

SanD-Planner: Sample-Efficient Diffusion Planner in B-Spline Space for Robust Local Navigation

Jincheng Wang¹, Lingfan Bao¹, Tong Yang², Diego Martinez Plasencia¹, Jianhao Jiao^{1,3,*}, and Dimitrios Kanoulas¹



Fig. 1: Real-world demonstrations of SanD-Planner. The proposed algorithm enables the robot to traverse cluttered environments with tight clearance, including navigating narrow passages, avoiding obstacles, and climbing stairs.

Abstract—The challenge of generating reliable local plans has long hindered practical applications in highly cluttered and dynamic environments. Key fundamental bottlenecks include acquiring large-scale expert demonstrations across diverse scenes and improving learning efficiency with limited data. This paper proposes SanD-Planner, a sample-efficient diffusion-based local planner that conducts depth image-based imitation learning within the clamped B-spline space. By operating within this compact space, the proposed algorithm inherently yields smooth outputs with bounded prediction errors over local supports, naturally aligning with receding-horizon execution. Integration of an Euclidean Signed Distance Field (ESDF)-based safety checker with explicit clearance and time-to-completion metrics further reduces the training burden associated with value-function learning for feasibility assessment. Experiments show that training with 500 episodes (merely 0.25% of the demonstration scale used by the baseline), SanD-Planner achieves state-of-the-art performance on the evaluated open benchmark, attaining success rates of 90.1% in simulated cluttered environments and 72.0% in indoor simulations. The performance is further proven by demonstrating zero-shot transferability to realistic experimentation in both 2D and 3D scenes. The pre-trained models are open-sourced, and further details are available on the project website: <https://wangjincheng1998.github.io/SanD-planner>.

¹Department of Computer Science, University College London, London, the UK. (jincheng.wang.22@ucl.ac.uk, lingfan.bao.24@ucl.ac.uk, ucacaji@ucl.ac.uk, d.plasencia@ucl.ac.uk, d.kanoulas@ucl.ac.uk)

²Department of Computer Science and Technology, Tsinghua University, P. R. China. (tong.yang@zju.edu.cn)

³Department of Aeronautical and Aviation Engineering, the Hong Kong Polytechnic University, Hong Kong, P. R. China.

*Corresponding Author

I. INTRODUCTION

The capability of robust navigating in cluttered and dynamic environments remains a fundamental challenge for general-purpose mobile robotics [1, 2]. Classic navigation frameworks typically utilize a modular pipeline that sequentially executes perception, SLAM, and planning sub-routines [3–7]. Albeit effective in structured settings, the dependency on precise mapping creates bottlenecks in dynamic environments. Furthermore, the sequential processing induces latency and compounded error from sensor noises and drifts, compromising the system’s ability to safely adapt to abrupt changes. This has catalyzed a shift toward end-to-end navigation approaches [8–11] that map onboard observations directly to actions or trajectories. While deep reinforcement learning (DRL) optimizes policies through environment interaction [12–15], it is often hindered by sample inefficiency and the difficulty of reward engineering [16]. In contrast, imitation learning (IL) bypasses these issues by learning directly from expert demonstrations. Recent advances have shown remarkable results by aggregating large-scale real-world datasets to achieve general navigation behaviors [17–19].

However, the success of these methods remains largely data-driven, relying heavily on demonstration coverage and diversity [9]. Performance gains often hinge on expanding the supervision pipeline rather than learning effectively from limited demonstrations, becoming more challenging in scenarios where expert data are costly and unsafe to collect [20].

In this paper, we explore a complementary direction by asking whether local navigation can achieve high performance in a *data desert* of limited expert demonstrations. We propose **SanD-Planner**, a **Sample-Efficient Diffusion-based Local Planner** that maps depth observations directly to smooth and parametric paths. Unlike waypoint-based methods [18] that scale the output dimension linearly with the planning horizon, SanD-Planner predicts a fixed set of eight B-spline control points. This representation embeds structural geometric priors as an inductive bias, enabling an extended horizon to mitigate myopic behavior [21] without increasing learning complexity. By construction, B-splines guarantee C^2 continuity as a structural property rather than a learnt feature, while their local support isolates far-horizon perceptual uncertainty (*e.g.*, depth noise or occlusions) to distal control points, preventing it from destabilizing the immediate execution segment. Finally, we decouple trajectory generation from safety assessment via an interpretable online critic. By relegating feasibility checks to this explicit geometric module, we reduce learning complexity and allow the policy to focus exclusively on distilling the expert trajectory distribution. The contributions can be summarized as follows:

- **SanD-Planner**: A sample-efficient diffusion-based local planner that generates smooth and collision-avoiding paths within a clamped cubic B-spline control-point space. As such, it achieves high-performance point-goal navigation using only 500 expert trajectories, which is approximately 0.25% of the demonstration scale required by current state-of-the-art (SoTA) algorithms.
- **Representation Study**: A systematic investigation conducted under a unified training and evaluation protocol, characterizing how trajectory representations, precisely waypoint discretization, cubic interpolation, and B-spline control point design, impact the performance and sample efficiency of imitation learning-based local planners.
- **Zero-Shot Sim-to-Real Validation**: Successful deployment on a Unitree Go2 quadruped robot across diverse real-world environments, as shown in Fig. 1. This demonstrates robust, depth-only 3D navigation in unseen complex scenes without any real-world fine-tuning.
- **Reproducibility**: We will release our complete training and evaluation framework, including the dataset and pre-trained models.

II. RELATED WORK

A. Learning-based End-to-End Visual Navigation

DRL is a prominent category in this domain, employing trial-and-error optimization to eliminate the need for labeled data [12–14, 22]. However, DRL is often sample-inefficient [23] and relies on complex reward engineering that often proves more difficult than providing expert demonstrations [24]. IL has emerged as a practical paradigm, enabling policies to learn directly from expert behaviors [8, 17, 25–29]. Building on this foundation, recent approaches have explored generative policies, particularly diffusion-based ones, which model local

paths as conditional distributions, thereby capture the inherent multi-modality of navigation [11, 19, 30–33]. However, their robustness is fundamentally conditioned by the diversity and long-tail coverage of training data [10]. Consequently, pushing the limits of end-to-end navigation has largely depended on scaling expert demonstrations [34]. This is typically achieved by aggregating large real-world datasets [8, 17] or by generating high-quality simulation supervision at scale, rather than learning effectively from limited data. Recent works such as NavDP [19] leverage high-fidelity simulators to generate extensive supervision, utilizing datasets spanning hundreds of kilometers and requiring substantial compute budgets.

In contrast to the trend of data scaling, this work demonstrates that a robust local planning policy can be learned effectively with limited data by injecting structural priors. We leverage B-spline parameterization to explicitly enforce path smoothness and inherent representation-level robustness. By combining this generative policy with an explicit geometric critic for trajectory selection, we decouple safety verification from policy learning. This design enables robust navigation with significantly fewer demonstrations.

B. B-spline Representation in Navigation

B-splines [35–37] are mathematical representations that encode complex geometric curves compactly into control-point space. This motivates us to perform optimization within the control-point space for the efficiency purpose while naturally respecting smoothness and feasibility constraints. Classical optimization-based planners exploit spline structures, such as the convex-hull property, to improve clearance and dynamic feasibility [38–40]. These results suggest that trajectory parameterization can serve as a powerful inductive bias for learning-based navigation. In previous works, many IL policies were supervised using single-step actions [26, 27, 41], making temporal consistency an explicit learning burden [42]. To mitigate this, recent diffusion-based planners [43] predict short-horizon waypoint rollouts or action sequences and execute them in a receding-horizon manner, encouraging coherent motion [18, 31]. Moving beyond discrete sequences, recent approaches predict sparse waypoints and convert them into cubic-spline paths [9, 10]. In contrast, this work learns local planning directly in a clamped B-spline control-point space, thereby inheriting the continuity and convex-hull properties. Our analysis in Section IV-B will demonstrate that B-splines offer distinct advantages for imitation-based planners compared to waypoint or standard cubic-spline parameterizations.

III. PROBLEM FORMULATION

This work addresses mapless local navigation in cluttered environments by learning a collision-free goal-reaching policy. At each planning step t , the robot receives a short history of four depth images $\mathcal{O}_t = \{\mathbf{D}_{t-3}, \mathbf{D}_{t-2}, \mathbf{D}_{t-1}, \mathbf{D}_t\}$, captured in the robot frame, and receives a three-dimensional relative goal $\mathbf{g}_t = [x, y, z]^\top$ expressed in the same frame. To ensure temporal consistency across consecutive replanning cycles, we also incorporate a motion context term $\mathbf{v}_t^{\text{prev}}$ that

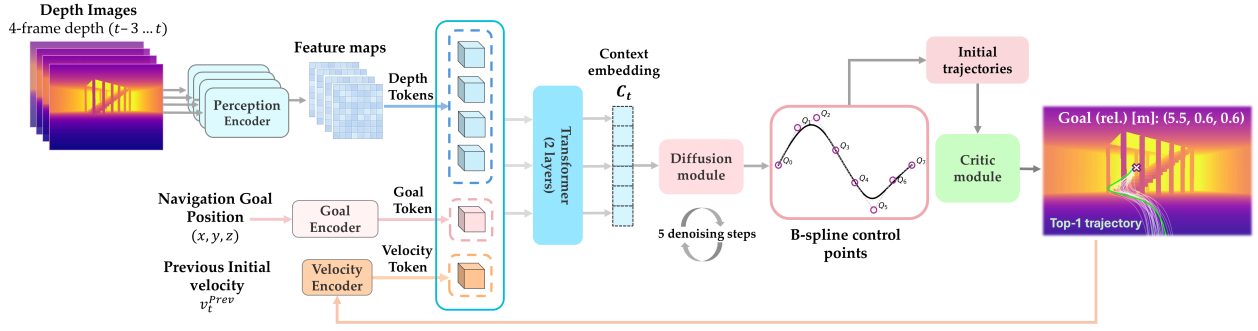


Fig. 2: **Overview of SanD-Planner.** The pipeline tokenizes and fuses historical depth images, relative point goals, and previous velocity using a two-layer Transformer encoder. The multi-modal context conditions a diffusion policy to generate B-spline control points through iterative denoising. Finally, a geometric critic module selects the optimal plan from candidates for execution and feeds its initial velocity back to the next planning step to maintain temporal consistency.

encodes the heading direction of the plan executed in the previous step. These inputs are projected via domain-specific encoders $f_{\cdot}$ and fused into a unified latent context $\mathcal{C}_t = \Phi(f_{\mathcal{O}}(\mathcal{O}_t), f_g(\mathbf{g}_t), f_v(\mathbf{v}_t^{\text{prev}}))$ using an encoder Φ . Instead of predicting waypoint sequences, our planner operates in the structured space of B-spline control points. The trajectory τ_t is continuously expressed in the robot frame, which consists of 3D waypoints (x, y, z) . In Eq. (1), we parameterize each output local trajectory τ_t as a sequence of N control points $\mathcal{Q}_t = \{\mathbf{Q}_{t,i}\}_{i=0}^{N-1} \in \mathbb{R}^{N \times 3}$ in the robot frame.

The objective is to generate a smooth, collision-free trajectory τ_t that efficiently connects the current pose to $\mathbf{g}_t$. To this end, we learn a conditional diffusion model $p_{\theta}(\mathcal{Q}_t | \mathcal{C}_t)$ from expert demonstrations that captures the distribution of feasible control points. During planning, the diffusion policy samples K candidate control-point sets $\{\mathcal{Q}_t^{(k)}\}_{k=1}^K$, which generate continuous trajectories $\{\tau_t^{(k)}\}_{k=1}^K$. The optimal plan τ_t^* is then selected from these candidates by minimizing a task-specific cost J , which encompasses safety constraints and efficiency.

IV. METHODOLOGY

As illustrated in Fig. 2, SanD-Planner presents a two-stage generate-and-select pipeline: 1) *trajectory generation* within a B-spline control-point space (Sections IV-B and IV-C), and 2) *critic-based selection* with explicit geometric objectives (Section IV-D). In the first stage, the policy network uses a conditional diffusion model to sample K candidate control-point sets $\{\mathcal{Q}_t^{(k)}\}_{k=1}^K$ conditioned on the context tensor $\mathcal{C}_t$. Subsequently, these discrete sets are mapped to continuous local trajectories via B-spline basis functions. In the second stage, a lightweight ESDF-based critic module ranks these candidates by evaluating a task-specific cost function. The trajectory with the lowest cost, τ_t^* , is selected as the local plan and executed in a receding-horizon manner. The initial heading velocity of the selected trajectory is fed back as $\mathbf{v}_{t+1}^{\text{prev}}$ to condition the next planning step.

A. Perception and Condition Encoding

At each planning step t , the agent receives historical observations $\mathcal{O}_t$ to capture environmental dynamics. Each depth

frame is encoded using a lightweight ResNet-18 backbone $f_{\mathcal{O}}$ [44]. To preserve fine-grained geometric details required for local navigation (e.g., obstacle boundaries for collision avoidance and stair-step structures for climbing), we extract spatial feature maps $\mathbf{F} \in \mathbb{R}^{H \times W \times D}$ for each frame from an intermediate layer rather than the final layer. These spatial features $\mathbf{F}$ are then flattened into a sequence of $H \times W$ visual tokens. To maintain spatial and temporal structure, we augment each token with 2D positional embeddings [45] and a learnable temporal embedding to distinguish tokens from different time steps. The resulting visual embedding sequence is denoted as $\mathbf{e}_{\mathcal{O}} = f_{\mathcal{O}}(\mathcal{O}_t) \in \mathbb{R}^{4 \times (HW) \times D}$.

Simultaneously, we project the relative goal $\mathbf{g}_t$ and the previously executed velocity $\mathbf{v}_t^{\text{prev}}$ into the same latent dimension D via dedicated MLPs, serving as encoders f_g and f_v . This yields the goal token $\mathbf{e}_g = f_g(\mathbf{g}_t)$ and the velocity token $\mathbf{e}_v = f_v(\mathbf{v}_t^{\text{prev}})$. If historical velocity is unavailable (e.g., $t = 0$), a learnable null token substitutes $\mathbf{e}_v$. All tokens are then concatenated into a sequence and aggregated by a two-layer Transformer encoder Φ . The final output serves as the multi-modal context $\mathcal{C}_t = \Phi([\mathbf{e}_{\mathcal{O}}, \mathbf{e}_g, \mathbf{e}_v])$, which is used to condition the diffusion denoising process.

B. Trajectory Parameterization via B-Spline Control Points

We represent each trajectory as a clamped cubic B-spline which inherently possesses C^2 continuity to ensure its smoothness [46–48]. Given a B-spline of degree p and a set of knot vectors $\mathcal{U} = \{u_0, \dots, u_{N+p}\}$, the trajectory τ_t is defined as

$$\tau_t(u) = \sum_{i=0}^{N-1} B_{i,p}(u) \mathbf{Q}_{t,i}, \quad u \in [u_p, u_N], \quad (1)$$

where $B_{i,p}(\cdot)$ are the B-spline basis functions. We employ a uniform knot vector, anchoring the curve at the endpoints so that $\tau_t(u_p) = \mathbf{Q}_{t,0}$ and $\tau_t(u_N) = \mathbf{Q}_{t,N-1}$. The control points $\mathbf{Q}_t$ are defined in the robot frame, hence $\mathbf{Q}_{t,0} = \mathbf{0}$.

This parameterization transforms the learning objective from a discrete sequence of waypoints to a compact set of control points. We view this reparameterization as a structured inductive bias for IL: it constrains the policy to a function

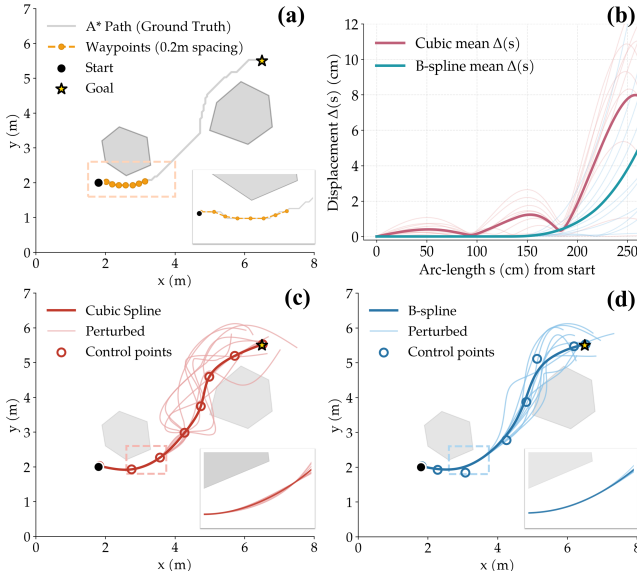


Fig. 3: Comparison of trajectory representations among discrete waypoints, cubic spline, and B-spline. (a) Ground-truth A* path (gray) and an 8-waypoint sequence (orange, 0.2m spacing). (b) Mean arc-length displacement $\Delta(s)$ under identical perturbations ($10\times$ disturbances). Over the first 1.5m, the B-spline deviation stays near zero, whereas the cubic spline oscillates. (c, d) We first fit the same ground-truth path with an 8-point interpolating cubic spline (c) and an 8-point B-spline (d), and then apply the random perturbations to the last four points within a 1m radius. Owing to local support and convex hull property, B-splines avoid overshoot and better preserve the global path shape, yielding smaller deviations than cubic splines under the same perturbation range.

class that is compact, smooth, and locally controlled. To make this bias concrete, we contrast B-splines against two prevalent IL parameterizations, discrete waypoints [8, 19, 31] and interpolating cubic splines [9, 10], with toy examples illustrated in Fig. 3.

1) *Parameter Efficiency*: As shown in Fig. 3 (a) and (c), B-spline control-point modeling decouples the number of predicted points from the spatial sampling resolution: with the same fixed output budget, the model predicts the shape of a continuous curve rather than a short sequence of densely spaced waypoints. This allows an 8-control-point B-spline to cover a longer local lookahead than an 8-waypoint rollout under the same output dimension, thereby mitigating waypoint-induced myopia [21]. We note that this compactness comes with a controlled loss of expressivity for longer-horizon planning, as the representation favors smooth, low-frequency trajectories rather than arbitrary path geometries. For short-horizon local planners studied here, 8 points empirically balance representational capability and training efficiency.

2) *Smoothness*: A trackable robot trajectory should exhibit continuous position, velocity, and acceleration profiles, otherwise the downstream controller is forced to absorb discontinuities that the planner emits. Rather than expecting

the policy to acquire this property from limited and noisy expert demonstrations, we encode smoothness directly into the representation: cubic B-splines are C^2 -continuous by construction, so any trajectory generated from the predicted control points is guaranteed to be smooth. Together with the compact, longer-horizon output discussed above, this structural prior relieves the policy of learning smoothness on top of dense waypoint placement, thereby simplifying the learning objective and enhancing trackability for downstream control.

3) *Representation-Level Robustness*: Depth sensing often suffers from partial, noisy observations and occlusions, which compromise the reliability of long-range planning. B-splines offer representation-level robustness against this issue through two key properties.

Local Support: A B-spline segment depends only on $p + 1$ local control points. Consequently, noise-induced deviations in distal control points remain isolated and do not reshape the entire path. This property aligns naturally with receding-horizon planning, allowing the robot to execute a stable near-horizon segment while iteratively correcting far-horizon errors in subsequent steps. In contrast, globally coupled interpolating splines propagate local perturbations across the entire trajectory, potentially destabilizing immediate actions.

Convex Hull Property: To ensure safety under distribution shifts, the representation should prevent prediction errors from amplifying into large path deviations. Since B-spline basis functions $B_{i,p}(u)$ satisfy the partition of unity [48] (i.e., $\sum B_{i,p}(u) = 1, B_{i,p} \geq 0$), every point on the trajectory $\tau_t(u)$ lies within the convex hull of its control points. For a predicted set with errors $\mathbf{Q}'_{t,i} = \mathbf{Q}_{t,i} + \Delta\mathbf{Q}_{t,i}$, the trajectory deviation is uniformly bounded by the maximum control-point error:

$$\max_u \|\tau'_t(u) - \tau_t(u)\| \leq \max_i \|\Delta\mathbf{Q}_{t,i}\|. \quad (2)$$

Eq. (2) shows that B-splines prevent error amplification by providing a strict upper bound on path deviation. As shown in Fig. 3, B-splines preserve the global trajectory shape even under perturbation, whereas interpolating splines often suffer from Runge’s phenomenon [49], where small nodal errors induce disproportionately large overshoots.

C. Diffusion Policy for Local Trajectory Generation

1) *Diffusion in B-Spline Control-Point Space*: We model local planning as conditional generation in B-spline control-point space. Given the conditioning context $\mathcal{C}_t$, the generator samples a set of control points $\mathcal{Q}_t$, which is deterministically mapped to a continuous trajectory τ_t via the B-spline representation. Let $\bar{\mathbf{Q}}_t \in \mathbb{R}^{N \times 3}$ denote the expert control points obtained by fitting a B-spline to an expert trajectory, and let $\mathbf{x}_0 := \text{vec}(\bar{\mathbf{Q}}_t)$ be the corresponding clean control-point vector. We train a conditional Denoising Diffusion Probabilistic Model (DDPM) [50] to model $p_\theta(\mathbf{x}_0 | \mathcal{C}_t)$. With a predefined noise schedule $\{(\alpha_s, \sigma_s)\}_{s=1}^S$ satisfying $\alpha_s^2 + \sigma_s^2 = 1$, the forward process is

$$q(\mathbf{x}_s | \mathbf{x}_0) = \mathcal{N}(\alpha_s \mathbf{x}_0, \sigma_s^2 \mathbf{I}), \quad s \in \{1, \dots, S\}, \quad (3)$$

where S denotes the total number of denoising steps and s is the current diffusion step. We employ the v -prediction parameterization [51] with the target defined as $\mathbf{v}_s := \alpha_s \epsilon - \sigma_s \mathbf{x}_0$, where $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. A denoiser is trained to predict the diffusion target conditioned on the context embedding and timestep. Specifically, we use a conditional 1D U-Net over the control-point sequence with cross-attention to $\mathcal{C}_t$ and timestep embeddings injected into each block, and predict the v -parameterization target $\hat{\mathbf{v}} = \hat{\mathbf{v}}_\theta(\mathbf{x}_s, \mathcal{C}_t, s)$. The training objective is the standard v -prediction loss [51]

$$\mathcal{L}_{\text{diff}}(\theta) = \mathbb{E}_{\mathbf{x}_0, \epsilon, s} \left[\left\| \mathbf{v}_s - \hat{\mathbf{v}}_\theta(\mathbf{x}_s, \mathcal{C}_t, s) \right\|_2^2 \right]. \quad (4)$$

During inference, given c_t , we initialize the diffusion process from Gaussian noise and run S reverse denoising steps to sample K candidate B-spline control-point sets $\mathcal{Q}_t^{(k)} \sim p_\theta(\mathcal{Q}_t \mid \mathcal{C}_t)$ for $k \in \{1, \dots, K\}$, which are subsequently mapped to continuous local plans $\{\tau_t^{(k)}\}_{k=1}^K$.

2) *Temporal Consistency*: Receding-horizon planning under partial observability is often inherently multi-modal, where valid paths exist in distinct topological modes [43] (e.g., bypassing an obstacle from left or right). If consecutive plans switch between these modes, the executed behavior may oscillate, yielding jittery headings and collisions [52]. To increase temporal consistency in the generated candidates, we augment the conditioning signal with an explicit previous-plan token $\mathbf{v}_t^{\text{prev}}$ representing the heading direction from the previously selected trajectory τ_{t-1}^* .

With a clamped cubic B-spline, the initial heading direction is proportional to the difference between the first and the second control points, i.e., $\dot{\tau}(0) \propto (\mathbf{Q}_{t-1,1} - \mathbf{Q}_{t-1,0})$. In our robot-centric frame, the first control point is anchored at the origin $\mathbf{Q}_{t-1,0} = \mathbf{0}$. The last heading direction is obtained as $\mathbf{v}_t^{\text{prev}} = \text{norm}(\mathbf{Q}_{t-1,1}^*)$, where $\mathbf{Q}_{t-1,1}^*$ denotes the second control point of τ_{t-1}^* . To simplify training and stabilize the learning process, we use privileged information for $\mathbf{v}_t^{\text{prev}}$. During training process, $\mathbf{v}_t^{\text{prev}}$ is computed from the ground-truth expert control points $\bar{\mathbf{Q}}_{t,2}$ rather than from the previous step. For the first planning step, we use a learned null token to indicate the absence of previous-plan information.

D. Safety-Aware Candidate Selection via the Critic Module

While the diffusion policy effectively captures the multi-modal distribution of expert behaviors, the sampled trajectories are not guaranteed to satisfy safety constraints [53, 54]. To decouple complex distribution modeling from strict safety verification, we employ a generate-and-select pipeline. At each planning step t , the diffusion policy samples K candidate control-point sets $\{\mathcal{Q}_t^{(k)}\}_{k=1}^K$ conditioned on the context $\mathcal{C}_t$, which generate continuous trajectories $\{\tau_t^{(k)}\}_{k=1}^K$. These are evaluated by an explicit geometric critic, which selects the optimal one τ_t^* by minimizing a cost function J :

$$\tau_t^* = \arg \min_{\tau \in \{\tau_t^{(k)}\}_{k=1}^K} J(\tau). \quad (5)$$

In contrast to learning-based trajectory ranking [19] which requires extensive training and often fails under distribution shifts, we propose an explicit, analytic geometric critic. This module leverages real-time sensor data for robust safety verification without additional supervision. We discretize each candidate spline $\tau_t^{(k)}$ into M waypoints $\{\mathbf{x}_j\}_{j=0}^{M-1}$ equidistantly spaced in arc length. The cost J of each path is then evaluated using the following terms.

1) *Discounted Safety Cost*: To score the safety of each trajectory candidate, we build a robot-centric ESDF map $E(\mathbf{x})$ using the current depth image $\mathbf{D}_t$. Each voxel stores the signed distance to the nearest obstacle surface, where positive values indicate free space and negative values lie inside obstacles. We prioritize obstacle clearance. Since the planner operates in a receding-horizon manner and executes only the near-horizon segment before replanning, later waypoints should not be weighted equally. We design a temporal discount factor $\gamma \in (0, 1)$ to prioritize near-field safety:

$$J_{\text{esdf}}(\tau_t^{(k)}) = \frac{1}{\sum_{j=0}^{M-1} \gamma^j} \sum_{j=0}^{M-1} \gamma^j \max(0, d_{\text{safe}} - E(\mathbf{x}_j^{(k)})), \quad (6)$$

where $E(\mathbf{x})$ is the ESDF value at position $\mathbf{x}$, and d_{safe} is the desired safety margin. This formulation imposes a weighted penalty on clearance violations, placing greater importance on the near-horizon segment of the execution path.

2) *Path Efficiency*: To encourage the efficiency of the selected plan, we penalize unnecessary detours and prioritize candidates that terminate closest to the local goal $\mathbf{g}_t$.

$$J_{\text{len}}(\tau_t^{(k)}) = \sum_{j=0}^{M-2} \left\| \mathbf{x}_{j+1}^{(k)} - \mathbf{x}_j^{(k)} \right\|_2, \quad (7)$$

$$J_{\text{goal}}(\tau_t^{(k)}) = \left\| \mathbf{x}_{M-1}^{(k)} - \mathbf{g}_t \right\|_2.$$

The total trajectory cost for each candidate path is a weighted sum of the three terms:

$$J(\tau_t^{(k)}) = \lambda_1 J_{\text{esdf}}(\tau_t^{(k)}) + \lambda_2 J_{\text{len}}(\tau_t^{(k)}) + \lambda_3 J_{\text{goal}}(\tau_t^{(k)}), \quad (8)$$

where $\lambda_1, \lambda_2, \lambda_3 > 0$ balance safety and efficiency.

V. EXPERIMENTS

This section conducts experiments to answer the following questions:

- **Q1.** How does SanD-Planner perform compared to other vision-based baselines? (Section V-B)
- **Q2.** How does SanD-Planner scale with the number of expert trajectories (i.e., sample efficiency)? (Section V-C)
- **Q3.** How does trajectory representation affect planning performance? (Section V-D)
- **Q4.** How much does critic-based candidate selection contribute, and how does performance scale with the number of candidates K ? (Section V-E)
- **Q5.** How does the velocity token affect planning consistency? (Section V-F)
- **Q6.** Can SanD-Planner generalize zero-shot to real-world and unseen complex environments? (Section V-G)

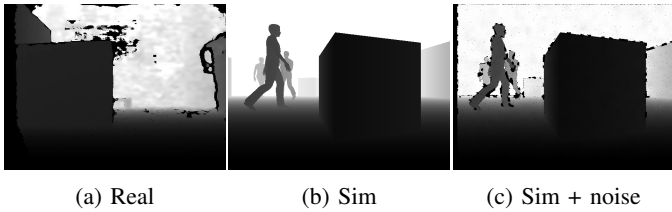


Fig. 4: Comparison of depth images. (a) Real-world capture. (b) Clean simulated depth. (c) Domain randomization depth.



Fig. 5: Simulation benchmark environments [19]. Left: **ClutteredEnv** provides diverse geometric obstacle layouts. Right: **InternScenes** features photorealistic indoor settings.

A. Experimental Setup

1) *Training Dataset and Implementation Details*: Our training dataset is collected entirely in simulation across 10 diverse environments, including two photorealistic Matterport3D scenes [55] and eight Gazebo worlds comprising tunnels and indoor layouts [3]. To facilitate 3D local planning, we explicitly incorporate multi-level structures (*e.g.*, stairs) and low-clearance obstacles. We collect 500 navigation episodes using the PCT-planner [6] as the expert to generate collision-free 3D trajectories while logging the egocentric depth stream. To maximize data efficiency, we construct training pairs by randomly sampling sub-trajectories from each expert episode. Each segment is defined by a random start pose and a future waypoint along the same expert path, which is then fitted with a clamped cubic B-spline using eight control points. This resampling strategy generates diverse and informative labels from a limited set of demonstrations, significantly enhancing the policy’s sample efficiency.

The model is trained on a desktop equipped with an NVIDIA RTX4080 GPU, requiring approximately 5 hours to converge on the 500-trajectory dataset. During inference, candidate trajectories are sampled using the deterministic DPM-Solver++ [56], achieving an operational frequency of approximately 8Hz. For real-world validation in Section V-G, the system is deployed on a laptop with an RTX4090 to ensure low-latency planning during autonomous navigation.

2) *Observation Domain Randomization*: To bridge the sim-to-real gap, we apply domain randomization to observations during training: 1) **Temporal Latency**: we introduce a temporal freezing mechanism to mimic compute and communication delays. With a probability of 0.1, we either freeze the observation buffer $\mathcal{O}_t$ to the latest frame or replicate individual frames

TABLE I: Quantitative comparison on InternNav benchmarks. Results of the baseline methods are taken from [19].

Benchmark Environment	Method	SR $\uparrow$	SPL $\uparrow$
ClutteredEnv	iPlanner	84.8	83.6
	ViPlanner	72.4	72.3
	NavDP	89.8	87.7
	SanD-Planner	90.1	84.0
InternScenes	iPlanner	48.8	46.7
	ViPlanner	54.3	52.5
	NavDP	65.7	60.7
	SanD-Planner	72.0	63.7

TABLE II: Training data and compute resources.

Method	Training Data	Training Time
iPlanner	30k depth images	1 $\times$ RTX3090 Ti, $\approx$ 20h
ViPlanner	80k start-goal pairs	1 $\times$ RTX3090, $\approx$ 6h
NavDP	200k traj. ($\approx$ 40M images)	32 $\times$ A100, 24h
Ours	0.5k traj. ($\approx$ 35k images)	1 $\times$ RTX4080, $\approx$ 5h

from their predecessors, simulating signal loss or cold starts. 2) **Sensor Noise**: We emulate noise patterns characteristic of the Intel RealSense D435. Following parameters in [57], we inject distance-dependent axial noise that scales quadratically with depth, as visualized in Fig. 4. We also synthesize random pixel dropouts with 0.1 probability, a stereo-occlusion band on the left margin, and spatially correlated speckle noise. Such perturbations force the policy to prioritize structural geometric cues over perfect depth measurements.

3) *Simulation Benchmark*: We evaluate SanD-Planner using the InternNav Benchmark [19], which provides photorealistic indoor environments designed to minimize the sim-to-real gap for local obstacle avoidance. The benchmark features diverse assets and predefined start-goal pairs for point-goal navigation. To ensure a fair comparison, all baselines are evaluated under the same protocol as shown in Fig. 5. Following the protocol, failure is defined as a collision-caused timeout.

B. Quantitative Performance Comparison

We evaluate SanD-Planner against three SoTA visual planners: iPlanner [9], ViPlanner [10], and NavDP [19], regarding the point-goal navigation task. NoMaD [18] is excluded as it targets image-goal navigation rather than our point-goal setting. Following the benchmark protocol, we report Success Rate (SR) and Success weighted by Path Length (SPL) [58] across 2020 episodes in ClutteredEnv and 4040 episodes in InternScenes. Notably, despite training on only 500 trajectories ($\approx$ **0.25%** of the 200k trajectories used by NavDP), SanD-Planner achieves competitive or superior performance across both benchmarks (Tab. I). On the realistic InternScenes, SanD-Planner outperforms NavDP by large margins (+6.3% SR and +3.0% SPL), demonstrating robust generalization to complex indoor layouts and diverse obstacle distributions. In ClutteredEnv, SanD-Planner achieves the highest 90.1% SR while maintaining a competitive 84.0% SPL. The slight SPL gap compared to NavDP is partly attributable to the B-spline parameterization, which embeds a structural bias toward

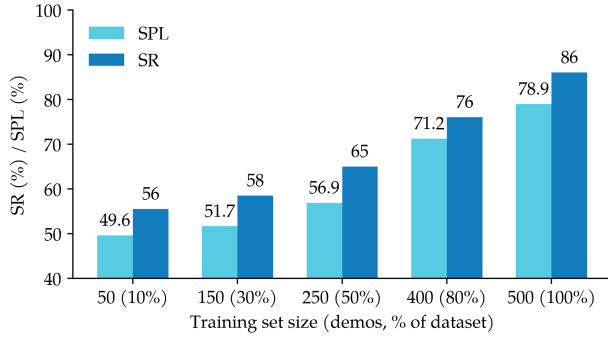


Fig. 6: Ablation study on training data size in ClutteredEnv. SR and SPL improve consistently with training dataset size.

safer, albeit more conservative, detours. However, analysis in Section V-C suggests this gap arises from limited data scale rather than methodological constraints and could be further narrowed with additional supervision. DD-PPO [59] achieved only 6.6% SR / 6.5% SPL on ClutteredEnv due to the mismatch between its step-wise discrete actions and our trajectory-level protocol, so we exclude it as a baseline.

The performance comparison between iPlanner (depth-only) and ViPlanner (depth and semantics) highlights the trade-offs of multi-modal inputs. While ViPlanner leverages semantics in InternScenes (+5.5% SR over iPlanner), it suffers a severe performance degradation in ClutteredEnv (−12.4% SR), where primitive obstacles lack well-defined semantic categories. This underscores the potential brittleness of relying on semantic cues under distribution shifts. Crucially, SanD-Planner outperforms ViPlanner on both benchmarks using depth observations alone. These results suggest that depth information provides sufficient geometric cues for general 2D collision avoidance, whereas semantics may degrade model generalization in out-of-distribution (OOD) scenes when training data is limited.

Tab. II summarizes the reported training data units and compute budgets of representative learning-based local planners. These numbers serve as contextual references only, since methods differ in supervision, data unit, and training protocol, making budgets not strictly commensurate. NavDP serves as the most relevant comparison because it is also a trajectory-level diffusion planner. NavDP is trained as a broader generalist model using 200k trajectories, approximately 40M images, across 3k scenes, whereas our current model is trained on 500 trajectories, approximately 35k logged depth images, from 10 scenes. The results show SanD-Planner demonstrates superior efficiency, requiring orders of magnitude less training data and compute to achieve SoTA performance.

C. Sample Efficiency and Data Scaling

We investigate the sample efficiency of SanD-Planner by training on trajectory-level subsets ranging from 10% to 100% of the full dataset. As shown in Fig. 6, our approach exhibits monotonic performance gains as the number of demonstrations increases. Notably, the model is highly efficient even in data-scarce regimes, achieving a non-trivial 55.5% SR with only

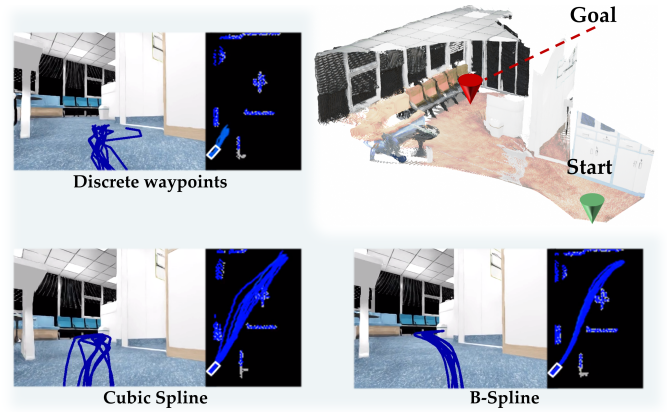


Fig. 7: Comparison of trajectory representations in simulation. Top-right: the scene reconstruction with the same start and goal. For each representation (discrete waypoints, cubic spline, and B-spline), we show the ego-view image and a bird-view radar visualization with 10 candidate trajectories overlaid.

TABLE III: Ablation on trajectory representation. Average SR and SPL on two ClutteredEnv and two InternScenes scenes.

Benchmark	Representation	SR [%]↑	SPL [%]↑
ClutteredEnv	Waypoints	75.5	65.8
	Cubic spline	75.0	70.6
	B-spline (Ours)	93.0	83.8
InternScenes	Waypoints	75.5	67.8
	Cubic spline	72.0	65.8
	B-spline (Ours)	83.5	72.1

50 expert trajectories (10% data). With 250 trajectories (50% data), the model attains 65.0% SR, recovering 76% of its peak performance. Although basic collision avoidance can be learned from minimal data, SPL continues to improve substantially as the dataset scales to its full size (49.6% → 78.9%). This suggests that small datasets are sufficient for acquiring safe behaviors, whereas larger demonstration sets help the planner learn more efficient bypass strategies and reduce unnecessary detours.

D. Ablation Study on Trajectory Representation

To validate the impact of trajectory parameterization, we ablate the output representation while keeping the rest of the navigation pipeline fixed. All variants predict eight anchor points to maintain identical output dimensionality. The discrete waypoint baseline utilizes a fixed 0.2m spacing, resulting in a 1.4m lookahead. In contrast, both interpolating cubic splines [9] and our B-spline representation parameterize an extended horizon of approximately 6m. Tab. III shows that the choice of representation influences performance, with B-splines consistently outperforming the alternatives. In ClutteredEnv (episodes > 20m), discrete waypoints suffer from the myopic behavior [21]; while they match cubic splines with ≈ 75% SR, they exhibit lower efficiency (65.8% vs. 70.6% SPL) due to suboptimal detours. Spline-based parameterizations mitigate this by enforcing smoothness and extending the

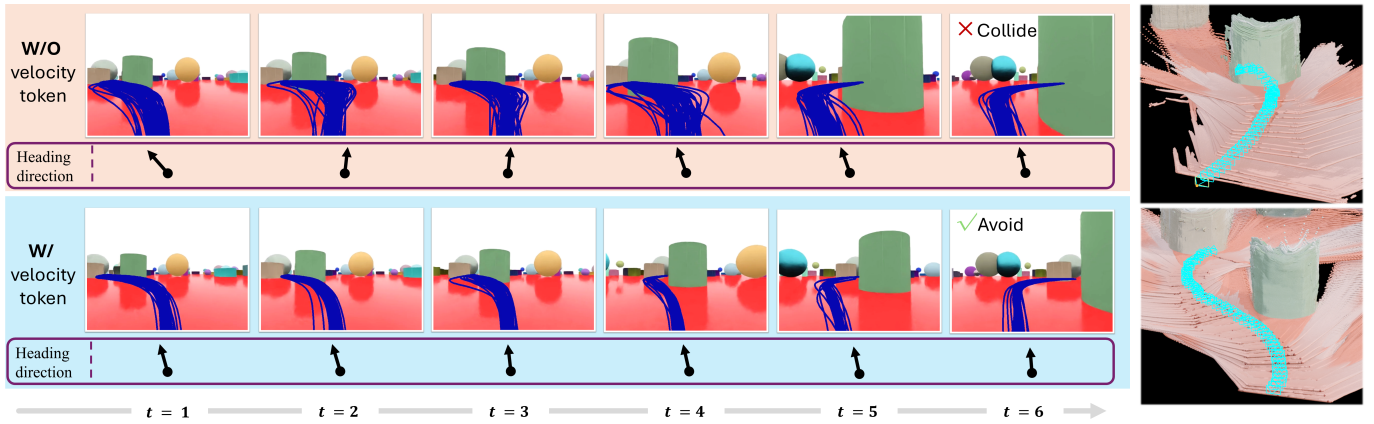


Fig. 8: **Representative comparison of trajectory consistency with and without the velocity token.** We visualize the diffusion-sampled bypass candidates (blue) at each planning step t . Without the velocity token (top), the candidate set remains highly multi-modal and dispersed, and the critic-selected solution alternates between bypass modes across consecutive replans, producing inconsistent heading directions (arrows) and eventually leading to a collision. With the velocity token (bottom), the sampled candidates are more temporally consistent, yielding stable headings and a safe bypass. The rightmost panel shows the corresponding 3D reconstructions from Depth Anything 3 [60], where the cyan curve indicates the camera/robot trajectory.

planning horizon within the same parameter budget.

The performance gap of B-splines highlights the critical role of local support. The interpolated cubic splines are globally coupled, causing prediction errors in occluded far-field regions to propagate backward and destabilize the immediate execution segment. Conversely, B-splines possess compact support, which effectively isolates far-field uncertainty and ensures near-field stability. In the confined InternScenes environment, the advantage of long-horizon planning is less pronounced due to heavy occlusions. This explains why myopic waypoints slightly outperform cubic splines in such settings. However, B-splines maintain the best overall performance, balancing near-horizon robustness with path efficiency. Qualitative results in Fig. 7 further illustrate this robustness: waypoints yield non-smooth, less trackable paths, and cubic splines amplify prediction noise across the entire curve. In contrast, B-splines generate consistent candidates with localized variations, demonstrating lower sensitivity to perceptual uncertainty.

E. Ablation Study on Candidate Selection

This section further investigates the contribution of the critic-based candidate selection strategy, since the final performance of SanD-Planner depends on both the B-spline diffusion policy and the online selection module. To disentangle these factors, we fix the trained B-spline diffusion model and vary the number of sampled candidates K during inference. When $K = 1$, the planner executes a single diffusion sample, and the critic-based minimization in Eq. (5) becomes a trivial selection over a singleton set. This setting, therefore, discards the candidate selection. For $K > 1$, the same diffusion model produces K B-spline candidates, and the ESDF-based critic selects the trajectory with the lowest cost according to Eq. (8).

Tab. IV reports the results on ClutteredEnv. With $K = 1$, the policy alone achieves 75.5% SR and 66.8% SPL, indicating that the learned B-spline policy can generate non-trivial

TABLE IV: Ablation on the number of diffusion candidates K on 2 ClutteredEnv scenes.

K	1	4	8	16	32
SR [%] $\uparrow$	75.5	88.0	90.0	90.0	93.0
SPL [%] $\uparrow$	66.8	78.5	80.9	81.0	83.8

feasible trajectories without selection. However, increasing K to 4 improves SR from 75.5% to 88.0% and SPL from 66.8% to 78.5%, showing that candidate selection provides a substantial improvement. Further increasing K yields non-decreasing performance, reaching 93.0% SR and 83.8% SPL at $K = 32$. The gains become smaller after $K = 8$, suggesting diminishing returns once the candidate set contains sufficiently diverse feasible modes.

These results clarify the role of the critic. The critic module significantly improves execution-time robustness by rejecting unsafe or inefficient samples. At the same time, the $K = 1$ result shows that the B-spline diffusion policy itself remains a meaningful planner rather than relying entirely on the critic. Together with the representation ablation in Section V-D, these results indicate that the two components contribute complementary benefits. The B-spline parameterization improves the quality and stability of the generated candidate distribution, whereas the critic exploits this candidate set to perform safety-aware selection.

F. Ablation Study on Velocity Token for Temporal Consistency

We evaluate the impact of previous plan information on planning consistency by ablating the velocity token $\mathbf{v}_t^{\text{prev}}$. As reported in Tab. V, removing this temporal conditioning leads to a consistent performance drop, which confirms its necessity for trajectory reliability. Analysis reveals that collisions without $\mathbf{v}_t^{\text{prev}}$ frequently occur near geometrically symmetric

TABLE V: Velocity-token ablation on ClutteredEnv scenes.

Scene	w/o token		w/ token	
	SR [%]↑	SPL [%]↑	SR [%]↑	SPL [%]↑
Scene 1	79.0	0.73	87.0	0.78
Scene 2	93.0	0.86	99.0	0.89
Average	86.0	0.80	93.0	0.84

obstacles (e.g., pillars), where an example is shown in Fig. 8. Under partial observability, such structures induce bimodal distributions of plausible bypass trajectories. Without velocity conditioning, the diffusion model generates dispersed, high-variance candidates at each step. Since the critic selects the optimal trajectory independently per step, the planner suffers from topological inconsistency, frequently switching between conflicting bypass modes. This results in oscillatory steering and eventual collisions. In contrast, the $\mathbf{v}_t^{\text{prev}}$ condition effectively regularizes the generative process, yielding concentrated candidate sets consistent with the robot’s established motion and ensuring a stable heading.

G. Real-world Experiments

Real-world experiments are deployed on a Unitree Go2 quadruped robot. The platform features an Intel RealSense D435 depth camera with an IR-pass filter, with algorithms executed on a tethered laptop. These experiments are designed to qualitatively verify the zero-shot sim-to-real capability of our system under unseen scenes, varying illumination, and noise from realistic depth sensors. The planner operates at approximately 10Hz, with trajectories tracked by an onboard Model Predictive Controller (MPC) [19]. To facilitate high-frequency replanning in a receding-horizon setting, we employ a warm-start strategy: each subsequent planning cycle is initialized from the previous solution and denoised only from an intermediate diffusion timestep, i.e., from step 6 in a 10-step sampling schedule. At each step, $K = 16$ candidate trajectories are sampled in parallel, with the optimal plan selected by the geometric critic module.

To assess SanD-Planner’s generalization, we conduct experiments across diverse scenarios with static and dynamic pedestrians, as shown in Fig. 1 and Fig. 9. Static tests include narrow mazes and cluttered office with tight clearances. We further evaluate outdoor navigation, where illumination changes and complex geometries present significantly more challenging perception conditions than simulation. Additionally, we demonstrate 3D navigation capability through stair-climbing trials. For these trials, we switch the Unitree Go2 to its built-in stair-climbing locomotion mode so that the same planner output drives genuine 3D motion through the controller. Despite these diverse settings, SanD-Planner generates smooth, collision-avoiding trajectories, confirming the effectiveness and zero-shot capability of our method.

VI. CONCLUSION AND FUTURE WORK

This paper presents SanD-Planner, a sample-efficient, diffusion-based planner using clamped cubic B-spline pa-

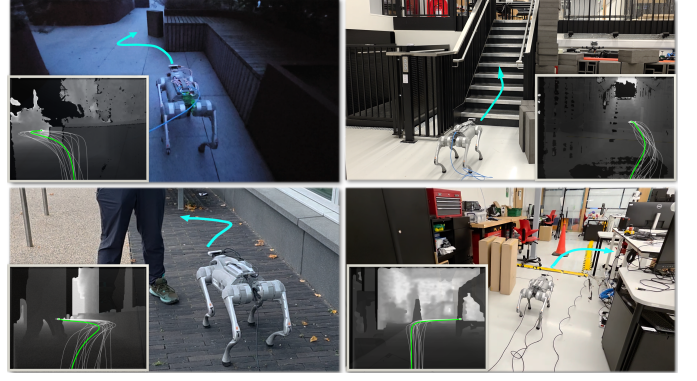


Fig. 9: Demonstrations of SanD-Planner in unseen environments. Depth inputs and predicted trajectories are also shown, with the selected trajectory highlighted in green.

rameterization. By predicting control points and decoupling the safety check via a critic module, our method reduces learning complexity and data dependency. Benchmarks show that with only 500 expert episodes, SanD-Planner matches or surpasses SoTA performance. Our findings suggest that: 1) B-splines isolate sensing uncertainty better than waypoints or interpolating splines, yielding more stable planning; and 2) depth sensing provides sufficient geometric cues for collision avoidance. Zero-shot real-world deployment further validates our method, enabling 2D and 3D navigation tasks such as point-goal navigation and stair traversal without fine-tuning.

Despite its strong performance, SanD-Planner inherits the limitations of depth sensing, especially for small, thin, or specular objects. Future work will investigate integrating monocular depth foundation models, such as DepthPro [61], as complementary geometric inputs to improve perceptual coverage. However, their scale ambiguity and prediction uncertainty require explicit handling; conservative ESDF margins can only partially compensate for such errors. Since our ESDF is constructed from local observations, partial observability remains. The planner emphasizes short-horizon reactive safety, while long-term spatial memory is better handled by a global planner. In dynamic scenes, receding-horizon replanning allows the system to react to moderate changes, but systematic dynamic-scene benchmarking remains future work. The modular critic, currently used for post-hoc candidate selection, could also be extended to test-time guidance or trajectory refinement under distribution shift. Finally, although this work focuses on sample efficiency, our data-scaling results suggest further capacity: training on larger and more diverse demonstrations may improve generalization across varied environments.

VII. ACKNOWLEDGEMENT

This work was supported by the UKRI Future Leaders Fellowship [MR/V025333/1] (RoboHike). For the purpose of Open Access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

REFERENCES

- [1] Faiza Gul, Wan Rahiman, and Syed Sahal Nazli Alhady. A comprehensive study for robot navigation techniques. *Cogent Engineering*, 6(1):1632046, 2019.
- [2] Yuri DV Yasuda, Luiz Eduardo G Martins, and Fabio AM Cappabianco. Autonomous visual navigation for mobile robots: A systematic literature review. *ACM Computing Surveys (CSUR)*, 53(1):1–34, 2020.
- [3] Chao Cao, Hongbiao Zhu, Fan Yang, Yukun Xia, Howie Choset, Jean Oh, and Ji Zhang. Autonomous exploration development environment and the planning algorithms. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 8921–8928. IEEE, 2022.
- [4] Lydia E Kavraki, Petr Svestka, J-C Latombe, and Mark H Overmars. Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE transactions on Robotics and Automation*, 12(4):566–580, 2002.
- [5] Nathan Ratliff, Matt Zucker, J Andrew Bagnell, and Siddhartha Srinivasa. Chomp: Gradient optimization techniques for efficient motion planning. In *2009 IEEE international conference on robotics and automation*, pages 489–494. IEEE, 2009.
- [6] Bowen Yang, Jie Cheng, Bohuan Xue, Jianhao Jiao, and Ming Liu. Efficient global navigational planning in 3-d structures based on point cloud tomography. *IEEE/ASME Transactions on Mechatronics*, 30(1):321–332, 2024.
- [7] Jianhao Jiao, Changkun Liu, Jingwen Yu, Boyi Liu, Qianyi Zhang, Yue Wang, and Dimitrios Kanoulas. Opennavmap: Structure-free topometric mapping via large-scale collaborative localization. *arXiv preprint arXiv:2601.12291*, 2026.
- [8] Dhruv Shah, Ajay Sridhar, Arjun Bhorkar, Noriaki Hirose, and Sergey Levine. Gnm: A general navigation model to drive any robot. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 7226–7233, 2023. doi: 10.1109/ICRA48891.2023.10161227.
- [9] Fan Yang, Chen Wang, Cesar Cadena, and Marco Hutter. iPlanner: Imperative Path Planning. In *Proceedings of Robotics: Science and Systems*, Daegu, Republic of Korea, July 2023. doi: 10.15607/RSS.2023.XIX.064.
- [10] Pascal Roth, Julian Nubert, Fan Yang, Mayank Mittal, and Marco Hutter. Viplanner: Visual semantic imperative learning for local navigation. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5243–5249. IEEE, 2024.
- [11] Jianwei Liu, Maria Stamatopoulou, and Dimitrios Kanoulas. Dipper: Diffusion-based 2d path planner applied on legged robots. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9264–9270. IEEE, 2024.
- [12] David Hoeller, Lorenz Wellhausen, Farbod Farshidian, and Marco Hutter. Learning a state representation and navigation in cluttered and dynamic environments. *IEEE Robotics and Automation Letters*, 6(3):5081–5088, 2021.
- [13] Haobin Shi, Lin Shi, Meng Xu, and Kao-Shing Hwang. End-to-end navigation strategy with deep reinforcement learning for mobile robots. *IEEE Transactions on Industrial Informatics*, 16(4):2393–2402, 2019.
- [14] Karmesh Yadav, Ram Ramrakhy, Arjun Majumdar, Vincent-Pierre Berges, Sachit Kuhar, Dhruv Batra, Alexei Baevski, and Oleksandr Maksymets. Offline visual representation learning for embodied navigation. In *Workshop on Reincarnating Reinforcement Learning at ICLR 2023*, 2023.
- [15] Shubhra Aich, Wenshan Wang, Parv Maheshwari, Matthew Sivaprakasam, Samuel Triest, Cherie Ho, Jason M Gregory, John G Rogers III, and Sebastian Scherer. Deep bayesian future fusion for self-supervised, high-resolution, off-road mapping. *arXiv preprint arXiv:2403.11876*, 2024.
- [16] Jincheng Wang, Penny Karanasou, Pengyuan Wei, Elia Gatti, Diego Martinez Plasencia, and Dimitrios Kanoulas. Long-short decision transformer: Bridging global and local dependencies for generalized decision-making. In *The Thirteenth International Conference on Learning Representations*, 2025.
- [17] Dhruv Shah, Ajay Sridhar, Nitish Dashora, Kyle Stachowicz, Kevin Black, Noriaki Hirose, and Sergey Levine. Vint: A foundation model for visual navigation. *arXiv preprint arXiv:2306.14846*, 2023.
- [18] Ajay Sridhar, Dhruv Shah, Catherine Glossop, and Sergey Levine. Nomad: Goal masked diffusion policies for navigation and exploration. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 63–70. IEEE, 2024.
- [19] Wenzhe Cai, Jiaqi Peng, Yuqiang Yang, Yujian Zhang, Meng Wei, Hanqing Wang, Yilun Chen, Tai Wang, and Jiangmiao Pang. Navdp: Learning sim-to-real navigation diffusion policy with privileged information guidance. *arXiv preprint arXiv:2505.08712*, 2025.
- [20] Yifan Peng, Yuze Hong, Ziyang Hong, Apple Pui-Yi Chui, and Junfeng Wu. Aquaticvision: Benchmarking visual slam in underwater environment with events and frames. In *2025 IEEE International Conference on Robotics and Automation (ICRA) Field Robotics Workshop*, 2025.
- [21] Matt Schmittle, Rohan Baijal, Nathan Hatch, Rosario Scalise, Mateo Guaman Castro, Sidharth Talia, Khimya Khetarpal, Byron Boots, and Siddhartha Srinivasa. Long range navigator (lrn): Extending robot planning horizons beyond metric maps. *arXiv preprint arXiv:2504.13149*, 2025.
- [22] Linghong Yao, Valerio Modugno, Andromachi Maria Delfaki, Yuanchang Liu, Danail Stoyanov, and Dimitrios Kanoulas. Local path planning among movable obstacles with deep reinforcement learning. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024. doi: 10.1109/IROS58592.2024.10802257. Acceptance Rate: 47.5%.
- [23] Ahmed Hussein, Mohamed Medhat Gaber, Eyad Elyan,

- and Chrisina Jayne. Imitation learning: A survey of learning methods. *ACM Computing Surveys (CSUR)*, 50(2):1–35, 2017.
- [24] Maryam Zare, Parham M Kebria, Abbas Khosravi, and Saeid Nahavandi. A survey of imitation learning: Algorithms, recent developments, and challenges. *IEEE Transactions on Cybernetics*, 54(12):7173–7186, 2024.
- [25] Mariusz Bojarski, Davide Del Testa, Daniel Dworakowski, Bernhard Firner, Beat Flepp, Prasoon Goyal, Lawrence D Jackel, Mathew Monfort, Urs Muller, Jiakai Zhang, et al. End to end learning for self-driving cars. *arXiv preprint arXiv:1604.07316*, 2016.
- [26] Mark Pfeiffer, Michael Schaeuble, Juan Nieto, Roland Siegwart, and Cesar Cadena. From perception to decision: A data-driven approach to end-to-end motion planning for autonomous ground robots. In *2017 IEEE international conference on robotics and automation (icra)*, pages 1527–1533. IEEE, 2017.
- [27] Mark Pfeiffer, Samarth Shukla, Matteo Turchetta, Cesar Cadena, Andreas Krause, Roland Siegwart, and Juan Nieto. Reinforced imitation: Sample efficient deep reinforcement learning for mapless navigation by leveraging prior demonstrations. *IEEE Robotics and Automation Letters*, 3(4):4423–4430, 2018.
- [28] Kiana Ehsani, Tanmay Gupta, Rose Hendrix, Jordi Salvador, Luca Weihs, Kuo-Hao Zeng, Kunal Pratap Singh, Yejin Kim, Winson Han, Alvaro Herrasti, et al. Spoc: Imitating shortest paths in simulation enables effective navigation and manipulation in the real world. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16238–16250, 2024.
- [29] Xinhao Liu, Jintong Li, Yicheng Jiang, Niranjana Sujay, Zhicheng Yang, Juexiao Zhang, John Abanes, Jing Zhang, and Chen Feng. Citywalker: Learning embodied urban navigation from web-scale videos. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 6875–6885, 2025.
- [30] Fernando Estévez Casado, Rodrigo Chacón Quesada, and Yiannis Demiris. Navigating uncertainty: Diffusion-based user intention estimation for wheelchair assistance. *IEEE Transactions on Robotics*, 42:80–97, 2025.
- [31] Samiran Gode, Abhijeet Nayak, Débora NP Oliveira, Michael Krawez, Cordelia Schmid, and Wolfram Burgard. Flownav: Combining flow matching and depth priors for efficient navigation. In *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 17762–17768. IEEE, 2025.
- [32] Maria Stamatopoulou, Jianwei Liu, and Dimitrios Kanoulas. Dippest: Diffusion-based path planner for synergistic trajectory generation applied on quadrupedal robots. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024. doi: 10.1109/IROS58592.2024.10802677. Acceptance Rate: 47.5%.
- [33] Jianwei Liu, Shirui Lyu, Denis Hadjivelichkov, Valerio Modugno, and Dimitrios Kanoulas. Vit-a*: Legged robot path planning using vision transformer a*. In *IEEE International Conference on Automation Science and Engineering (CASE) / Humanoids*, 2023. Acceptance Rate: 63%.
- [34] Roy Schwartz, Jesse Dodge, Noah A Smith, and Oren Etzioni. Green ai. *Communications of the ACM*, 63(12):54–63, 2020.
- [35] Xin Zhou, Zhepei Wang, Hongkai Ye, Chao Xu, and Fei Gao. Ego-planner: An esdf-free gradient-based local planner for quadrotors. *IEEE Robotics and Automation Letters*, 6(2):478–485, 2020.
- [36] Boyu Zhou, Fei Gao, Luqi Wang, Chuhao Liu, and Shaojie Shen. Robust and efficient quadrotor trajectory generation for fast autonomous flight. *IEEE Robotics and Automation Letters*, 4(4):3529–3536, 2019.
- [37] Ngoc Thinh Nguyen, Lars Schilling, Michael Sebastian Angern, Heiko Hamann, Floris Ernst, and Georg Schildbach. B-spline path planner for safe navigation of mobile robots. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 339–345. IEEE, 2021.
- [38] Boyu Zhou, Fei Gao, Jie Pan, and Shaojie Shen. Robust real-time uav replanning using guided gradient-based optimization and topological paths. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1208–1214. IEEE, 2020.
- [39] Yunfan Ren, Fangcheng Zhu, Guozheng Lu, Yixi Cai, Longji Yin, Fanze Kong, Jiarong Lin, Nan Chen, and Fu Zhang. Safety-assured high-speed navigation for mavs. *Science Robotics*, 10(98):eado6187, 2025.
- [40] Maria Stamatopoulou and Dimitrios Kanoulas. Efficiency-seeking, robot-specific path planning. In *International Conference on Robot Intelligence Technology and Applications (RiTA)*, 2025.
- [41] Lei Tai, Shaohua Li, and Ming Liu. A deep-network solution towards model-less obstacle avoidance. In *2016 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pages 2759–2764. IEEE, 2016.
- [42] Lei Tai, Jingwei Zhang, Ming Liu, and Wolfram Burgard. Socially compliant navigation through raw depth inputs with generative adversarial imitation learning. In *2018 IEEE international conference on robotics and automation (ICRA)*, pages 1111–1117. IEEE, 2018.
- [43] Cheng Chi, Zhenjia Xu, Siyuan Feng, Eric Cousineau, Yilun Du, Benjamin Burchfiel, Russ Tedrake, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion. *The International Journal of Robotics Research*, 44(10-11):1684–1704, 2025.
- [44] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. In *European conference on computer vision*, pages 630–645. Springer, 2016.
- [45] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin

- transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 10012–10022, 2021.
- [46] Kaihuai Qin. General matrix representations for b-splines. In *Proceedings Pacific Graphics’ 98. Sixth Pacific Conference on Computer Graphics and Applications (Cat. No. 98EX208)*, pages 37–43. IEEE, 1998.
- [47] Wenchao Ding, Wenliang Gao, Kaixuan Wang, and Shaojie Shen. An efficient b-spline-based kinodynamic replanning framework for quadrotors. *IEEE Transactions on Robotics*, 35(6):1287–1306, 2019.
- [48] Yi Zhou. A tutorial on uniform b-spline. *arXiv preprint arXiv:2309.15477*, 2023.
- [49] Bengt Fornberg and Julia Zuev. The runge phenomenon and spatially variable shape parameters in rbf interpolation. *Computers & Mathematics with Applications*, 54(3):379–398, 2007.
- [50] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.
- [51] Tim Salimans and Jonathan Ho. Progressive distillation for fast sampling of diffusion models. In *International Conference on Learning Representations*, 2022. URL <https://openreview.net/forum?id=TIIdIXIpzh0I>.
- [52] Marcel Torne, Andy Tang, Yuejiang Liu, and Chelsea Finn. Learning long-context diffusion policies via past-token prediction. *arXiv preprint arXiv:2505.09561*, 2025.
- [53] Yunhao Luo, Chen Sun, Joshua B Tenenbaum, and Yilun Du. Potential based diffusion motion planning. *arXiv preprint arXiv:2407.06169*, 2024.
- [54] Huang Huang, Balakumar Sundaralingam, Arsalan Mousavian, Adithyavairavan Murali, Ken Goldberg, and Dieter Fox. Diffusionseeder: Seeding motion optimization with diffusion for rapid motion planning. In *Conference on Robot Learning*, pages 4392–4409. PMLR, 2025.
- [55] Angel Chang, Angela Dai, Thomas Funkhouser, Maciej Halber, Matthias Niebner, Manolis Savva, Shuran Song, Andy Zeng, and Yinda Zhang. Matterport3d: Learning from rgb-d data in indoor environments. In *2017 International Conference on 3D Vision (3DV)*, pages 667–676, 2017. doi: 10.1109/3DV.2017.00081.
- [56] Cheng Lu, Yuhao Zhou, Fan Bao, Jianfei Chen, Chongxuan Li, and Jun Zhu. Dpm-solver++: Fast solver for guided sampling of diffusion probabilistic models. *Machine Intelligence Research*, pages 1–22, 2025.
- [57] Min Sung Ahn, Hosik Chae, Donghun Noh, Hyunwoo Nam, and Dennis Hong. Analysis and noise modeling of the intel realsense d435 for mobile robots. In *2019 16th International Conference on Ubiquitous Robots (UR)*, pages 707–711. IEEE, 2019.
- [58] Peter Anderson, Angel Chang, Devendra Singh Chaplot, Alexey Dosovitskiy, Saurabh Gupta, Vladlen Koltun, Jana Kosecka, Jitendra Malik, Roozbeh Mottaghi, Manolis Savva, et al. On evaluation of embodied navigation agents. *arXiv preprint arXiv:1807.06757*, 2018.
- [59] Erik Wijmans, Abhishek Kadian, Ari Morcos, Stefan Lee, Irfan Essa, Devi Parikh, Manolis Savva, and Dhruv Batra. Dd-ppo: Learning near-perfect pointgoal navigators from 2.5 billion frames. In *International Conference on Learning Representations*, 2019.
- [60] Haotong Lin, Sili Chen, Junhao Liew, Donny Y Chen, Zhenyu Li, Guang Shi, Jiashi Feng, and Bingyi Kang. Depth anything 3: Recovering the visual space from any views. *arXiv preprint arXiv:2511.10647*, 2025.
- [61] Alexey Bochkovskiy, Amaël Delaunoy, Hugo Germain, Marcel Santos, Yichao Zhou, Stephan Richter, and Vladlen Koltun. Depth pro: Sharp monocular metric depth in less than a second. In *The Thirteenth International Conference on Learning Representations*, 2025.