CORRECTED 9 JANUARY 2026; SEE LAST PAGE
SCIENCE ROBOTICS | RESEARCH ARTICLE

ROBOT LOCALIZATION

Resilient odometry via hierarchical adaptation

Shibo Zhao[1]*, Sifan Zhou[1], Yuchen Zhang[1], Ji Zhang[1], Chen Wang[2], Wenshan Wang[1]†, Sebastian Scherer[1]†

Resilient and robust odometry is crucial for autonomous systems operating in complex and dynamic environments. Existing odometry systems often struggle with severe sensory degradations and extreme conditions such as smoke, sandstorms, snow, or low-light conditions, threatening both the safety and functionality of robots. To address these challenges, we present Super Odometry, a sensor fusion framework that dynamically adapts to varying levels of environmental degradation. Super Odometry uses a hierarchical structure to integrate four core modules from lower-level to higher-level adaptability, including adaptive feature selection, adaptive state direction selection, adaptive engine selection, and a learning-based inertial odometry. The inertial odometry, trained on more than 100 hours of heterogeneous robotic platforms, captures comprehensive motion dynamics. Super Odometry elevates the inertial measurement unit to equal importance with camera and light detection and ranging (LiDAR) systems in the sensor fusion framework, providing a reliable fallback when exteroceptive sensors fail. Super Odometry has been validated across 200 kilometers and 800 operational hours on a fleet of aerial, wheeled, and legged robots and under diverse sensor configurations, environmental degradation, and aggressive motion profiles. It marks an important step toward safe and long-term robotic autonomy in all-degraded environments.

INTRODUCTION

Odometry is an important technique to estimate the position and orientation of robots over time, while also allowing for the three-dimensional (3D) geometry reconstruction of surrounding environments. It plays a crucial role in robotics, enabling spatial understanding and serving as a foundation for both high-level tasks such as navigation and exploration and low-level functions like path planning and control (*1*). As a result, odometry systems are widely used in robotic applications, including off-road driving (*2*), search-and-rescue (*3*), and planetary exploration (*4*).

However, maintaining robust and resilient odometry performance in real-world applications remains challenging because of a range of environmental degradations, such as poor lighting, geometrical degradation, motion blur from aggressive maneuvers, and extreme weather. These conditions make it difficult for odometry systems to extract and track reliable features from sensors over time, often leading to incorrect data associations (*5*) and eventual drift or failure.

To overcome these environmental degradations, researchers have explored multimodal sensor fusion methods, implemented as either loosely coupled (*6*–*11*) or tightly coupled (*12*–*15*) frameworks. Although these approaches have substantially advanced the field, none can fully handle various forms of environmental degradation, particularly when visual and geometric signals degrade simultaneously (*16*). This shortcoming continues to hinder the development of robust and resilient autonomous navigation systems capable of long-term operation in diverse environments. The key challenges can be summarized as follows.

Resilience under severe and persistent sensor degradation

Existing odometry solutions face substantial challenges when multiple sensors experience degradation or remain off-nominal over extended periods. For example, persistent smoke in wildfire environments can affect both visual and light detection and ranging (LiDAR) sensors, substantially hindering navigation (*17*).

Balancing robustness and efficiency

Although many approaches improve robustness by integrating data from all available sensors to ensure redundancy and safety, they often introduce high computational costs because of complex fusion strategies (*5*). Achieving robustness without sacrificing efficiency remains an open challenge.

Generalization across varied degradation scenarios

Most methods need substantial tuning for different types of environments because they are built on a "rigid" framework with a "fixed" parameter setting (*5*). For example, a method optimized for outdoor conditions may fail in indoor environments such as subterranean environments. This lack of generalization limits these methods' ability to generalize across diverse degradation conditions.

To address the first challenge of resilience, we drew inspiration from the human sensing system. Humans navigate degraded environments such as dense fog or darkness by integrating external perception with internal cues. When vision becomes unreliable, humans rely on vestibular and proprioceptive feedback to estimate motion, a process known as path integration (*18*). This allows them to maintain robust motion tracking even in the absence of external observation. Followed by this insight, we introduced a complementary internal sensing mechanism into robotic systems: a learned inertial module that infers motion from inertial measurement unit (IMU) data alone. Specifically, we developed a deep inertial network trained on diverse motion patterns to serve as an internal motion prior when external observations become unreliable.

To integrate this internal sensing into the odometry pipeline, we propose a reciprocal fusion strategy that combines traditional model-based odometry with the learned inertial module based on imperative learning (*19*). Instead of operating independently, these two components learn from each other: Under nominal conditions, the traditional estimator refines the inertial network using accurate pose estimates, allowing it to adapt online. As environmental degradation worsens, the learned IMU network, having captured the system's motion dynamics, takes over estimation and ensures continued robustness. In this way, robustness becomes adaptive, evolving with the robot's operating conditions.

[1]Carnegie Mellon University, Pittsburgh, PA, USA. [2]University at Buffalo, Buffalo, NY, USA.
*Corresponding author. Email: shiboz@andrew.cmu.edu
†These authors contributed equally to this work.

To address the second and third challenges, efficiency and generalization, we introduce a hierarchical adaptation framework that dynamically adjusts its behavior on the basis of the severity of environmental degradation. The framework operates through a multilevel scheme: Lower levels provide rapid and resource-efficient adaptations to address mild disturbances. If these measures are insufficient, higher levels provide more complex and computationally intensive interventions to support state estimation recovery. This layered design enables the system to maintain efficiency under nominal conditions and robustness under extreme scenarios, adapting both its computational load and sensing strategy to meet the demands of diverse environments.

In summary, we present Super Odometry (Movie 1), a sensor fusion framework designed for resilient, efficient, and generalizable state estimation under diverse and extreme sensor degradations. Our key contributions are as follows.

Hierarchical adaptation mechanism

We propose a hierarchical adaptive framework to address degradation along a spectrum from mild to severe and scale its response to maintain efficiency and robustness. It is dynamically reconfigurable, functioning as a multilevel scheme. Specifically, the system initiates with adaptive feature selection to mitigate mild visual degradation, then transitions to adaptive state direction for moderate geometric degradation, proceeds to adaptive factor (engine) selection for mixed degradation, and ultimately relies on a learning-based inertial odometry module when suffering from complete degradation, as illustrated in Fig. 1.

Heterogeneous learning–based inertial odometry

Inspired by human path integration (*18*), we introduce a heterogeneous learning–based IMU odometry model to achieve internal motion estimation and predict 3D poses in real time. Trained on hundreds of hours of diverse robotic data, it achieved strong generalization and maintained low-drift accuracy.

Learning and adaptation over time

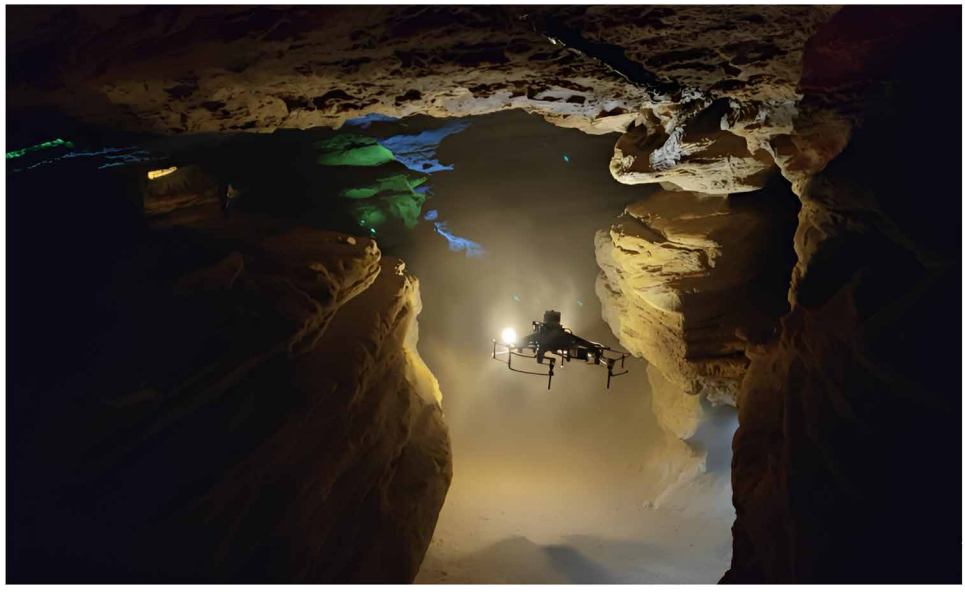
We developed a self-supervised online adaptation algorithm for the data-driven inertial odometry using self-supervised imperative learning (*19*). We formulated the system as a bilevel optimization: The IMU network learns motion patterns from free pose labels generated by a lower-level factor graph (*20*), whereas the upper-level inertial network provides motion priors back to the graph. This closed loop enables mutual correction between the learning-based and traditional estimators. We elevate IMUs to an independent role in sensor fusion, positioning them as equally important as cameras and LiDAR systems.

Extensive evaluation

Super Odometry was rigorously validated across over 200 km and 800 operational hours over 6 years on aerial, wheeled, legged, and handheld robots under diverse and extreme degradation, consistently outperforming state-of-the-art techniques. We believe Super Odometry represents a notable step toward a resilient and robust sensor fusion solution capable of operating reliably anytime, anywhere, and in any degraded environment. We expect this advancement to enhance long-term robotic autonomy and ensure the safety of robots. For more details on motivation and methods, refer to movie S1.

RESULTS

Real-world environments are often complex, involving multiple types of degradation ranging from mild to extreme. Therefore, an odometry solution should be adaptable and adjustable.



Movie 1. Summary of Super Odometry. This video introduces the core idea behind our resilient odometry framework and results. By elevating inertial sensing to a central role in state estimation and enabling it to self-evolve, robots gain an inner sense of motion that forms the foundation for robust odometry in extreme environments. The video presents diverse robot navigation results powered by the Super Odometry pipeline, highlighting strong adaptation to various forms of degradation and broad generalization across platforms. It also includes a single-run, 13-type degradation test covering all major failure modes. Together, these results point to a new direction for resilient SLAM, one that relies more on inner sensing rather than external perception alone.

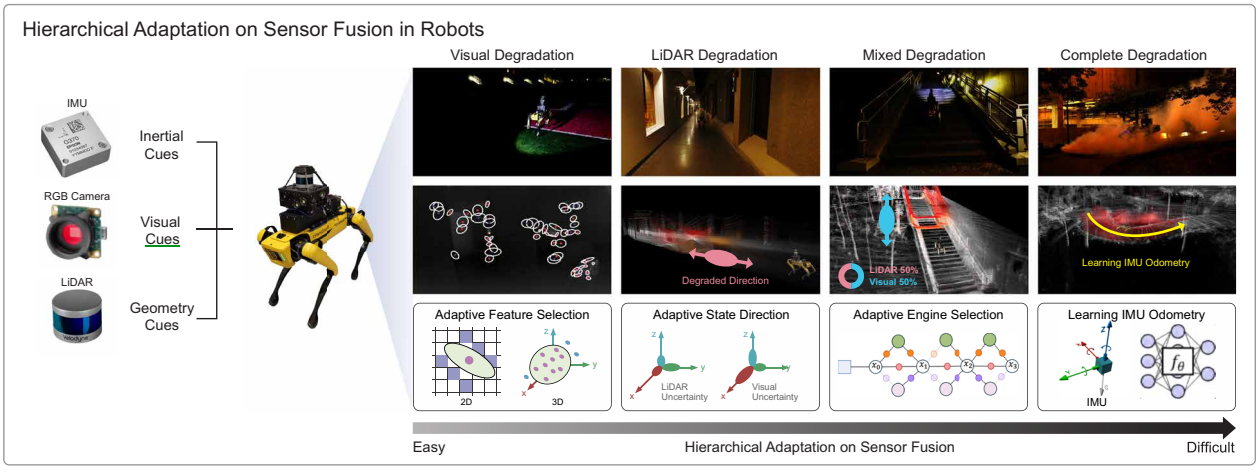


Fig. 1. Hierarchical adaptation on sensor fusion in robots. We propose a hierarchical adaptation strategy that initiates with adaptive feature selection for visual degradation, moves to adaptive state direction for LiDAR degradation, applies adaptive factor (engine) selection for mixed degradation, and uses learning-based IMU odometry during complete degradation.

We categorize degradation into four types: Visual degradation refers to adverse visual conditions (e.g., low light, motion blur, and darkness) that impair feature extraction from cameras. Geometric degradation arises in environments lacking meaningful 3D structure (e.g., open spaces and long corridors). Mixed degradation occurs when visual and geometric cues degrade intermittently. Complete degradation represents extreme scenarios where both visual and LiDAR sensors are off nominal, such as in dense smoke.

For visual degradation, our approach relies on adaptive feature selection to reject outliers. In geometric and mixed degradation, it actively selects the state direction and switches between sensor engines to ensure that the most reliable measurements are used. When facing extreme degradation such as complete degradation, our approach adopts a learning-based inertial odometry to overcome off-nominal sensor conditions such as smoke, dust, and snow.

In this section, we describe how our method performs across various degraded conditions. First, the method was evaluated through a comprehensive experiment, testing it under different degradation conditions in one run. Second, to illustrate the functionality of the method, we begin with visual degradation as an example, highlighting the role of adaptive feature selection. Next, we used geometrically degraded environments to demonstrate adaptive state selection and mixed degraded environments for sensor engine switching. Last, we explain how the learning-based IMU estimator helps overcome complete degradation in extreme smoke environments.

Comprehensive degradation in one run

University campus environment

To thoroughly evaluate the robustness of our odometry system, we conducted a comprehensive degradation experiment on a legged robot in a single run. This experiment included visual, geometrical, mixed, and complete degradation scenarios. The test route, located on Carnegie Mellon University's campus, covered a distance of 2966 m and took ~46 min to complete. The route incorporates more than 13 complex degradation scenarios, involving various combinations of challenging conditions. These scenarios include geometrical degradations such as long, textureless corridors (Fig. 2D), narrow and steep multifloor environments (Fig. 2, E and F), and transitions between indoor and outdoor environments (Fig. 2L). In addition, we introduced visual degradations (Fig. 2, G to N) such as smoke, white walls, low lighting, darkness, glass corridors, featureless surfaces, and lens flare. Mixed degradations, like steep staircases in darkness (Fig. 2B) and expansive dark courts (Fig. 2C), were also tested. Last, we examined complete degradation, where dense smoke enveloped the robot (Fig. 2A).

Our odometry system successfully completed real-time state estimation without any failures, primarily using a LiDAR-inertial setup. The final end-point drift was only 0.2 m over a total distance of 2966 m, resulting in an exceptionally low drift rate of 0.006%. The complete mapping results are shown in Fig. 2, highlighting high-fidelity views of the map under selected degraded conditions. The mapping results indicate the high accuracy and robustness of our pose estimates. Notably, this performance was achieved without using any backend optimization techniques, such as loop closure. For more details on results, refer to movie S2.

Adaptive feature selection for visual degradation

In this section, we primarily use visual degradation as an example to explain how adaptive feature selection works. We first conducted photometric calibration on red-green-blue (RGB) images by adjusting exposure time (*21*), which ensures that the images are consistent over time regardless of changes in lighting conditions. Then, we leveraged our previously proposed method ThermalPoint (*22*) to achieve robust feature detection on RGB images. Unlike existing methods tracking all features passively (*23*, *24*), we took into account the quality of feature correspondences and estimated their covariance for each 2D feature. By doing this, our method can robustly select the most informative features in very challenging scenarios such as textureless corridors, low lighting conditions, and even air obscurants shown in Fig. 3.

Adaptive state direction for geometric degradation

In this section, we use long corridors to illustrate the adaptive state direction mechanism. Long corridors, with repetitive and symmetrical features such as uniform walls lacking distinct geometry, present substantial challenges for LiDAR systems (*25*). The similarity of

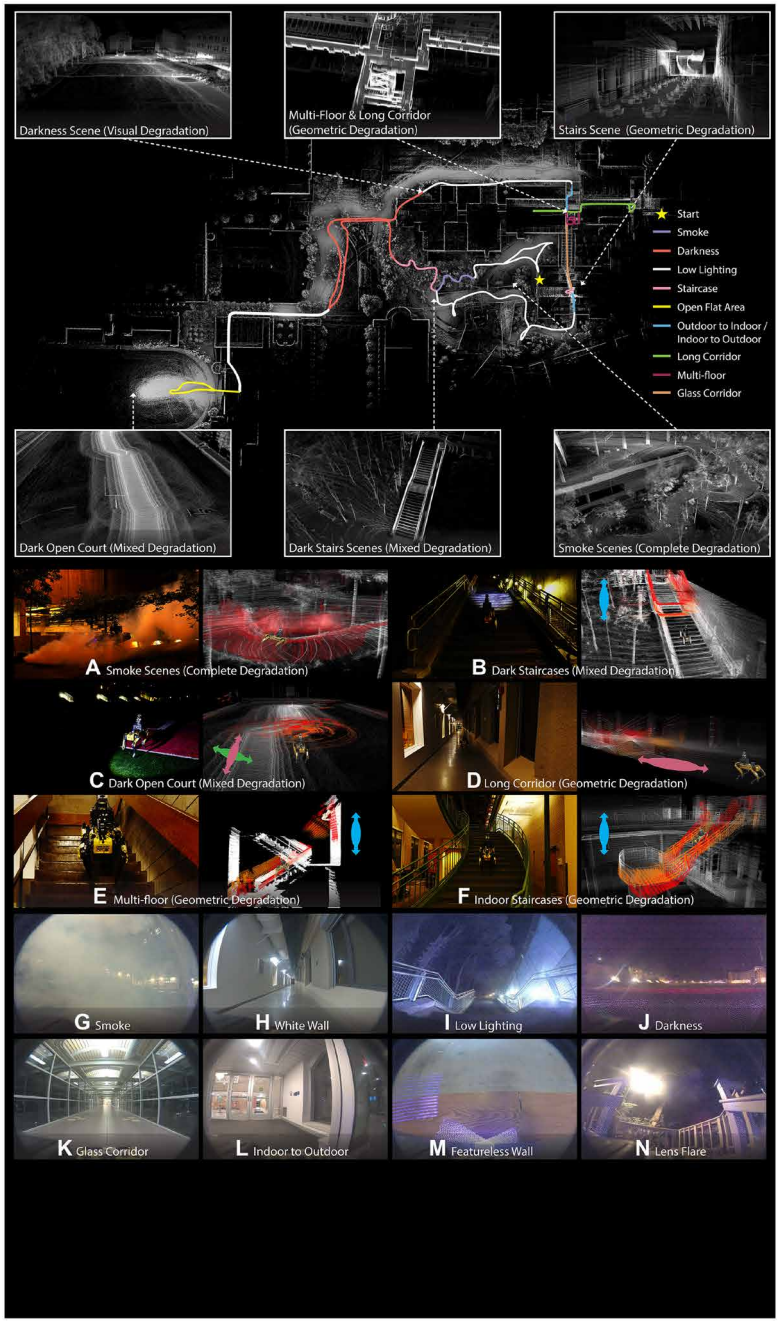


Fig. 2. Evaluation of 13 types of degradation in a single run. (**A** to **N**) The color-coded trajectory depicts our estimated odometry of a legged robot navigating through more than 13 complex degradation scenarios. Despite these difficulties, the final end-point drift was only 20 cm over a total distance of 2966 m.

these features across different positions makes iterative closest point (ICP) registration (*26*) poorly constrained in the forward direction.

The experiment involved navigating a completely dark, structureless corridor with highly reflective walls (Fig. 4A). Despite these difficulties, we successfully completed the testing route of 668 m. The vehicle operated aggressively, maintaining a speed of 4 m/s for up to 90% of the route and reaching a maximum speed of 5 m/s (Fig. 4B).

To achieve this, we developed an adaptive state direction method, which actively predicted the degeneracy direction of the optimization and incorporated other pose priors (visual) to mitigate the ill-optimization problem before failure. Figure 4C shows the confidence metric of LiDAR odometry reliability in the X, Y, and Z directions in the long corridor environments. It was observed that the confidence in the X direction, corresponding to the robot-centered forward direction, was relatively low compared with those in the Y and Z directions during the run.

Figure 4D illustrates the relative confidence levels of visual and LiDAR odometry in the degeneracy direction. Our method assigned lower confidence to LiDAR odometry in the forward direction,

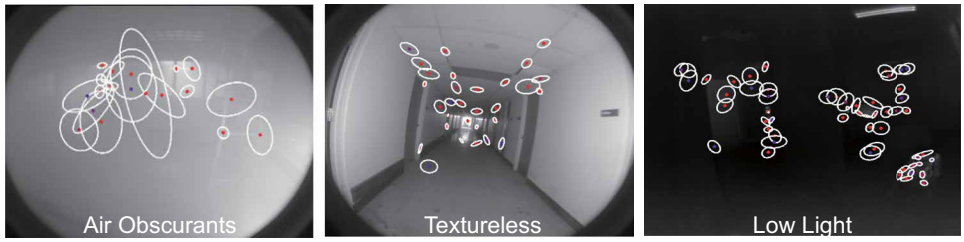


Fig. 3. Robust feature selection in low light, low texture, and obscured situations. The ellipses represent the uncertainty of 2D feature landmarks; a larger ellipse indicates greater uncertainty. The ellipse orientation reflects the ambiguity in feature tracking along a specific 2D direction.

whereas visual odometry maintained higher confidence. The weights for LiDAR odometry in the *Y* and *Z* directions remained unchanged, given that there was no degradation in those axes. In addition, our method outperformed state-of-the-art methods like FAST-LIVO (*27*) in achieving reliable scan registration in poorly constrained long-corridor environments.

By analyzing the relative confidence of LiDAR and visual odometry, we color-coded the trajectory on the LiDAR map on the basis of each sensor's contribution, as shown in Fig. 4B. In the corridor's middle section, visual odometry was actively integrated into sensor fusion and addressed degeneracies, whereas at the end section, the system increasingly relied on LiDAR odometry. This process used soft fusion between LiDAR and visual odometry across each axis, providing additional constraints at optimal moments. For more results, refer to movie S3.

Adaptive engine selection for mixed degradation

Real-world environments present various challenges, where systems can experience visual or geometric degradation at different stages, classified as mixed degradation. A typical example is sensor-drop scenarios, where either visual or LiDAR data may be compromised and suffer from interruptions for extended periods. To solve this, we believe an odometry solution should be fail safe and failure aware. It should detect imminent failures like signal loss or degradation and provide recovery mechanisms.

We initially blocked the LiDAR's side view using the cardboard shown in Fig. 5A. Our algorithm, adopting an adaptive state-direction mechanism, promptly predicted the degradation in the *Y* direction and reduced the LiDAR confidence weight to 60% in that direction (Fig. 5B). Simultaneously, the visual odometry confidence weight was increased to 40% to compensate for the LiDAR's reduced performance. This process not only predicted the degraded direction but also adapted the necessary weights and constraints between the visual and LiDAR modality to prevent optimization failure.

Similarly, when the front and back views of the LiDAR were blocked, the LiDAR's confidence weight in the *X* direction dropped to 60%, whereas the weight of the visual odometry in that direction increased to 40%. Upon fully blocking the LiDAR with cardboard, its weight in all directions dropped to 10%, whereas the visual odometry's confidence weight rose to 90%. Last, when we blocked the camera, the LiDAR's confidence weight across all directions returned to greater than 90%. If a modality's contribution to the joint state falls below 10% and remains low for more than 2 to 4 s, the system will disable that modality. This threshold and temporal window were empirically chosen and tunable to ensure fail-safe and failure-aware operation. Figure 5C illustrates the odometry and mapping accuracy to evaluate the sensor dropout tolerance of our method. The largest absolute trajectory error (ATE) is only 0.184 m over a 600-s run.

Learning-based inertial odometry for complete degradation

In degraded environments like dense smoke, dust storms, and heavy snow, almost all the odometry algorithms (*25*, *27*, *28*) face severe challenges, because they encounter both visual and geometric degradation over an extended period. Visual sensors struggle to capture the high-quality images needed to extract reliable features. LiDAR systems, which depend on laser pulses for distance measurements, experience reduced range and accuracy because of air obscurants.

To address these limitations, we propose a learning-based IMU odometry, given that IMU is unaffected by perceptual degradation. Our solution is straightforward: When other sensors, such as visual or LiDAR, are functioning well, the IMU odometry receives feedback from them and fine-tunes its path integration capabilities (*18*). However, when these sensors are compromised, such as in extreme environments, the IMU odometry takes over state estimation.

IMU pretrained model on large dataset

Developing robust IMU odometry is challenging, especially in ensuring generalization across platforms. Most methods lack adaptability beyond specific settings (*29*). We addressed this by proposing a general IMU model trained on large, diverse datasets (*16*, *30*), covering drones, quadrupeds, cars, and humans. This broad training enabled it to outperform specific models (*31*–*34*) and adapt to various platforms. Our approach aimed at strong generalization, rapid adaptation, and seamless integration into sensor fusion pipelines. To validate these objectives, we investigated three core research questions: (i) Can our model generalize across platforms and outperform specialized models? (ii) Can it adapt online to previously unseen environments? (iii) How does it enhance sensor fusion robustness in extreme environments?

Generalization of IMU-pretrained model

To address (i), we focused on two critical factors for training a more generalized IMU model: a heterogeneous shared backbone and large, high-quality, diverse data. To demonstrate the importance of a heterogeneous shared backbone, we evaluated different IMU models on data from four platform types: aerial, wheeled, legged, and handheld. For a fair comparison, all models, including AI-IMU (*35*), RNIN-VIO (*33*), TLIO (*36*), and IMO (*29*), were trained on the same large datasets and tested on identical unseen data.

As shown in Fig. 6A, our IMU pretrained model, built with a heterogeneous shared backbone, demonstrated strong few-shot capabilities. When tested across the four robotic platforms, it consistently outperformed specialized models tailored for specific

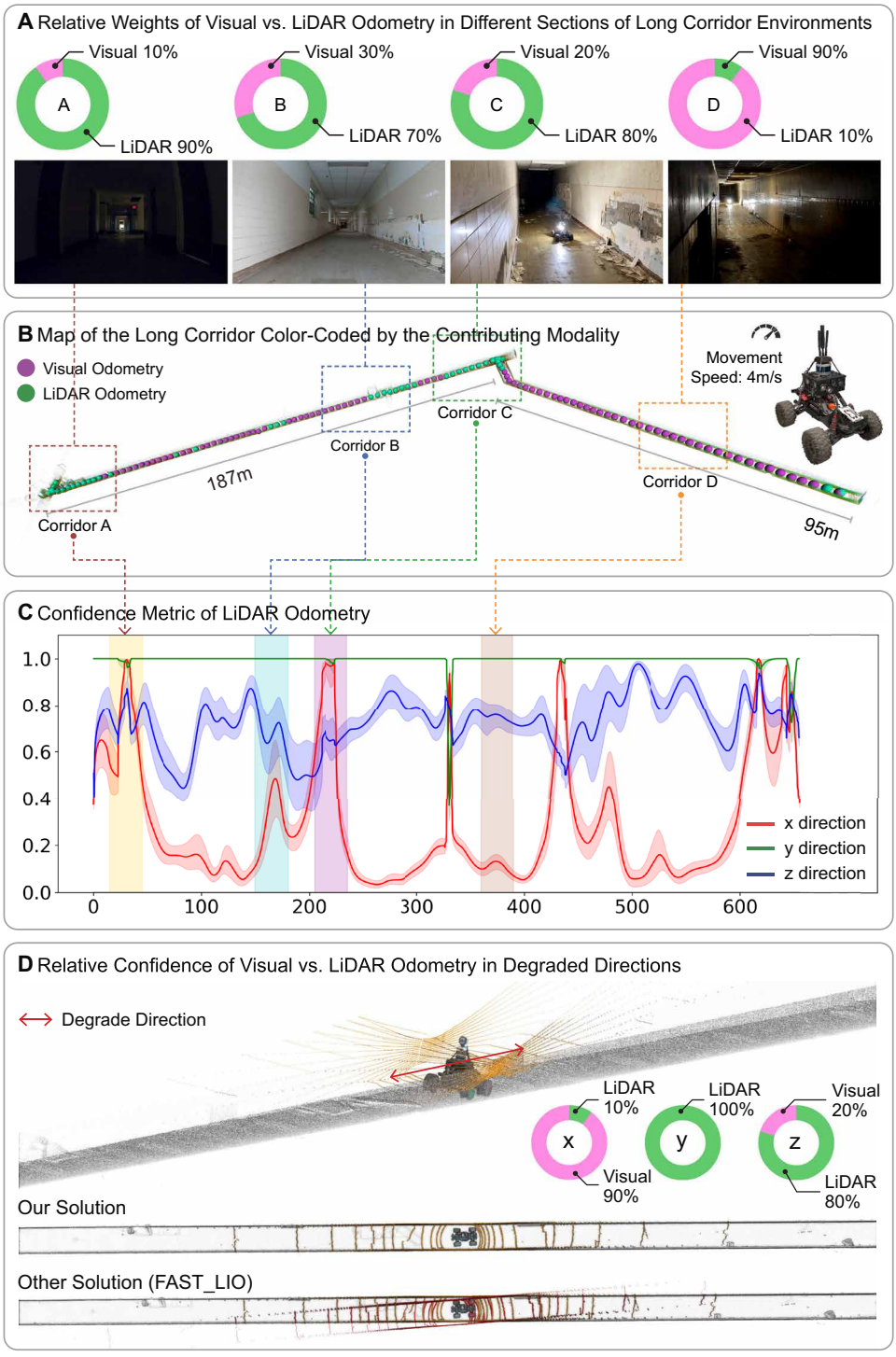


Fig. 4. Illustration of adaptive state direction selection in a long corridor from the abandoned Pittsburgh hospital. (**A**) Relative weights of visual and LiDAR odometry across different sections of long corridor environments. (**B**) Map of the long corridor, color-coded by the contributing modality; the purple circle indicates areas where visual odometry is actively integrated into sensor fusion, whereas the green circle denotes areas dominated by LiDAR odometry. Note: A modality is considered "dominant" if its weight in the sensor fusion process is greater than 50%. (**C**) Confidence metric of LiDAR odometry over a distance of 668 m. (**D**) Relative confidence of visual and LiDAR odometry in the degeneracy direction, demonstrating more reliable scan registration achieved with our method.

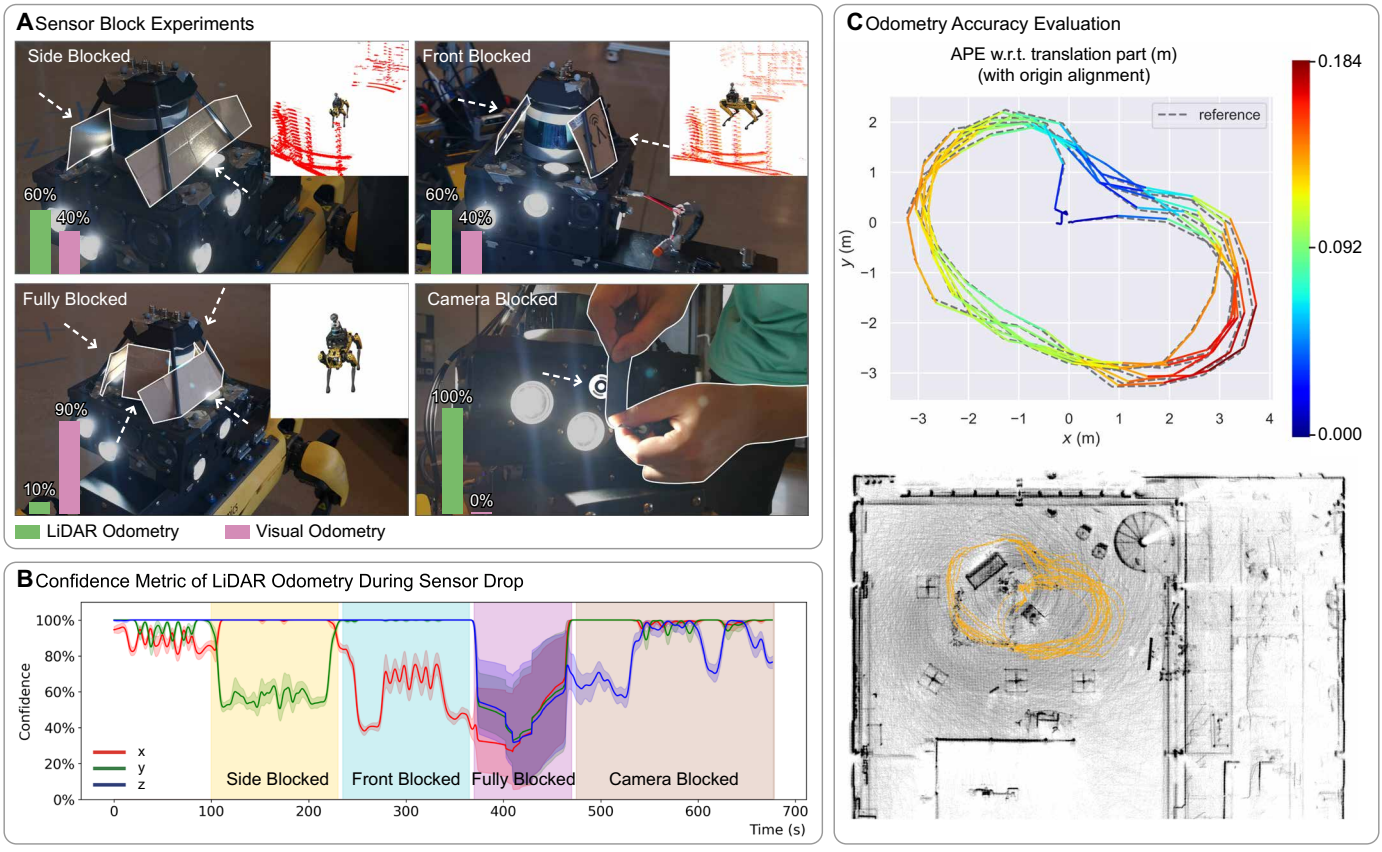


Fig. 5. Adaptive engine selection for mixed degradation in sensor-drop scenarios. (**A**) We used cardboard to sequentially block the LiDAR's side, front, and back views and eventually completely obstructed both the LiDAR and visual sensor to evaluate the dropout tolerance for 600 s. (**B**) The corresponding confidence metric of the LiDAR odometry system. (**C**) Odometry accuracy evaluation: The trajectory color encodes APE (absolute pose error) (*37*) (top), and the bottom shows mapping results with yellow indicating the estimated trajectory. w.r.t., with respect to.

systems. Without requiring fine-tuning, our model provided the most reliable and accurate odometry estimates across a variety of motion patterns (see Fig. 6A, last column). On average, it improved ATE and time-relative trajectory error (T-RTE) by 35.5 and 41.0%, respectively, compared with the second-best model (see Table 1). These results underscore the effectiveness of the heterogeneous shared backbone in achieving robust generalization across diverse robotic motion patterns.

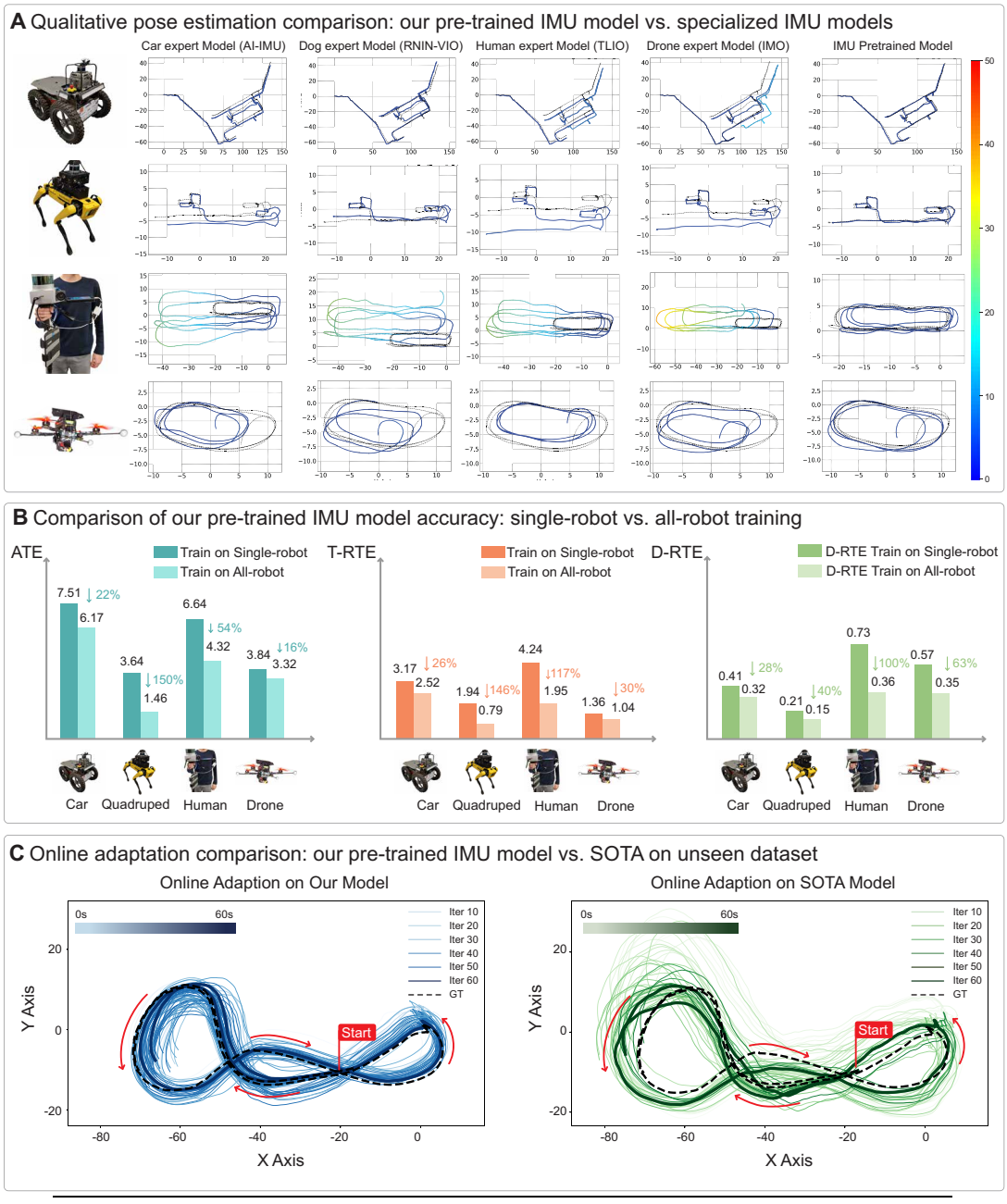
We also investigated the effect of data diversity on model performance. Specifically, we compared the performance of our IMU pretrained model when trained on data from a single platform versus data from various robot types. As illustrated in Fig. 6B, models trained on diverse datasets showed substantial performance improvements. For instance, in the "human handheld" sequence, a model trained on data from aerial, wheeled, legged, and quadruped platforms outperformed one trained exclusively on human handheld data. The ATE decreased by up to 54%, whereas T-RTE and distance-relative trajectory error (D-RTE) decreased by 117 and 100%, respectively. These findings confirm the critical role of large, high-quality, and diverse datasets in improving model generalization. Training on a wide variety of robot data enables the model to perform well even when there are substantial differences between the training and testing domains. For more details on results, refer to movie S4.

Fast online adaptation on test data from TartanDrive

Although our model was trained on a large dataset [as described in (i)], it still faces generalization issues when deployed on a robot platform with different motion patterns. Moreover, collecting sufficient data for fine-tuning during deployment is often impractical. To address (ii), we used online fine-tuning during deployment—we learn as we operate.

We assessed our model's ability to generalize to an unseen platform. The model was trained exclusively on the SubT dataset, which features a small ground vehicle with a maximum speed of 5 m/s. To evaluate its adaptability, we tested it on the TartanDrive dataset, which represents a full-size all-terrain vehicle with a maximum speed of 15 m/s. During the online fine-tuning process, the model was provided with only 60 s of IMU data, enabling it to adapt in real time from the SubT UGV (unmanned ground vehicle) model to the TartanDrive car model.

We compared our model against a state-of-the-art baseline model (*33*), trained on the same SubT dataset as ours, following an identical training process. In Fig. 6C, over a 60-s fine-tuning period, our model (blue lines) demonstrated substantially better odometry predictions, closely aligning with the ground truth (dashed line). In contrast, the state-of-the-art model (green lines), although showing some improvement, exhibited substantial trajectory errors at the 60-s



Timeline	0 s	10 s	20 s	30 s	40 s	50 s	60 s
SOTA Model ATE (m)	34.14	29.24	18.20	15.56	13.28	9.48	4.94
Our Model ATE (m)	**32.87**	**27.06**	**13.49**	**8.27**	**5.34**	**4.03**	**1.73**
↑ Improve (%)	3.72	8.06	25.88	46.86	59.79	57.49	64.98

Fig. 6. Evaluation on learning-based inertial odometry. (**A**) Qualitative pose comparison shows that our method (last column) achieves the highest accuracy against ground truth (dashed gray). (**B**) Accuracy improves significantly when the IMU pretrained model is trained on multirobot data versus single-robot data. (**C**) Online adaptation on unseen datasets demonstrates that our model outperforms SOTA (state of the art), as supported by a quantitative comparison of ATE over time. Each iteration shown corresponds to 1 s of adaptation time (e.g., iteration 10 represents 10 s). GT, ground truth.

Table 1. Quantitative pose comparison: Pretrained versus specialized IMU models. Our IMU pretrained model outperformed different specialized IMU models, achieving an average of 35.5% on ATE and an average of 41% on T-RTE, respectively, compared with the second-best model across various robotic platforms, including wheeled, legged, handheld, and aerial systems.

IMU model	Wheeled (car) (*16*)	Legged (quadruped) (*16*)	Handheld (human) (*58*)	Aerial (drone) (*63*)
		AI-IMU (*35*)		
ATE ↓	7.68	3.23	8.26	4.14
T-RTE ↓	3.33	1.6	4.89	1.45
		RNIN-VIO (*33*)		
ATE ↓	7.82	3.1	7.62	4.32
T-RTE ↓	5.06	1.58	5.61	1.51
		TLIO (*36*)		
ATE ↓	8.12	3.61	6.96	3.93
T-RTE ↓	3.73	1.73	4.82	1.4
		IMO (*29*)		
ATE ↓	8.12	3.35	10.19	3.72
T-RTE ↓	3.73	1.64	5.67	1.34
		IMU-pretrain (ours)		
ATE ↓	6.17	1.46	4.32	3.32
T-RTE ↓	2.52	0.79	1.95	1.04
		Improvement		
ATE ↓	↑ 19.63%	↑ 54.83%	↑ 47.69%	↑ 19.81%
T-RTE ↓	↑ 24.36%	↑ 50.63%	↑ 60.05%	↑ 28.97%

mark. This online adaptation scenario shows the potential of our pretrained IMU model, which can adapt to new domains more rapidly and converges with fewer iterations than the state-of-the-art methods, as detailed in the table from Fig. 6C.

Enhance odometry robustness in smoke environments

To answer (iii), we designed experiments where the robot is required to operate in a smoke-filled environment to demonstrate how learning-based IMU odometry can improve the robustness of odometry systems. It included two processes: self-supervised online adaptation [explained in (ii)] and learning-based IMU model deployment in the sensor fusion pipeline.

Self-supervised online adaptation. As discussed in (ii), Fig. 7C illustrates the self-supervised online adaptation process used to fine-tune our pretrained IMU model from the SubT vehicle to the TartanDrive vehicle. Initially, the TartanDrive vehicle operated in a smoke-free environment, where our factor graph pipeline continuously provided supervisory signals to refine the learning-based IMU odometry. This online training process required only 45 s of data collection to successfully evolve the pretrained SubT IMU model into the TartanDrive model.

Learning-based inertial odometry deployment. After fine-tuning the IMU model, we integrated the learning-based IMU odometry into our adaptive sensor fusion pipeline and evaluated its performance online using data from human-driven vehicles. We evaluated its performance in two challenging environments: an extremely smoke-filled area and a clear, open, flat area, as depicted in Fig. 7A. In these environments, LiDAR odometry became highly unreliable. When the vehicle entered the smoke or open flat areas, our adaptive pipeline automatically prioritized the learning-based IMU odometry solution (purple trajectory). This is because the learning-based IMU odometry remained unaffected by these environmental conditions and continued to provide reliable state estimation. Conversely, when the vehicle exited these areas and conditions improved, the pipeline seamlessly transitioned back to the LiDAR odometry solution (blue trajectory). This switching behavior was governed by the confidence level of the LiDAR odometry, as illustrated in Fig. 7D.

This transition was not a "hard switch" but a soft fusion approach that blended the outputs of both LiDAR and IMU odometry across each axis. This strategy mirrored the adaptive engine selection mechanism used in previous mixed degradation scenarios.

To validate our method, we conducted figure-eight driving tests, repeatedly transitioning between smoky and clear areas. The adaptive switching occurred multiple times, demonstrating the pipeline's responsiveness to environmental changes. The estimated trajectory and the corresponding map are shown in Fig. 7B. Although the reconstructed map exhibits severe noise because of the smoke, our sensor fusion pipeline consistently delivered robust state estimation.

This experiment demonstrated that incorporating learning-based IMU odometry significantly enhances the robustness of odometry systems in severely degraded environments. For more details on results, refer to movie S5.

Quantitative comparison with state-of-the-art methods

Accuracy evaluation

To further validate the precision of pose estimation, we conducted ATE (*37*) analysis using our odometry system on the SubT-MRS dataset (*16*). This dataset encompasses challenging environments featuring sensor degradation, aggressive locomotion, and extreme weather conditions. The eight sequences in the dataset are categorized into two groups for testing: geometric degradation and mixed degradation. ATE results of competing systems were sourced from an open odometry challenge (*16*). Participating teams were allowed

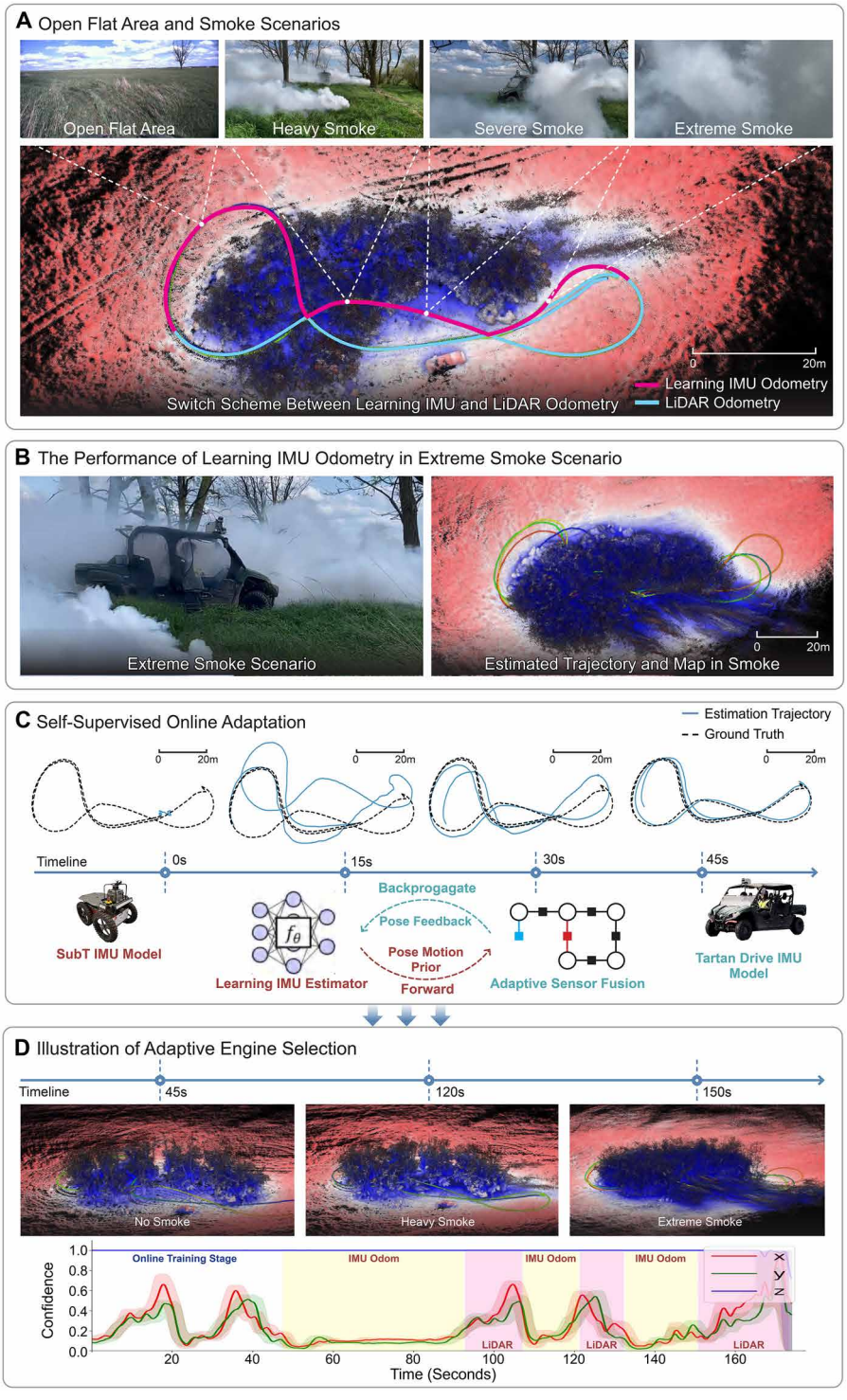


Fig. 7. Robust performance in a smoke scenario. (**A**) Adaptive engine selection between learning-based IMU and LiDAR odometry in challenging environments, with purple and blue trajectories indicating their respective use. (**B**) Trajectory and map generated using learning-based IMU odometry in heavy smoke. (**C**) Self-supervised adaptation fine-tunes the IMU model from SubT car to TartanDrive. (**D**) Confidence-driven adaptive fusion combines IMU and LiDAR odometry outputs using soft fusion, avoiding hard switches.

to use loop closure techniques and offline processing to achieve high accuracy. Despite this, our method achieved an average ATE of 0.271 in real time. This performance surpassed the second-best method by 54%, showing our high accuracy and robustness in Table 2.

Robustness evaluation

An ideal evaluation metric should assess both accuracy and completeness (*16*). Although the ATE effectively measures trajectory accuracy, it does not adequately capture trajectory completeness. To address this limitation, we used robustness metrics (*38*), which account for both aspects. Table 3 presents a detailed comparison of these metrics across various scenarios and algorithms, highlighting that our method demonstrates superior robustness compared with competing approaches.

This comparison excludes completely degraded or smoke-related experiments (as illustrated in Fig. 7). The exclusion was due to the

Table 2. ATE performance on SubT-MRS (*38*). – denotes incomplete trial.

	Liu* (*64, 65*)	Weitong* (*20, 66*)	Kim* (*25, 67*)	Yibin (*68*)	Zhong* (*69, 70*)	Zheng (*69*)	Ours
Geometric degradation							
F01	0.307	0.26	0.331	1.060	1.205	–	0.238
F02	0.095	0.096	0.092	0.220	0.695	–	0.074
F03	0.629	0.617	0.787	0.750	–	–	0.396
U01	0.122	0.120	0.123	0.470	1.175	–	0.026
U02	0.235	0.222	0.270	0.620	1.72	–	0.104
Mixed degradation							
Cave03	0.260	0.402	0.279	9.140	2.08	3.786	0.204
Corridor01	1.454	1.254	2.100	2.990	–	55.205	0.817
Floor01	0.401	0.577	0.650	5.500	–	19.769	0.315
Average	0.588	0.663	3.825	4.312	1.209	26.254	0.271

*Incorporation of loop closure.

Table 3. Robustness performance on SubT-MRS (*38*). – denotes incomplete trial.

	Weitong* (*20, 66*)	Liu* (*64, 65*)	Kim* (*25, 67*)	Yibin (*68*)	Zhong* (*69, 70*)	Zheng (*69*)	Ours
R_p: *Geometric degradation (SubT-MRS)*							
Final01	0.922	0.920	0.884	0.849	0.278	–	0.929
Final02	0.929	0.928	0.929	0.897	0.91	–	0.933
Final03	0.906	0.904	0.906	0.827	0.827	–	0.916
Urban01	0.933	0.931	0.933	0.875	0.905	–	0.937
Urban02	0.919	0.917	0.915	0.795	0.877	–	0.925
R_p: *Mix degradation*							
Cave03	0.83	0.832	0.83	0.751	–	0.618	0.973
Corridor01	0.889	0.885	0.89	0.502	–	0.579	0.899
Floor01	0.909	0.910	0.252	0.738	–	0.389	0.891
R_r: *Geometric degradation (SubT-MRS)*							
Final01	0.931	0.930	0.923	0.924	0.296	–	0.938
Final02	0.931	0.930	0.931	0.928	0.921	–	0.942
Final03	0.907	0.908	0.908	0.898	0.851	–	0.925
Urban01	0.938	0.937	0.938	0.933	0.922	–	0.943
Urban02	0.924	0.922	0.92	0.914	0.905	–	0.941
R_r: *Mix degradation*							
Cave03	0.848	0.848	0.848	0.846	–	0.629	0.964
Corridor01	0.89	0.885	0.89	0.874	–	0.882	0.944
Floor01	0.914	0.908	0.26	0.886	–	0.489	0.924

*Loop closure usage.

fact that all other algorithms fail under such conditions. For details on the robustness metric, refer to Materials and Methods.

DISCUSSION

In this section, we provide insights into developing a robust odometry for degraded environments.

Hierarchical adaptation is a key factor for resilience

State estimation in challenging environments demands not only sensor redundancy but also computational efficiency. However, most existing odometry frameworks rely on rigid multimodal fusion strategies that prioritize robustness by incorporating additional sensors yet still fail to generalize across diverse degradation scenarios (*17*, *39*).

This limitation arises from two common assumptions: Measurement noise follows a Gaussian distribution, and averaging redundant measurements yields accurate estimates. However, in practice, these assumptions break down in degraded environments such as smoke or darkness, where sensor signals deviate from expected distributions and corrupted measurements dominate certain directions of the state space. As a result, state estimation becomes unreliable, and optimization grows computationally expensive because of the increasing difficulty of distinguishing inliers from outliers. This highlights a critical need for foundational research on resilient systems.

To address this, we propose a hierarchical adaptation framework for resilient odometry. Our approach dynamically reconfigures the system on the basis of degradation severity, starting with lightweight adaptations and escalating to more complex ones when needed. This multilevel strategy enables reliable state estimation across diverse degradation types, including visual, geometric, mixed, and full-sensor failures (see Fig. 2).

Low-cost IMUs are more useful than expected

Historically, IMUs have been treated as auxiliary or secondary sensors in odometry pipelines, typically limited to preintegration (*40*). However, we argue that IMUs hold far greater potential and can serve as independent sources for state estimation, contributing as meaningfully as LiDAR and visual sensors.

We introduce a reciprocal fusion strategy in which learning-based IMU odometry and traditional model-based sensor fusion continuously learn from each other to enhance overall performance. Our approach enables the IMU model to adapt and improve using feedback from the sensor fusion pipeline while simultaneously providing motion priors and acting as a reliable fallback solution under degraded conditions. This bidirectional adaptation leads to improved accuracy and robustness across diverse environments.

Improving sensor calibration and time synchronization

Currently, the performance of our system depends heavily on accurate calibration and precise time synchronization. Implementing online calibration and time-sync techniques (*41*) could eliminate the need for manual parameter tuning. A promising direction is to leverage continuous-time batch optimization (*42*) combined with observability-aware calibration (*41*), which actively selects informative motion segments to align sensor trajectories. This would enable the joint estimation of intrinsic parameters, extrinsic transformations, and temporal offsets in a unified framework.

Improving generalizability of the IMU model

The learning-based IMU model requires faster adaptation to new robots and environments. Despite generalizing well across platforms, it struggles with unseen domains because of distribution gaps between training and testing data. Incorporating both real-world and simulated IMU data could reduce this gap and improve generalization.

In conclusion, the proposed hierarchical adaptation system offers robust, adaptable, and efficient state estimation in degraded environments. Although challenges remain, such as time synchronization, calibration, and improvements to the generalization of learning IMU odometry, this work lays a strong foundation for advancing resilient odometry under real-world conditions.

MATERIALS AND METHODS

Before outlining our methodology, we first clarify the key technical terms used in this paper. We propose hierarchical multilevel mechanisms to address environmental degradation, ranging from mild to extreme. These mechanisms include adaptive feature selection, adaptive state direction, adaptive engine selection, and learning-based inertial odometry. For clarity, we group adaptive feature selection, state direction, and engine selection under the term adaptive sensor fusion, because they were based on classic factor graph optimization. We treat learning-based inertial odometry separately, because it relied on a deep learning network. Together, these components formed what we defined as a hierarchical adaptation framework.

Problem definition

State estimation definition

Following the SLAM (simultaneous localization and mapping) formulation (*1*, *5*), the general goal of state estimation is to infer the agent's state x_t, and optionally the map m_t, from a sequence of control inputs $u_{1:t}$ and observations $o_{1:t}$, via the posterior

$$P(x_t, m_t | o_{1:t}, u_{1:t}) \tag{1}$$

In contrast with SLAM, the odometry problem focuses solely on estimating the agent's trajectory over time, without building a persistent map. The objective is to estimate the current or next pose x_{t+1} given a sequence of past observations $o_{1:t}$ and control inputs $u_{1:t}$. Odometry aims to compute

$$P(x_{t+1} | o_{1:t}, u_{1:t}) \tag{2}$$

Assuming a probabilistic motion model $P(x_{t+1}|x_t, u_{t+1})$ and an observation model $P(o_{t+1}|x_{t+1})$, the belief over the current state is recursively updated using Bayes' rule

$$P(x_{t+1}|o_{1:t+1}, u_{1:t+1}) = \eta \cdot P(o_{t+1}|x_{t+1}) \int P(x_{t+1}|x_t, u_{t+1}) P(x_t|o_{1:t}, u_{1:t}) dx_t \tag{3}$$

where η is a normalization constant. On the basis of the above, the odometry problem is incrementally estimating the next pose x_{t+1} by given local observations and short-term motion models.

Robustness definition

Existing evaluation metrics such as ATE (*43*) have limitations in evaluating the odometry's robustness in real-world applications. ATE primarily focused on the overall trajectory error and cannot effectively capture the robustness of the algorithm. To evaluate robustness, we used the relative pose between frames as robustness

metrics (*38*). Unlike traditional relative pose error (RPE) evaluations, our metric assesses both the local accuracy of the trajectory and its completeness across the entire trajectory. The robustness metric is the area under the curve (AUC) of the F_1 score

$$F_1(e) = \frac{2P(e<T)R(e<T)}{P(e<T)+R(e<T)} \tag{4}$$

where the precision P quantifies the precision of the relative pose, namely, how closely the estimated relative pose matches the ground truth point and the recall rate R measures the completeness, and to what extent the estimated relative pose covers all ground truth points. Specifically, an estimated error e is regarded as an inlier if it is smaller than a threshold T. To evaluate robustness across different tolerance levels, we applied an exponential mapping $\exp(-10T)$ to normalize the threshold range to [0, 1]. We then computed the AUC of the F_1 score by sweeping across this range of thresholds. This provides a comprehensive, threshold-independent measure of overall robustness.

Position robustness R_p and rotation robustness R_r

The AUC of the F_1 score can be defined using both linear velocity and angular velocity, which can reflect the robustness of position and rotation estimation, respectively. Specifically, we define $R_p = \mathrm{AUC}\left(F_1\left(v_e\right)\right)$ and $R_r = \mathrm{AUC}\left(F_1\left(\omega_e\right)\right)$, where v_e and ω_e are the estimated error of relative position and orientation between frames, respectively.

Overview of approach

The core of our method was a hierarchical adaptation framework that integrated adaptive sensor fusion with learning-based IMU odometry in a self-supervised manner, as shown in Fig. 8A. This framework was structured as a bilevel optimization (*19*), where adaptive sensor fusion is in the lower level (a "teacher"), providing optimization-based constraints for the IMU odometry in the upper level (a "student"). When vision or LiDAR functions well, IMU odometry receives feedback to learn robot motion. When vision or LiDAR degrades, the IMU module takes over to maintain reliable estimation. Together, this cooperative symbiosis substantially enhances the robustness, resilience, and adaptability of odometry systems in challenging environments. For details, see the "System overview" section in the Supplementary Materials.

Adaptive sensor fusion

We proposed adaptive sensor fusion that hierarchically selects information from parameters, features, states, and engines across multiple levels. It included adaptive feature selection, adaptive state selection, adaptive engine selection, and adaptive parameter tuning.

Adaptive feature selection

As shown in Fig. 8B, step 1, not all features (visual or geometric) contributed equally to state estimation. Thus, selecting the most informative ones is essential. For visual features, we used the Hessian matrix of KLT (Kanade-Lucas-Tomasi tracking) tracking (*44*) and analyzed its covariance $\Sigma_{\mathrm{SE}(2)}$ using small patch sizes (*45*) from intensity residuals. Given the estimated $\boldsymbol{R}_0 \in \mathrm{SO}(2)$, the rotated covariance for each 2D feature is

$$\Sigma_{2\mathrm{D},\theta} = \boldsymbol{R}_\theta \Sigma_{\mathrm{SE}(2)} \boldsymbol{R}_\theta^{\top} \tag{5}$$

For LiDAR, we evaluated the quality of correspondences on the basis of the fit of neighboring points to point, line, or plane distributions (*46*)

$$w_i^{\mathrm{po}\to\mathrm{li}} = \frac{\sigma_1-\sigma_2}{\mu}, w_i^{\mathrm{po}\to\mathrm{pl}} = \frac{\sigma_2-\sigma_3}{\mu}, w_i^{\mathrm{po}\to\mathrm{po}} = \frac{\sigma_3}{\mu} \tag{6}$$

where σ_i are the eigenvalues of the principal components analysis (PCA) used for the point-to-point, point-to-line, point-to-plane distance error metric (*38*). The normalization coefficient $\mu = \sum_{i=1}^{3}\sigma_i$ is used, and $W_l = \left\{ w_i^{\mathrm{po}\to\mathrm{po}}, w_i^{\mathrm{po}\to\mathrm{li}}, w_i^{\mathrm{po}\to\mathrm{pl}} \right\} \in [0,1]$ describes the linearity $\left(w_i^{\mathrm{po}\to\mathrm{li}}\right)$, planarity $\left(w_i^{\mathrm{po}\to\mathrm{pl}}\right)$, and scatter $\left(w_i^{\mathrm{po}\to\mathrm{po}}\right)$ of the features. This approach enables the selection of the meaningful features while reducing the computational load.

Adaptive state direction

Various forms of degradation hinder the robustness of existing odometry systems (*25*, *47*). However, existing approaches require prior knowledge of the environment to determine well-conditioned state estimation and cannot adapt very well to changing environments (*48*–*51*). Moreover, some methods perform degeneracy analysis (*52*) or estimate a robust kernel function (*53*) during or after the optimization. However, this may be too late to predict potential failures (*48*), and insufficient data association already causes unreliable optimization.

To address potential failures, we performed a degeneracy analysis by evaluating the observability of LiDAR odometry at the front end (Fig. 8C, step 2). We focused on LiDAR rather than vision, because it offers more reliable observability and uncertainty estimates for each state direction. Once LiDAR quantifies directional uncertainty, visual weights can be adjusted accordingly. Specifically, we predicted ICP alignment risk and computed the observability **Obs***i* for each point along each state direction. Each correspondence was then labeled on the basis of the state it primarily influences. Due to space constraints, refer to the Supplementary Materials for details on alignment risk prediction and observability for each point-to-plane constraint.

To evaluate observability over a scan, we developed a statistical method that analyzes label distributions across state directions. Let $\mathcal{O} = \{x, y, z, \mathrm{roll}, \mathrm{pitch}, \mathrm{yaw}\}$ be the set of observable dimensions in our state space. For each dimension $i \in \mathcal{O}$, we defined N_i as the count of observability labels in that dimension. The total observability count was given by $N_{\mathrm{total}} = \sum_{i\in\mathcal{O}} N_i$. We defined the normalized confidence metrics as

$$\gamma_{\mathrm{trans}} = |\mathcal{O}| \left[\frac{N_x}{N_{\mathrm{total}}}, \frac{N_y}{N_{\mathrm{total}}}, \frac{N_z}{N_{\mathrm{total}}} \right]^T \tag{7}$$

$$\gamma_{\mathrm{rot}} = |\mathcal{O}| \left[\frac{N_{\mathrm{roll}}}{N_{\mathrm{total}}}, \frac{N_{\mathrm{pitch}}}{N_{\mathrm{total}}}, \frac{N_{\mathrm{yaw}}}{N_{\mathrm{total}}} \right]^T \tag{8}$$

Here, the factor $|\mathcal{O}| = 6$ scales the metrics to [0, 1] under uniform distribution across all dimensions. These confidence metrics construct a diagonal 6-by-6 covariance matrix

$$\Sigma_{\mathrm{cov}} = \mathrm{diag}\left(\gamma_{\mathrm{trans}}, \gamma_{\mathrm{rot}}\right) \tag{9}$$

It is important to note that Eq. 9 assesses the ratio of observability labels in a specific direction to the total number of labels. This relative approach is based on the hypothesis that in well-structured environments, observability should be uniformly distributed across all state directions.

By analyzing label distributions across scan correspondences, our approach detects weakly observed directions needing reinforcement.

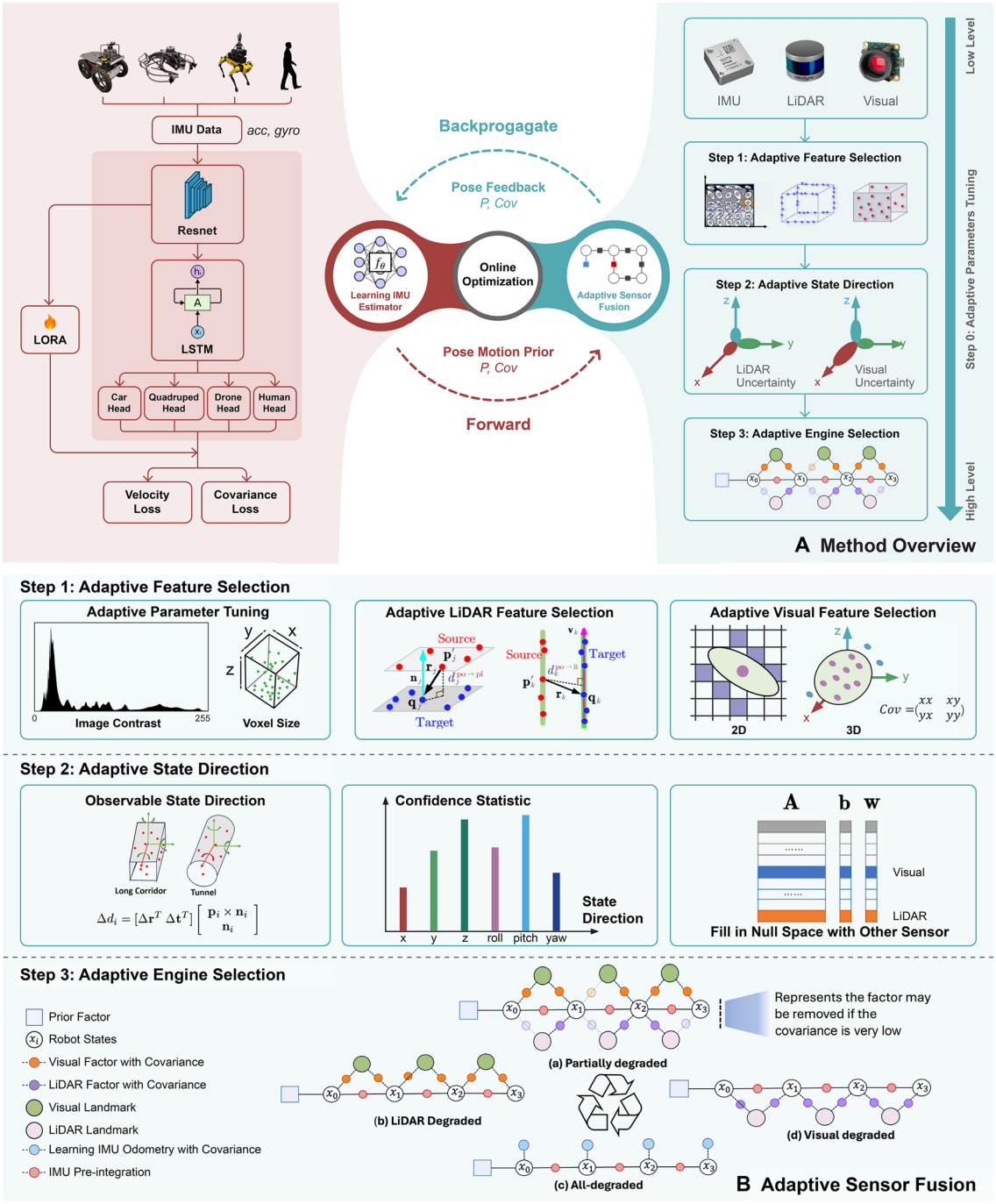


Fig. 8. System overview. (**A**) A hierarchical adaptation framework that integrates adaptive sensor fusion with a learning-based IMU odometry. The sensor fusion module provides free pose supervision to guide the learning process of the IMU odometry, helping to reduce drift and capture motion patterns. In turn, the IMU odometry offers motion priors and serves as a fail-safe solution. (**B**) Illustration of adaptive feature selection, adaptive state direction estimation, and adaptive engine selection.

This allows the system to fuse sensor information more effectively, especially along those vulnerable directions.

Adaptive engine selection

In real-world settings, sensors often operate under persistent degradation (e.g., darkness). To reduce computation and filter faulty data, we proposed an adaptive factor selection strategy that dynamically reconfigured the factor graph on the basis of sensor quality (Fig. 8B, step 3). Factor graph optimization should be dynamically reconfigurable on the basis of degradation type. By evaluating factor quality, the system decided whether to include or exclude a sensor's data. For example, during prolonged visual degradation, where the visual sensor's contribution dropped below 10% for 2 to 4 s (such as in darkness), the system excluded visual factors and relied on LiDAR. If LiDAR became unreliable because of loss of observability, its

factors were also rejected. In mixed degradation scenarios, the system selected the most reliable sensor data. When both LiDAR and vision failed, the system fell back to learning-based IMU odometry.

We adopted a factor graph to represent all residuals but actively selected them in optimization on the basis of degeneracy. Using prior observability analysis, the covariance matrix **cov** guided the inclusion of only well-constrained factors

$$\min_{\mathbf{T}_{i+1}} \left\{ \sum_{p\in\mathbb{F}_i} \left\| e_i^{\text{po}\to\text{po,li,pl}} \right\|^2_{\mathbf{cov}_\text{L}} + \sum_{(i,i+1)\in\mathcal{B}} \left\| e_{\text{vo}}^{\text{prior}} \right\|^2_{\mathbf{cov}_{\text{io}}} + \sum_{(i,i+1)\in\mathcal{B}} \left\| e_{\text{vo}}^{\text{prior}} \right\|^2_{\mathbf{cov}_{\text{vo}}} \right\} \tag{10}$$

In this context, $e_{\text{io}}^{\text{prior}}$ represents the relative pose factor provided by learning-based IMU odometry, weighted by the proposed $\mathbf{cov}_{\mathbf{io}}$. Similarly, $e_{\text{vo}}^{\text{prior}}$ was the relative pose factor from visual odometry, weighted by $\mathbf{cov}_{\mathbf{vo}}$. The term $e_i^{\text{po}\to\text{po,li,pl}}$ indicated point-to-point (line, distance) residuals between two keyframes with covariance $\mathbf{cov}_\text{L}$ (*38*). For more on the relative pose factor, refer to the Supplementary Materials. This adaptability ensured that our sensor fusion remains flexible and resilient, dynamically optimizing performance under varying conditions.

Adaptive parameter tuning

As shown in Fig. 8B (step 1), our method used self-tuning strategies to ensure adaptability across diverse platforms and environments (*54, 55*). For visual odometry, we adjusted image contrast to improve feature detection (*21*). LiDAR odometry, however, may diverge in narrow spaces (e.g., stairwells or corridors) because of parameters optimized for open environments. To address this, the algorithm dynamically adjusted the weighting between visual and LiDAR odometry to adapt to both confined and open areas. For details, see the adaptive state direction in multifloor environments experiment in the Supplementary Materials.

Other parameters, such as voxel size, were adapted to the scale of the environment. The algorithm also used intensity values to filter airborne obscurants (*56*) and gravity vectors to estimate initial rotational alignment between LiDAR and IMU (*57*).

Learning-based inertial odometry

We introduced a learning-based inertial odometry system designed to capture motion patterns across various robots. The structure is shown in Fig. 8A. It consisted of two key components: a pretrained IMU model trained on large datasets and an adapter network for unseen environments. To enable generalization and few-shot capability, the pretrained model was trained on diverse datasets, including wheeled, legged, and aerial robots and handheld devices. It includes five modules: IMU data preprocessing, data augmentation, a heterogeneous shared backbone, a multihead design, and a loss function (Fig. 8A). We first describe the pretrained IMU model and then its adapter network.

Data preprocessing

Our training dataset included more than 100 hours of real-world IMU data from diverse platforms—wheeled robots, drones, legged robots, and human-held devices—sourced from SubT-MRS (*16*), TartanDrive (*30*), IDOL (*58*), Blackbird (*59*), and UZH (*60*). These datasets provide both raw IMU measurements and ground-truth trajectories across eight platforms with varying dynamics. To ensure consistency, we unified all IMU axes to a common frame (*X*-forward, *Y*-left, *Z*-up) and resampled all data to 200 Hz via interpolation. This preprocessing reduced variability across sources and ensured uniform input quality during training and inference.

Data augmentation

Inspired by the work of Cao *et al.* (*61*), we rotated the IMU data around the *X*, *Y*, and *Z* axes by a specified degree. Simultaneously, the corresponding ground truth trajectory was also rotated by the same angle. A pretrained neural network for the IMU was trained on both the original and rotated datasets, as illustrated in Fig. 8B. This approach substantially enlarges the IMU training dataset and enhances the model's generalization capability.

Heterogeneous shared backbone

After preprocessing and augmentation, IMU data were passed through a heterogeneous shared backbone designed to generalize across diverse robot platforms. This backbone maps data into a unified high-dimensional latent space that captures common motion knowledge, enabling efficient adaptation to new platforms with minimal additional training.

As shown in Fig. 8A, the network consisted of three key components: a 1D ResNet backbone, LSTM (long short-term memory) layers, and a multihead output. The ResNet extracted motion features from the IMU data; the LSTM modeled temporal dependencies; and the multihead layer predicted local velocity and its covariance, completing the motion estimation pipeline.

Multihead

Building on a shared heterogeneous backbone, we introduced a multihead architecture to map latent features to velocity estimates, with each head tailored to a specific robot type (Fig. 8B). This architecture offered two main advantages: (i) It enabled learning of diverse motion patterns in parallel, enhancing adaptability across different robot types; and (ii) it stabilized training by avoiding conflicting learning objectives. The mathematical formulation is as follows

$$\begin{aligned} (\hat{v}, \hat{u}) &= f\left(\left({}^B\mathbf{a}_{n-N}, {}^B\omega_{n-N} \right), \ldots, \left({}^B\mathbf{a}_n, {}^B\omega_n \right), \mathbf{h}_{n-N} \right) \\ {}^B\mathbf{a}_n &= {}^B_W\mathbf{R}_n \left({}^W_B\mathbf{R}_n \left(\mathbf{a} - \mathbf{b}_a \right) - {}^W\mathbf{g} \right) \\ {}^B\omega_n &= \omega - \mathbf{b}_g \end{aligned} \tag{11}$$

Here, $f(\cdot)$ represents the function defined by the neural network that processes inputs from the IMU sensor. The network receives as input the acceleration $\mathbf{a}$, angular velocity $\boldsymbol{\omega}$, and the hidden state $\mathbf{h}_{n-N}$ produced by the LSTM at the previous time step. For each IMU measurement, we removed the gravity vector ${}^W\mathbf{g} = [0, 0, 9.8]$ as well as the acceleration bias $\mathbf{b}_a$ and gyro bias $\mathbf{b}_g$. At each time step, the network predicts the current motion on the basis of the hidden state $\mathbf{h}_{n-N}$ and a local window of N samples of acceleration and angular velocity in the inertial frame B. The output of the network consists of the estimated relative velocity $\hat{v}$ and its associated uncertainties $\hat{u}$.

Unlike existing learning-based IMU odometry methods (*29*, *36*), which process IMU data in world coordinates (W), our method operates entirely in the body frame (B). Specifically, we first transformed all IMU data into a unified body frame (*Z*-up, *Y*-left, *X*-forward). We then directly used raw acceleration (α) and angular velocity (ω) in the body frame as input to the network and regressed velocity in the body frame. Because both the input and loss functions were defined in the body frame, our method reduced the risk

of overfitting to specific trajectories in the world frame and exhibited improved generalization.

Loss function

Our network used relative loss functions, as illustrated in Fig. 8B. The relative loss function helped the network capture motion dynamics within a single window, improving the smoothness of the predicted odometry. To optimize the relative loss, we used the mean square error (MSE), defined as follows

$$L_{\mathrm{RL}}^{\mathrm{MSE}}\left(v,\hat{v}\right)=\frac{1}{n}\sum_{i=1}^{n}\left(\sum_{j=m}^{m+M}\left(v_{j\to j+1}-\hat{v}_{j\to j+1}\right)\right) \tag{12}$$

We defined the loss in the body frame. Specifically, $\hat{v}_{j\to j+1}$ represents the 3D velocity output of the network for the j-th window in the body frame, whereas $v_{j\to j+1}$ denotes the corresponding ground truth in the same frame. Variable n is the batch size during training, m is the starting window of the sequence, and M is the total number of LSTM windows.

Covariance

The covariance estimation follows the methodology outlined in (*36*). The network outputs a 3D vector, where each element represents the logarithm of a diagonal element of the covariance matrix Σ. For more implementation details on learning IMU odometry, refer to the Supplementary Materials.

$$L^{\mathrm{NLL}}=\frac{1}{2}\left(v-\hat{v}\right)^{T}\Sigma^{-1}\left(v-\hat{v}\right)+\frac{1}{2}\ln|\Sigma| \tag{13}$$

Adaptor network

Although the pretrained IMU model was designed to generalize across platforms, it cannot guarantee "few-shot" capability for every robot or motion pattern. To address this limitation, we introduced a lightweight low-rank adaptation (LoRA) module (*62*) that enabled the model to acquire knowledge and improve continuously as it processed additional data. This adapter was integrated into the pretrained IMU model (Fig. 8B). To update each weight matrix $W_0\in\mathbb{R}^{d\times k}$ of the pretrained IMU model, we applied a low-rank decomposition

$$W_0+\Delta W=W_0+BA \tag{14}$$

where $B\in\mathbb{R}^{d\times r}$ and $A\in\mathbb{R}^{r\times k}$, r is the rank, and $r\ll\min(d,k)$. During adaptation, W_0 remained fixed, whereas A and B are trainable parameters. Both W_0 and $\Delta W=BA$ process the input, and their outputs are summed, resulting in

$$h=W_0x+\Delta Wx=W_0x+BAx \tag{15}$$

where A is initialized by a random Gaussian distribution and B starts at zero. By freezing the original model parameters, our adaptive sensor fusion framework transferred knowledge solely to the LoRA adapter, preventing catastrophic forgetting. The adapter's small parameter set enabled rapid assimilation of new information for real-time adaptation. For more details, refer to our experiments on catastrophic forgetting in the Supplementary Materials.

This approach ensured flexible adaptation to diverse robotic systems while preserving the pretrained model's generalization ability. It enabled efficient incorporation of new knowledge and system-specific optimization without compromising the core capabilities.

Statistical analysis

Statistical analysis was performed in Python using the NumPy library to compute the ATE (*37*) and robustness metrics (*16*). Evaluations in Fig. 5 and Tables 2 and 3 were conducted on the SubT-MRS dataset (*16*), which contains eight sequences categorized into geometric and mixed degradation groups. Each algorithm was evaluated across multiple runs per sequence using our ICCV (International Conference on Computer Vision) 2023 SLAM Challenge submission system, and results from competing methods were provided by the original authors. For fair comparison, all odometry trajectories were up-sampled to 200 Hz. The mean ATE quantified trajectory accuracy, and the robustness metrics jointly assessed accuracy and completeness, providing an integrated measure of system resilience. Sequences where all baseline algorithms failed, e.g., dense smoke in Figs. 2 and 7, were excluded from the averaging and omitted from Tables 2 and 3.

Supplementary Materials

The PDF file includes:

Supplementary Text
Figs. S1 to S7
Table S1
References (*72*–*83*)

Other Supplementary Material for this manuscript includes the following:

Movies S1 to S6

REFERENCES AND NOTES

1. S. Thrun, Probabilistic robotics. *Commun. ACM* **45**, 52–57 (2002).
2. J. Levinson, J. Askeland, J. Becker, J. Dolson, D. Held, S. Kammel, J. Zico Kolter, D. Langer, O. Pink, V. Pratt, M. Sokolsky, G. Stanek, D. Stavens, A. Teichman, M. Werling, S. Thrun, "Towards fully autonomous driving: Systems and algorithms" in *IEEE Intelligent Vehicles Symposium (IV)* (IEEE, 2011), pp. 163–168.
3. S. Scherer, V. Agrawal, G. Best, C. Cao, K. Cujic, R. Darnley, R. DeBortoli, E. Dexheimer, B. Drozd, R. Garg, I. Higgins, J. Keller, D. Kohanbash, L. Nogueira, R. Pradhan, M. Tatum, V. K. Viswanathan, S. Willits, S. Zhao, H. Zhu, D. Abad, T. Angert, G. Armstrong, R. Boirum, A. Dongare, M. Dworman, S. Hu, J. Jaekel, R. Ji, A. Lai, Y. H. Lee, A. Luong, J. Mangelson, J. Maier, J. Picard, K. Pluckter, A. Saba, M. Saroya, E. Scheide, N. Shoemaker-Trejo, J. Spisak, J. Teza, F. Yang, A. Wilson, H. Zhang, H. Choset, M. Kaess, A. Rowe, S. Singh, J. Zhang, G. A. Hollinger, M. Travers, Resilient and modular subterranean exploration with a team of roving and flying robots. *Field Robot.* **2**, 678–734 (2022).
4. F. Mahlknecht, D. Gehrig, J. Nash, F. M. Rockenbauer, B. Morrell, J. Delaune, D. Scaramuzza, Exploring event camera-based odometry for planetary robots. *IEEE Robot. Autom. Lett.* **7**, 8651–8658 (2022).
5. C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, J. J. Leonard, Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Trans. Robot.* **32**, 1309–1332 (2016).
6. S. Shen, Y. Mulgaonkar, N. Michael, V. Kumar, "Multi-sensor fusion for robust autonomous flight in indoor and outdoor environments with a rotorcraft MAV" in *2014 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2014), pp. 4974–4981.
7. E. A. Wan, R. Van Der Merwe, S. Haykin, "The unscented Kalman filter" in *Kalman Filtering and Neural Networks* (Wiley, 2001), pp. 221–280.
8. J. Zhang, S. Singh, "LOAM: LiDAR odometry and mapping in real-time" in *Robotics: Science and Systems* (RSS, 2014), vol. 2, p. 07.
9. J. Zhang, S. Singh, "Visual-lidar odometry and mapping: Low-drift, robust, and fast" in *2015 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2015), pp. 2174–2181.
10. S. Khattak, H. Nguyen, F. Mascarich, T. Dang, K. Alexis, "Complementary multi–modal sensor fusion for resilient robot pose estimation in subterranean environments" in *2020 International Conference on Unmanned Aircraft Systems (ICUAS)* (IEEE, 2020), pp. 1024–1029.
11. M. Camurri, M. Ramezani, S. Nobili, M. Fallon, Pronto: A multi-sensor state estimator for legged robots in real-world scenarios. *Front. Robot. AI* **7**, 68 (2020).
12. J. Graeter, A. Wilczynski, M. Lauer, "Limo: Lidar-monocular visual odometry" in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2018), pp. 7872–7879.
13. X. Zuo, P. Geneva, W. Lee, Y. Liu, G. Huang, "Lic-fusion: Lidar-inertial-camera odometry" in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2019), pp. 5848–5854.
14. D. Wisth, M. Camurri, S. Das, M. Fallon, Unified multi-modal landmark tracking for tightly coupled lidar-visual-inertial odometry. *IEEE Robot. Autom. Lett.* **6**, 1004–1011 (2021).

15. J. Lin, F. Zhang, "R3LIVE: A robust, real-time, RGB-colored, LiDAR-inertial-visual tightly-coupled state estimation and mapping package" in *2022 International Conference on Robotics and Automation (ICRA)* (IEEE, 2022), pp. 10672–10678.
16. S. Zhao, Y. Gao, T. Wu, D. Singh, R. Jiang, H. Sun, M. Sarawata, Y. Qiu, W. Whittaker, I. Higgins, Y. Du, S. Su, C. Xu, J. Keller, J. Karhade, L. Nogueira, S. Saha, J. Zhang, W. Wang, C. Wang, S. Scherer, "SubT-MRS dataset: Pushing SLAM towards all-weather environments" in *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, 2024), pp. 22647–22657.
17. K. Ebadi, L. Bernreiter, H. Biggie, G. Catt, Y. Chang, A. Chatterjee, C. E. Denniston, S. P. Deschênes, K. Harlow, S. Khattak, L. Nogueira, M. Palieri, P. Petráček, M. Petrlík, A. Reinke, V. Krátký, S. Zhao, A. A. Agha-mohammadi, K. Alexis, C. Heckman, K. Khosoussi, N. Kottege, B. Morrell, M. Hutter, F. Pauling, F. Pomerleau, M. Saska, S. Scherer, R. Siegwart, J. L. Williams, L. Carlone, Present and future of SLAM in extreme environments: The DARPA SubT challenge. *IEEE Trans. Robot.* **40**, 936–959 (2024).
18. M. Bostelmann, P. Ruggeri, A. Rita Circelli, F. Costanzo, D. Menghini, S. Vicari, P. Lavenex, P. Banta Lavenex, Path integration and cognitive mapping capacities in Down and Williams syndromes. *Front. Psychol.* **11**, 571394 (2020).
19. C. Wang, K. Ji, J. Geng, Z. Ren, T. Fu, F. Yang, Y. Guo, H. He, X. Chen, Z. Zhan, Q. du, S. Su, B. Li, Y. Qiu, Y. du, Q. Li, Y. Yang, X. Lin, Z. Zhao, Imperative learning: A self-supervised neural-symbolic learning framework for robot autonomy. *Int. J. Robot. Res.* **2025**, 10.1177/02783649251353181 (2025).
20. F. Dellaert, M. Kaess, Factor graphs for robot perception. *Found. Trends Robot.* **6**, 1–139 (2017).
21. P. Bergmann, R. Wang, D. Cremers, Online photometric calibration of auto exposure video for real-time visual odometry and slam. *IEEE Robot. Autom. Lett.* **3**, 627–634 (2018).
22. S. Zhao, P. Wang, H. Zhang, Z. Fang, S. Scherer, "Tp-tio: A robust thermal-inertial odometry with deep thermalpoint" in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2020), pp. 4505–4512.
23. P. Geneva, K. Eckenhoff, W. Lee, Y. Yang, G. Huang, "OpenVINS: A research platform for visual-inertial estimation" in *2020 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2020), pp. 4666–4672.
24. T. Qin, P. Li, S. Shen, Vins-mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Trans. Robot.* **34**, 1004–1020 (2018).
25. W. Xu, F. Zhang, Fast-lio: A fast, robust lidar-inertial odometry package by tightly-coupled iterated Kalman filter. *IEEE Robot. Autom. Lett.* **6**, 3317–3324 (2021).
26. A. Segal, D. Haehnel, S. Thrun, "Generalized-ICP" in *Robotics: Science and Systems* (RSS, 2009), vol. 2, p. 435.
27. C. Zheng, Q. Zhu, W. Xu, X. Liu, Q. Guo, F. Zhang, "FAST-LIVO: Fast and tightly-coupled sparse-direct LiDAR-inertial-visual odometry" in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2022), pp. 4003–4009.
28. T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, D. Rus, "LIO-SAM: Tightly-coupled Lidar inertial odometry via smoothing and mapping" in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2020), pp. 5135–5142.
29. G. Cioffi, L. Bauersfeld, E. Kaufmann, D. Scaramuzza, Learned inertial odometry for autonomous drone racing. *IEEE Robot. Autom. Lett.* **8**, 2684–2691 (2023).
30. M. Sivaprakasam, P. Maheshwari, M. G. Castro, S. Triest, M. Nye, S. Willits, A. Saba, W. Wang, S. Scherer, TartanDrive 2.0: More modalities and better infrastructure to further self-supervised learning research in off-road driving tasks. arXiv:2402.01913 [cs.RO] (2024); https://arxiv.org/abs/2402.01913.
31. C. Chen, X. Lu, A. Markham, N. Trigoni, "Ionet: Learning to cure the curse of drift in inertial odometry" in *AAAI Conference on Artificial Intelligence* (PKP Publishing Services, 2018), vol. 32; https://doi.org/10.1609/aaai.v32i1.12102.
32. X. Liu, Z. Gao, H. Cheng, P. Wang, B. M. Chen, "Learning-based low light image enhancement for visual odometry" in *2020 IEEE 16th International Conference on Control & Automation (ICCA)* (IEEE, 2020), pp. 1143–1148.
33. D. Chen, N. Wang, R. Xu, W. Xie, H. Bao, G. Zhang, "RNIN-VIO: Robust neural inertial navigation aided visual-inertial odometry in challenging scenes" in *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (IEEE, 2021), pp. 275–283.
34. H. Yan, Q. Shan, Y. Furukawa, "RIDI: Robust IMU double integration" in *European Conference on Computer Vision (ECCV)* (ECVA, 2018), pp. 621–636.
35. M. Brossard, A. Barrau, S. Bonnabel, AI-IMU dead-reckoning. *IEEE Trans. Intell. Veh.* **5**, 585–595 (2020).
36. W. Liu, D. Caruso, E. Ilg, J. Dong, A. I. Mourikis, K. Daniilidis, V. Kumar, J. Engel, TLIO: Tight learned inertial odometry. *IEEE Robot. Autom. Lett.* **5**, 5653–5660 (2020).
37. J. Sturm, N. Engelhard, F. Endres, W. Burgard, D. Cremers, "A benchmark for the evaluation of RGB-D SLAM systems" in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2012), pp. 573–580.
38. S. Zhao, H. Zhang, P. Wang, L. Nogueira, S. Scherer, "Super odometry: IMU-centric LiDAR-visual-inertial estimator for challenging environments" in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2021), pp. 8729–8736.
39. A. Prorok, M. Malencia, L. Carlone, G. S. Sukhatme, B. M. Sadler, V. Kumar, Beyond robustness: A taxonomy of approaches towards resilient multi-robot systems. arXiv:2109.12343 [cs.RO] (2021); https://arxiv.org/abs/2109.12343.
40. C. Forster, L. Carlone, F. Dellaert, D. Scaramuzza, "IMU preintegration on manifold for efficient visual-inertial maximum-a-posteriori estimation" in *Robotics: Science and Systems XI* (RSS, 2015), p. 06.
41. J. Lv, X. Zuo, K. Hu, J. Xu, G. Huang, Y. Liu, Observability-aware intrinsic and extrinsic calibration of LiDAR-IMU systems. *IEEE Trans. Robot.* **38**, 3734–3753 (2022).
42. P. Furgale, T. D. Barfoot, G. Sibley, "Continuous-time batch estimation using temporal basis functions" in *2012 IEEE International Conference on Robotics and Automation* (IEEE, 2012), pp. 2088–2095.
43. M. Burri, J. Nikolic, P. Gohl, T. Schneider, J. Rehder, S. Omari, M. W. Achtelik, R. Siegwart, The EuRoC micro aerial vehicle datasets. *Int. J. Robot. Res.* **35**, 1157–1163 (2016).
44. M. Hwangbo, J.-S. Kim, T. Kanade, "Inertial-aided KLT feature tracking for a moving camera" in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2009), pp. 1909–1916.
45. D. Muhle, L. Koestler, N. Demmel, F. Bernard, D. Cremers, "The probabilistic normal epipolar constraint for frame-to-frame rotation optimization under uncertain feature positions" in *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, 2022), pp. 1819–1828.
46. J. Demantké, C. Mallet, N. David, B. Vallet, Dimensionality based scale selection in 3D lidar point clouds, in *Laserscanning* (University of Calgary, 2011), vol. XXXVIII-5/W12, pp. 97–102.
47. T. Shan, B. Englot, C. Ratti, D. Rus, "LVI-SAM: Tightly-coupled LiDAR-visual-inertial odometry via smoothing and mapping" in *2021 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2021), pp. 5692–5698.
48. J. Zhang, M. Kaess, S. Singh, "On degeneracy of optimization-based state estimation problems" in *2016 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2016), pp. 809–816.
49. W. Zhen, S. Scherer, "Estimating the localizability in tunnel-like environments using LiDAR and UWB" in *2019 International Conference on Robotics and Automation (ICRA)* (IEEE, 2019), pp. 4903–4908.
50. X. Ding, F. Han, T. Yang, Y. Wang, R. Xiong, Degeneration-aware localization with arbitrary global-local sensor fusion. *Sensors* **21**, 4042 (2021).
51. A. Hinduja, B.-J. Ho, M. Kaess, "Degeneracy-aware factors with applications to underwater SLAM" in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2019), pp. 1293–1299.
52. K. Ebadi, M. Palieri, S. Wood, C. Padgett, A.-a. Agha-mohammadi, DARE-SLAM: Degeneracy-aware and resilient loop closing in perceptually-degraded environments. *J. Intell. Robot. Syst.* **102**, 2 (2021).
53. N. Chebrolu, T. Läbe, O. Vysotska, J. Behley, C. Stachniss, Adaptive robust kernels for non-linear least squares problems. *IEEE Robot. Autom. Lett.* **6**, 2240–2247 (2021).
54. M. Ramezani, K. Khosoussi, G. Catt, P. Moghadam, J. Williams, P. Borges, F. Pauling, N. Kottege, Wildcat: Online continuous-time 3D LiDAR-inertial slam. arXiv:2205.12595 [cs.RO] (2022); https://arxiv.org/abs/2205.12595.
55. H. Lim, D. Kim, B. Kim, H. Myung, "AdaLIO: Robust adaptive LiDAR-inertial odometry in degenerate indoor environments" in *2023 International Conference on Ubiquitous Robots (UR)* (Korea Robotics Society, 2023), pp. 48–53.
56. G. Best, R. Garg, J. Keller, G. A. Hollinger, S. Scherer, "Resilient multi-sensor exploration of multifarious environments with a team of aerial robots" in *Robotics: Science and Systems (RSS)* (RSS, 2022), p. 004.
57. R. Nemiroff, K. Chen, B. T. Lopez, "Joint on-manifold gravity and accelerometer intrinsics estimation for inertially aligned mapping" in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2023), pp. 1388–1394.
58. S. Sun, D. Melamed, K. Kitani, "IDOL: Inertial deep orientation-estimation and localization" in *AAAI Conference on Artificial Intelligence* (PKP Publishing Services, 2021), vol. 35, pp. 6128–6137.
59. A. Antonini, W. Guerra, V. Murali, T. Sayre-McCord, S. Karaman, "The Blackbird dataset: A large-scale dataset for UAV perception in aggressive flight" in *International Symposium on Experimental Robotics* (Springer, 2020), pp. 130–139.
60. J. Delmerico, T. Cieslewski, H. Rebecq, M. Faessler, D. Scaramuzza, "Are we ready for autonomous drone racing? The UZH-FPV drone racing dataset" in *2019 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2019), pp. 6713–6719.
61. X. Cao, C. Zhou, D. Zeng, Y. Wang, "RIO: Rotation-equivariance supervised learning of robust inertial odometry" in *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, 2022), pp. 6614–6623.
62. E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, W. Chen, "LoRA: Low-rank adaptation of large language models" in *International Conference on Learning Representations* (ICLR, 2022), pp. 1–13.
63. A. Antonini, W. Guerra, V. Murali, T. Sayre-McCord, S. Karaman, The Blackbird UAV dataset. *Int. J. Robot. Res.* **39**, 1346–1364 (2020).

64. W. Xu, Y. Cai, D. He, J. Lin, F. Zhang, FAST-LIO2: Fast direct LiDAR-inertial odometry. *IEEE Trans. Robot.* **38**, 2053–2073 (2022).
65. Z. Liu, X. Liu, F. Zhang, Efficient and consistent bundle adjustment on LiDAR point clouds. *IEEE Trans. Robot.* **39**, 4366–4386 (2023).
66. W. Wu, C. Chen, B. Yang, X. Zou, F. Liang, Y. Xu, X. He, DALI-SLAM: Degeneracy-aware LiDAR-inertial SLAM with novel distortion correction and accurate multi-constraint pose graph optimization. *ISPRS J. Photogramm. Remote Sens.* **221**, 92–108 (2025).
67. H. Lim, S. Yeon, S. Ryu, Y. Lee, Y. Klm, J. Yun, E. Jung, D. Lee, H. Myung, "A single correspondence is enough: Robust global registration to avoid degeneracy in urban environments" in *2022 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2022), pp. 8010–8017.
68. Y. Wu, T. Guadagnino, L. Wiesmann, L. Klingbeil, C. Stachniss, H. Kuhlmann, "LIO-EKF: High frequency LiDAR-inertial odometry using extended Kalman filters," *2024 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2024), pp. 13741–13747.
69. K. Chen, B. T. Lopez, A.-A. Agha-Mohammadi, A. Mehta, Direct LiDAR odometry: Fast localization with dense point clouds. *IEEE Robot. Autom. Lett.* **7**, 2000–2007 (2022).
70. G. Kim, A. Kim, "Scan context: Egocentric spatial descriptor for place recognition within 3D point cloud map" in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2018), pp. 4802–4809.
71. S. Zhao, S. Zhou, Y. Zhang, J. Zhang, C. Wang, W. Wang, S. Scherer, Resilient odometry via hierarchical adaptation, Zenodo (2025); https://doi.org/10.5281/zenodo.17569700.
72. A. Tagliabue, J. Tordesillas, X. Cai, A. Santamaria-Navarro, J.P. How, L. Carlone, A.A. Agha-mohammadi, "Lion: LiDAR-inertial observability-aware navigator for vision-denied environments" in *Experimental Robotics: The 17th International Symposium* (Springer, 2021), pp. 380–390.
73. T. Tuna, J. Nubert, Y. Nava, S. Khattak, M. Hutter, X-ICP: Localizability-aware LiDAR registration for robust localization in extreme environments. *IEEE Trans. Robot.* **40**, 452–471 (2024).
74. W. Talbot, J. Nash, M. Paton, E. Ambrose, B. Metz, R. Thakker, R. Etheredge, M. Ono, V. Ila, "Principled ICP covariance modelling in perceptually degraded environments for the EELS mission concept" in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2023), pp. 10763–10770.
75. M. Brossard, S. Bonnabel, A. Barrau, A new approach to 3D ICP covariance estimation. *IEEE Robot. Autom. Lett.* **5**, 744–751 (2020).
76. J. Nubert, T. Tuna, J. Frey, C. Cadena, K. J. Kuchenbecker, S. Khattak, M. Hutter, Holistic fusion: Task-and setup-agnostic robot localization and state estimation with factor graphs. arXiv:2504.06479 [cs.RO] (2025); https://arxiv.org/abs/2504.06479.
77. X. Zhong, Y. Li, S. Zhu, W. Chen, X. Li, J. Gu, "LVIO-SAM: A multi-sensor fusion odometry via smoothing and mapping" in *2021 IEEE International Conference on Robotics and Biomimetics (ROBIO)* (IEEE, 2021), pp. 440–445.
78. R. Buchanan, V. Agrawal, M. Camurri, F. Dellaert, M. Fallon, Deep IMU bias inference for robust visual-inertial odometry with factor graphs. *IEEE Robot. Autom. Lett.* **8**, 41–48 (2023).
79. N. Gelfand, L. Ikemoto, S. Rusinkiewicz, M. Levoy, "Geometrically stable sampling for the ICP algorithm" in *International Conference on 3-D Digital Imaging and Modeling (3DIM)* (IEEE, 2003), pp. 260–267.
80. T. Wen, Y. Fang, B. Lu, X. Zhang, C. Tang, LIVER: A tightly coupled LiDAR-inertial-visual state estimator with high robustness for underground environments. *IEEE Robot. Autom. Lett.* **9**, 2399–2406 (2024).
81. X. Zheng, J. Zhu, Traj-LIO: A resilient multi-LiDAR multi-IMU state estimator through sparse Gaussian process. arXiv:2402.09189 [cs.RO] (2024); https://arxiv.org/abs/2402.09189.
82. C. Campos, R. Elvira, J. J. G. Rodriguez, J. M. Montiel, J. D. Tardós, ORB-SLAM3: An accurate open-source library for visual, visual–inertial, and multimap SLAM. *IEEE Trans. Robot.* **37**, 1874–1890 (2021).
83. M. Dreissig, D. Scheuble, F. Piewak, J. Boedecker, Survey on LiDAR perception in adverse weather conditions. arXiv:2304.06312 [cs.RO] (2023); https://arxiv.org/abs/2304.06312.

Acknowledgments: We thank Y. Qiu, M. Kaess, S. Suresh, and S. Tulsiani for suggestions for the manuscript. We appreciate the help of H. Zhu, R. Jiang, H. Sun, T. Wu, Y. Gao, D. Singh, L. Nogueira, G. Chen, P. Maheshwari, M. Sivaprakasam, S. Triest, M. Nye, Y. Liu, S. Willits, J. Keller, J. Karhade, Y. He, M. Yu, A. Jong, and J. Rogers in real-world experiments. **Funding:** This work was funded by US Army Research Lab W911NF-23-S-0001, US Army Research Lab W911NF2120152, US Army Research Lab W911NF2420125, and US Army Research Lab W911NF-17-S-0003.
Author contributions: Conceptualization: S. Zhao, W.W., C.W., J.Z., and S.S. Methodology: S. Zhao, S. Zhou, W.W., and C.W. Investigation: S. Zhao, S. Zhou, and Y.Z. Visualization: S. Zhao, S. Zhou, J.Z., C.W., and S.S. Funding acquisition: S.S. and W.W. Project administration: S. Zhao, W.W., and S.S. Supervision: W.W. and S.S. Writing—original draft: S. Zhao and S. Zhou. Writing—review and editing: S. Zhao, C.W, W.W., S. Zhou, J.Z., S.S., and Y.Z. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to support the conclusions of this manuscript are included in the main text or Supplementary Materials. The data for this study have been deposited in the database https://zenodo.org/records/17569700 (*71*).

Submitted 4 January 2025
Accepted 13 November 2025
Published 10 December 2025
10.1126/scirobotics.adv1818

Correction (9 January 2026): The citation information for references 66 and 68 was incorrect and has now been updated to Wu *et al*., "DALI-SLAM: Degeneracy-aware LiDAR-inertial SLAM with novel distortion correction and accurate multi-constraint pose graph optimization," and Wu *et al*., "LIO-EKF: High frequency LiDAR-inertial odometry using extended Kalman filters," respectively. No data or conclusions were affected.

Science Robotics

Resilient odometry via hierarchical adaptation

Shibo Zhao, Sifan Zhou, Yuchen Zhang, Ji Zhang, Chen Wang, Wenshan Wang, and Sebastian Scherer

Sci. Robot. **10** (109), eadv1818. DOI: 10.1126/scirobotics.adv1818

Science Robotics (ISSN 2470-9476) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Robotics* is a registered trademark of AAAS.