC) (a) Costmap with robot rear RoI (yellow rectangle). (b) Jackal LiDAR field-of-view.

Simulation Approach: In the simulation stage, INVENTEC team’s strategy for improving the baseline *LfLH* model was threefold. First, utilizing FI for obstacle checking with the latest LiDAR data (shown in Fig. 7); Second, checking the costmap for history obstacles to assess the safety of

the robot’s rear side; Finally, clipping maximum velocity to 0.7m/s, which provided a significant performance improvement. The footprint inflation was 0.04 m. As a result, whenever the inflated footprint overlaps any detected obstacles, the robot transitions into recovery behaviors. Two illustrations of the inflated footprint are depicted in Fig. 7, in which the green region and red region indicate safe and unsafe conditions respectively.

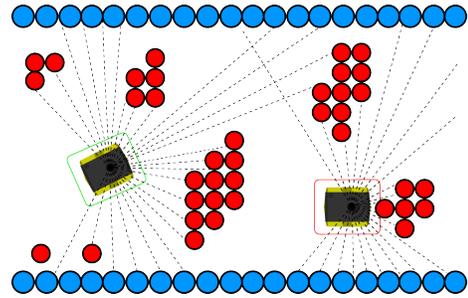


Fig. 7: (INVENTEC) Footprint inflation for collision detection.

During the simulation qualifier, this method enabled the robot to navigate in close proximity to obstacles without collision. However, there exists a trade-off between the size of the inflated footprint and the maximum velocity. If the size is large, the robot will have enough time for braking at high speed but is unable to navigate in highly constrained environments as it frequently stops. Conversely, if the inflated footprint size is small, there is a higher probability of front collision due to insufficient time for braking. The team conducted a test and determined that a 0.7m/s maximum velocity paired with a 0.04m offset provided the optimal balance, allowing the robot to stably navigate without encountering a front collision.

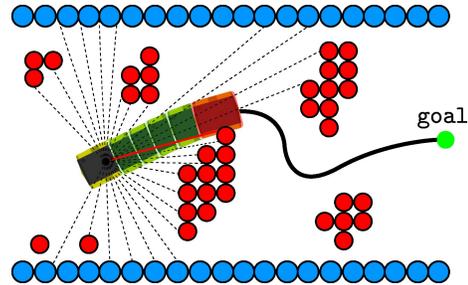


Fig. 8: (INVENTEC) MPC footprint forward safety check based on LiDAR sensor.

Real-world Approach: The only difference between simulation and real-world methods was in the safety check when moving forward. A Model Predictive Control (MPC) approach was adopted for the safety check during forward movement in the real world. MPC takes into account not only the robot’s current footprint but also integrates future position information, enabling the robot to proactively plan its movement and adjust its trajectory accordingly. MPC’s suitability for predictive collision results in adaptability and

real-time responsiveness in real-world navigation.

During the *LfLH* operation, the MPC predicts future steps and detects collisions. At every step, it predicts 20 steps into the future (taking about 200 milliseconds). If at any point in the future trajectory an obstacle overlaps with the footprint of the robot, the current command is deemed unsafe. This stops the *LfLH* state and the FSM switches to recovery behaviors. Fig. 8 depicts this process, with green indicating safe portions of the trajectory and red showing a detected future collision.

C. University of Almeria and Teladoc Health

The University of Almeria team’s implementation was built upon the Mobile Robots Programming Toolkit (MRPT) [17], an open-source C++ framework specifically developed for robotics applications, including libraries for navigation. The MRPT ROS nodes were fine-tuned to align with the specifications of the Clearpath Robotics Jackal robot [3].

A notable difference between the simulation qualifier and physical finals is that in simulation the goal was well-known, with fixed coordinates in a map known a priori. In contrast, in the physical finals there were no absolute coordinates of the goal. In practice, this makes the real-world navigation close to pure *exploration*.

Those caveats aside, the architecture of the system comprises the subsystems enumerated in the following paragraphs, each one implemented as an independent ROS node. Most of the nodes were used unmodified with respect to their open source repositories.⁴

Localization: A custom implementation of particle filter-based localization [18] with adaptive number of particles using the KLD approach [19] has been used by the University of Almeria team. This package is similar to the standard `amcl` ROS package, but capable of using several metric maps at once for localization, with more than one implementation of particle filter algorithms [20].

Local obstacles map: The purpose of this node is to perform real-time acquisition and processing of sensor data, in this case the LiDAR scans. This raw sensor data was subsequently transformed into a 2D point cloud representation, localized within the robot’s own coordinate frame and decimated into lower resolution for faster processing. Phony points were also added in the real-world phase along the past robot poses within a certain time window, in an attempt to discourage the navigation system to go back and revisit parts of the environment that were already traversed. This was revealed to be important when going through large open spaces, where turning back becomes possible.

Path planner (higher layer): At a relatively low rate (1Hz or slower), the local obstacles were considered to find a kinematically feasible path using a path planning algorithm, which is then sent to a local path follower. The team’s path planner uses a custom algorithm based on A^* on a discrete lattice of the state-vector space of the vehicle, i.e., $SE(2)$ pose plus velocities. Arcs between the

lattice nodes were explored efficiently using Parameterized Trajectory Generators (PTGs), a concept derived from past works [21], which defines families of paths to help exploring the environment with kinematically and dynamically feasible paths. Disregarding the use of these path libraries, the team’s method has resemblances with the implementation of the dubbed *Smac Planner* in Nav2 [22], although a systematic comparison with the state-of-the-art and a proper paper reporting the method’s details still remain as future work. The whole planner can be seen as an improved version of a formerly published RRT planner [23]. This package is open-sourced and released to ROS.⁵ This path planner is one of the main strengths of the whole stack, especially for simulation, since complex maneuvers in tight spaces can be computed in a safe manner.

Local planner (lower layer): Once a path is found reaching (or, at least, approaching) the target, the task of generating motor commands to follow it, including avoiding any new obstacles, is accomplished by a reactive navigation system, which is also based on Trajectory Parameter Space (TP-Space) [21]. In TP-Space, the robot, regardless of its physical shape and kinematic constraints, is transformed into a free-flying point within a newly formulated parameter space. This transformation incorporates the robot’s shape and kinematic restrictions, thereby allowing for efficient navigation by taking into account the robot’s specific physical characteristics and movement capabilities. On the other hand, PTGs define a set of potential trajectories for the robot, parameterized by variables such as path shape, robot speed, and turning radius. By dynamically adjusting these parameters based on real-time sensor data, the robot can select the optimal path for avoiding obstacles and reaching its destination. The reactive navigation system, through its integration with TP-Spaces and PTGs, acts as an advanced behavior planning algorithm. The implementation used the ROS node developed based on these principles. It processed the 2D point cloud data to dynamically generate an optimal path for the robot, while simultaneously accounting for the robot’s kinematic constraints and potential obstacles within its environment. This local planner can run at a higher rate than the path planner, typically between 5 and 10Hz.

Exploration: This module is in charge of generating targets for the path planner to find empty passages around. Since The BARN Challenge typically requires the robot to “move forward”, The University of Almeria team’s implementation simply generated target goals in a region of interest a few meters in front of the robot, but with a certain random perturbation to avoid getting the global planner trapped in a local minima.

In summary, the system’s ability to process and respond to real-time sensor data in a rapid and efficient manner, while accounting for the robot’s physical and kinematic characteristics, makes it a robust solution for navigating complex and constrained environments.

⁴<https://github.com/mrpt-ros-pkg/>

⁵https://github.com/MRPT/mrpt_path_planning/

V. DISCUSSIONS

While many discussion points from The 1st BARN Challenge [2] are still valid this year, we discuss new findings and lessons from The 2nd BARN Challenge and point out promising future research directions to push the boundaries of efficient mobile robot navigation in highly constrained spaces.

A. More participants from industry (hybrid academia-industry teams)

One interesting change in the 2nd year of The BARN Challenge is the participation from robotics industry. Compared to last year's competition, in which all participants were from universities worldwide, five out of the ten teams who participated in The 2nd BARN Challenge were from industry or hybrid academia-industry teams (KU Leuven with Flanders Make, INVENTEC, University of Almeria with Teladoc Health Inc., Staxel, and Lumnonicity from JIO AICOE). Notably all three winning teams have industry ties.

The significant increase in industry participation indicates that the problem The BARN Challenge aims to solve is of significant interest to real-world robotics manufacturers, providers, and users. As discussed in last year's report [2], even very experienced robotics researchers in academia may have the impression that such an ostensibly simple problem has already been solved. However, the fact that robotics industry is still working on such a problem and using The BARN Challenge as a testbed for their methods suggests that they still do not have a satisfactory solution for this problem in the real world. Therefore, the organizers suggest academic researchers to consider real-world problems the robotics industry faces in order to realize technology transfer from academia to industry. The 1st and 3rd place winners, KUL+FM and University of Almeria (with Teladoc) are very good examples of applying state-of-the-art academic research to real-world robotics industry and identifying real-world problems from industry to solve together with academia.

B. Simulation performance approaching the current upper bound

Another observation from the simulation qualifier is that the simulation performance has approached the current metric's upper bound, i.e., 0.25. The upper and lower bound is determined by the two hyper-parameters 4 and 8 in the evaluation metric to assure navigating extremely quickly or slowly in easy or difficult environments respectively won't disproportionately scale the score. Approaching 0.25 means most of the navigation trials in the BARN environments can be successfully finished within four times of the optimal traversal time, even in the most difficult ones with very dense obstacles. Therefore, the organizers will change these hyper-parameters in next year's challenge to increase the upper and lower bound to encourage the teams to score higher by achieving less traversal time than four times of the optimal time, without colliding with any obstacle.

C. The first team that finished all physical courses

KUL+FM is the first team in The BARN Challenge history that was able to successfully navigate the Jackal through all three physical obstacle courses in all nine final counted trials. However, they did fail three out of the total 15 timed trials in obstacle courses 1 and 3 as shown in Tab. II. While the failure trial in obstacle course 1 was likely due to the need of initial parameter fine-tuning to fit the physical environments, the two failure trials in the last obstacle course were both due to the desire to push on faster navigation speed. Considering that the physical obstacle courses constructed this year were qualitatively more difficult than last year's, all three winning teams' stable performance in the first two obstacle courses and KUL+FM's all nine successful trials in all three courses suggest that the physical performance of the navigation stacks have been improved compared to last year.

D. Final ranking still decided by success rate, not speed

Despite the improved performance and more confident deployment experience (e.g., less struggling to navigate to the goal and to avoid obstacles, but more focus on fine-tuning for robustness and speed) observed by the organizers, completely collision-free navigation is still out of reach for most teams regardless of speed. Therefore the final ranking in the physical finals is still decided by success rate, not navigation speed. Although the 2nd place winner, INVENTEC, had a chance to win the competition before they started the 3rd obstacle course by pushing on reducing traversal time, they eventually failed all five attempts in the last course. So the community is still waiting for the first BARN Challenge, in which the final ranking is determined by traversal speed in highly constrained obstacle environments, after the success rate has been guaranteed to be 100%.

E. Less obvious sim-to-real gap

Compared to last year when the winning approach in simulation suffered from significant collisions in the real world [2], the rankings in Tab. I and II suggest an decreasing sim-to-real gap between the simulation qualifier and physical finals this year. Based on the approaches taken by the teams, it is possible that such a smaller sim-to-real gap is caused by the decreased usage on learning-based navigation methods [24]. Despite the popularity of using machine learning to address visual inputs [25]–[27], off-road conditions [25], [28]–[33], social contexts [34]–[38], and multi-robot navigation [39], [40], only one team, INVENTEC (except Temple TRAIL who did not compete in the physical finals), used a learning-based method to navigate in highly constrained obstacle environments in The 2nd BARN Challenge. Their approach is mostly based on the Learning from Hallucination [9], [14], [15], [41] paradigm, especially the latest LfLH [9] approach (one of the baselines provided by the competition organizers). Furthermore, INVENTEC designed sophisticated recovery behaviors to address real-world scenarios where the learning approach did not work well (see detailed discussions below), which also helped to reduce the sim-to-real gap.

F. Importance of good recovery behaviors to deploy end-to-end learning-based systems in the real world

During The 1st BARN Challenge last year, the end-to-end learning approach trained by Deep Reinforcement Learning by Temple TRAIL [42] experienced significant sim-to-real gap because the training was conducted in simulation on a different, smaller robot platform (Turtlebot2) and there was no recovery behavior. This year, Temple TRAIL was not able to compete in-person and the system deployed by the organizers on behalf of the team did not perform well without fine-tuning. As the other team that adopted machine learning for their navigation system, INVENTEC used the LfLH [9] policy trained using simulated data, but also devised a set of recovery behaviors to address real-world scenarios where the learning approach did not work well. By developing good recovery behaviors to complement an end-to-end learned motion policy, INVENTEC's approach outperformed the original LfLH approach, which is assisted only by a set of very simple recovery behaviors, and achieved very good sim-to-real transfer. This observation suggests the potential of end-to-end learning approaches for navigation when being augmented by a sophisticated mechanism to complement learning during out-of-distribution real-world scenarios [43], [44].

G. Tuning still necessary for all classical systems facing different obstacle environments

Similar to last year's observation, the original intention of "out-of-the-box" deployment of the navigation systems submitted to the simulation qualifier directly in the physical finals was still impossible for all the teams: all teams that used classical systems had to extensively fine-tune their navigation stack, while INVENTEC, the only team that used a learning-based approach, fine-tuned their recovery behaviors, the classical part of their system. Although slight fine-tuning to adapt a system from simulation to the real world is reasonable, all teams needed to fine-tune their navigation systems before competing in all three obstacle courses. It is unclear whether a single navigation system configuration that works for all obstacle courses exists or not. The organizers suggest the community to keep such an intention in mind when developing their navigation systems, because the option of fine-tuning for every deployment scenario is not possible, especially when the goal is to deploy autonomous mobile robots at scale in the wild. One single navigation system configuration or parameterization to address all possible scenarios is certainly ideal, but another promising approach is autonomous parameter tuning to adapt to different obstacle configurations around the robot [10], [45]–[49].

H. Interest from outside the robotics community

Surprisingly, The 2nd BARN Challenge has even raised interest from outside the robotics community. For example, Harald Carlens, a Machine Learning researcher who runs <https://mlcontests.com>, which lists machine learning competitions, has reported The 1st BARN Challenge after ICRA 2022 in "The State of Competitive Machine

Learning" [50]. Harald Carlens was able to attend and observe The 2nd BARN Challenge in-person in London and had long discussions with the organizers regarding The BARN Challenge and other Machine Learning competitions. Another Artificial Intelligence student researcher from RWTH Aachen University, Malte Schwerin, reached out after The 2nd BARN Challenge and requested raw results of the simulation qualifier to further study ranking methods for algorithm competitions. A professor in robotics and psychology from Graz University of Technology in Austria, Dr. Gerald Steinbauer-Wagner, reached out during the physical finals and was interested in assessing robotic capabilities using methods from psychology, i.e., test theory, which had been applied for centuries to assess the latent capabilities of people, e.g., intelligent and PISA test. This well-grounded theory allows to create objective, reliable, and valid tests for skills, even for nondeterministic domains. Dr. Steinbauer-Wagner's research group is currently working on analyzing the BARN simulation runs via Item Response Theory [51] and One-Parameter Rasch Model [52]. Such a trend indicates that benchmarking robotic navigation capabilities is not only of interest for the sake of improving autonomous navigation for roboticists alone, but also has broader impacts on the study of machine learning, artificial intelligence, test theory, psychology, and potentially other disciplines.

VI. CONCLUSIONS

The results of The 2nd BARN Challenge at ICRA 2023 suggest that the mobile robot community has been making steady progress on autonomous metric ground navigation, an ostensibly simple but largely unsolved problem. The competition has doubled its size, attracted many participants from robotics industry, approached the performance upper bound in the simulation qualifier, had the first winning team that finished all physical trials, experienced a smaller sim-to-real gap between the simulation qualifier and physical finals, and attracted interest from outside the robotics community. However, work remains to be done in order to guarantee collision avoidance and to reduce the need of system tuning to efficiently navigate in different obstacle environments.

REFERENCES

- [1] "The barn challenge 2023," https://cs.gmu.edu/~xiao/Research/BARN_Challenge/BARN_Challenge23.html, accessed: 2023-07-23.
- [2] X. Xiao, Z. Xu, Z. Wang, Y. Song, G. Warnell, P. Stone, T. Zhang, S. Ravi, G. Wang, H. Karnan *et al.*, "Autonomous ground navigation in highly constrained spaces: Lessons learned from the benchmark autonomous robot navigation challenge at icra 2022 [competitions]," *IEEE Robotics & Automation Magazine*, vol. 29, no. 4, pp. 148–156, 2022.
- [3] "Jackal ugv - small weatherproof robot - clearpath," <https://clearpathrobotics.com/jackal-small-unmanned-ground-vehicle/>, accessed: 2022-07-21.
- [4] D. Perille, A. Truong, X. Xiao, and P. Stone, "Benchmarking metric ground navigation," in *2020 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. IEEE, 2020, pp. 116–121.
- [5] A. Nair, F. Jiang, K. Hou, Z. Xu, S. Li, X. Xiao, and P. Stone, "Dynabarn: Benchmarking metric ground navigation in dynamic environments," in *2022 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. IEEE, 2022, pp. 347–352.

- [6] D. Fox, W. Burgard, and S. Thrun, "The dynamic window approach to collision avoidance," *IEEE Robotics & Automation Magazine*, vol. 4, no. 1, pp. 23–33, 1997.
- [7] S. Quinlan and O. Khatib, "Elastic bands: Connecting path planning and control," in *[1993] Proceedings IEEE International Conference on Robotics and Automation*. IEEE, 1993, pp. 802–807.
- [8] Z. Xu, B. Liu, X. Xiao, A. Nair, and P. Stone, "Benchmarking reinforcement learning techniques for autonomous navigation," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 9224–9230.
- [9] Z. Wang, X. Xiao, A. J. Nettekoven, K. Umasankar, A. Singh, S. Bommakanti, U. Topcu, and P. Stone, "From agile ground to aerial navigation: Learning from learned hallucination," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 148–153.
- [10] Z. Xu, G. Dhamankar, A. Nair, X. Xiao, G. Warnell, B. Liu, Z. Wang, and P. Stone, "Applr: Adaptive planner parameter learning from reinforcement," in *2021 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2021, pp. 6086–6092.
- [11] "ROS movebase navigation stack," http://wiki.ros.org/move_base, accessed: 2021-09-9.
- [12] R. T. Rodrigues, N. Tsiogkas, N. Huebel, and H. Bruyninckx, "Clutter-resilient autonomous mobile robot navigation with computationally efficient free-space features," in *Robotics Research*, A. Billard, T. Asfour, and O. Khatib, Eds. Cham: Springer Nature Switzerland, 2023, pp. 522–537.
- [13] S. Kohlbrecher, J. Meyer, O. von Stryk, and U. Klingauf, "A flexible and scalable slam system with full 3d motion estimation," in *Proc. IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR)*. IEEE, November 2011.
- [14] X. Xiao, B. Liu, G. Warnell, and P. Stone, "Toward agile maneuvers in highly constrained spaces: Learning from hallucination," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1503–1510, 2021.
- [15] X. Xiao, B. Liu, and P. Stone, "Agile robot navigation through hallucinated learning and sober deployment," in *2021 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2021, pp. 7316–7322.
- [16] H. Mandala and G. Christmann, "The barn challenge 2023 – autonomous navigation in highly constrained spaces – inventec team," *arXiv preprint arXiv:2307.14580*, 2023.
- [17] "Mobile robot programming toolkit (MRPT)," <https://github.com/MRPT/mrpt>, accessed: 2023-07-28.
- [18] S. Thrun, D. Fox, W. Burgard, and F. Dellaert, "Robust monte carlo localization for mobile robots," *Artificial intelligence*, vol. 128, no. 1-2, pp. 99–141, 2001.
- [19] D. Fox, "Kld-sampling: Adaptive particle filters," *Advances in neural information processing systems*, vol. 14, 2001.
- [20] J.-L. Blanco, J. González, and J.-A. Fernández-Madriral, "Optimal filtering for non-parametric observation models: applications to localization and slam," *The International Journal of Robotics Research*, vol. 29, no. 14, pp. 1726–1742, 2010.
- [21] J.-L. Blanco, J. González, and J.-A. Fernández-Madriral, "Extending obstacle avoidance methods through multiple parameter-space transformations," *Autonomous Robots*, vol. 24, no. 1, pp. 29–48, 2008. [Online]. Available: <https://doi.org/10.1007/s10514-007-9062-7>
- [22] S. Macenski, F. Martín, R. White, and J. Ginés Clavero, "The marathon 2: A navigation system," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020. [Online]. Available: <https://github.com/ros-planning/navigation2>
- [23] J. L. Blanco, M. Bellone, and A. Gimenez-Fernandez, "Tp-space rrt-kinematic path planning of non-holonomic any-shape vehicles," *International Journal of Advanced Robotic Systems*, vol. 12, no. 5, p. 55, 2015.
- [24] X. Xiao, B. Liu, G. Warnell, and P. Stone, "Motion planning and control for mobile robot navigation using machine learning: a survey," *Autonomous Robots*, pp. 1–29, 2022.
- [25] G. Kahn, P. Abbeel, and S. Levine, "Badgr: An autonomous self-supervised learning-based navigation system," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1312–1319, 2021.
- [26] H. Karnan, G. Warnell, X. Xiao, and P. Stone, "Voila: Visual-observation-only imitation learning for autonomous navigation," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 2497–2503.
- [27] K. S. Sikand, S. Rabiee, A. Uccello, X. Xiao, G. Warnell, and J. Biswas, "Visual representation learning for preference-aware path planning," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 11 303–11 309.
- [28] Y. Pan, C.-A. Cheng, K. Saigol, K. Lee, X. Yan, E. A. Theodorou, and B. Boots, "Imitation learning for agile autonomous driving," *The International Journal of Robotics Research*, vol. 39, no. 2-3, pp. 286–302, 2020.
- [29] X. Xiao, J. Biswas, and P. Stone, "Learning inverse kinodynamics for accurate high-speed off-road navigation on unstructured terrain," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 6054–6060, 2021.
- [30] H. Karnan, K. S. Sikand, P. Atreya, S. Rabiee, X. Xiao, G. Warnell, P. Stone, and J. Biswas, "Vi-ikd: High-speed accurate off-road navigation using learned visual-inertial inverse kinodynamics," in *To Appear in 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022.
- [31] P. Atreya, H. Karnan, K. S. Sikand, X. Xiao, G. Warnell, S. Rabiee, P. Stone, and J. Biswas, "High-speed accurate robot control using learned forward kinodynamics and non-linear least squares optimization," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022.
- [32] A. Datar, C. Pan, M. Nazeri, and X. Xiao, "Toward wheeled mobility on vertically challenging terrain: Platforms, datasets, and algorithms," *arXiv preprint arXiv:2303.00998*, 2023.
- [33] A. Datar, C. Pan, and X. Xiao, "Learning to model and plan for wheeled mobility on vertically challenging terrain," *arXiv preprint arXiv:2306.11611*, 2023.
- [34] R. Mirsky, X. Xiao, J. Hart, and P. Stone, "Prevention and resolution of conflicts in social navigation—a survey," *arXiv preprint arXiv:2106.12113*, 2021.
- [35] C. Mavrogiannis, F. Baldini, A. Wang, D. Zhao, P. Trautman, A. Steinfeld, and J. Oh, "Core challenges of social robot navigation: A survey," *arXiv preprint arXiv:2103.05668*, 2021.
- [36] H. Karnan, A. Nair, X. Xiao, G. Warnell, S. Pirk, A. Toshev, J. Hart, J. Biswas, and P. Stone, "Socially compliant navigation dataset (scand): A large-scale dataset of demonstrations for social navigation," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 11 807–11 814, 2022.
- [37] X. Xiao, T. Zhang, K. M. Choromanski, T.-W. E. Lee, A. Francis, J. Varley, S. Tu, S. Singh, P. Xu, F. Xia, S. M. Persson, L. Takayama, R. Frostig, J. Tan, C. Parada, and V. Sindhwani, "Learning model predictive controllers with real-time attention for real-world navigation," in *Conference on robot learning*. PMLR, 2022.
- [38] D. M. Nguyen, M. Nazeri, A. Payandeh, A. Datar, and X. Xiao, "Toward human-like social robot navigation: A large-scale, multimodal, social human navigation dataset," in *2023 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2023.
- [39] P. Long, T. Fan, X. Liao, W. Liu, H. Zhang, and J. Pan, "Towards optimally decentralized multi-robot collision avoidance via deep reinforcement learning," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 6252–6259.
- [40] Y. F. Chen, M. Liu, M. Everett, and J. P. How, "Decentralized non-communicating multiagent collision avoidance with deep reinforcement learning," in *2017 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2017, pp. 285–292.
- [41] J.-S. Park, X. Xiao, G. Warnell, H. Yedidsion, and P. Stone, "Learning perceptual hallucination for multi-robot navigation in narrow hallways," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 10 033–10 039.
- [42] Z. Xie and P. Dames, "Drl-vo: Learning to navigate through crowded dynamic scenes using velocity obstacles," *IEEE Transactions on Robotics*, 2023.
- [43] B. Liu, X. Xiao, and P. Stone, "A lifelong learning approach to mobile robot navigation," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1090–1096, 2021.
- [44] Z. Xu, A. Nair, X. Xiao, and P. Stone, "Learning real-world autonomous navigation by self-supervised environment synthesis," *arXiv preprint arXiv:2210.04852*, 2022.
- [45] X. Xiao, B. Liu, G. Warnell, J. Fink, and P. Stone, "Appld: Adaptive planner parameter learning from demonstration," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4541–4547, 2020.
- [46] Z. Wang, X. Xiao, B. Liu, G. Warnell, and P. Stone, "Appli: Adaptive planner parameter learning from interventions," in *2021 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2021, pp. 6079–6085.
- [47] Z. Wang, X. Xiao, G. Warnell, and P. Stone, "Apple: Adaptive planner

- parameter learning from evaluative feedback,” *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7744–7749, 2021.
- [48] H. Ma, J. S. Smith, and P. A. Vela, “Navtuner: Learning a scene-sensitive family of navigation policies,” in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 492–499.
- [49] X. Xiao, Z. Wang, Z. Xu, B. Liu, G. Warnell, G. Dhamankar, A. Nair, and P. Stone, “Appl: Adaptive planner parameter learning,” *Robotics and Autonomous Systems*, vol. 154, p. 104132, 2022.
- [50] “The state of competitive machine learning - ml contests,” <https://mlcontests.com/state-of-competitive-machine-learning-2022/>, accessed: 2022-07-23.
- [51] C. E. Cantrell, “Item response theory: Understanding the one-parameter rasch model.” 1997.
- [52] R. Hambleton, H. Swaminathan, and H. Rogers, “Principles and applications of item response theory,” 1985.