FIELD ROBOTS

Reinforcement learning–based framework for whale rendezvous via autonomous sensing robots

Ninad Jadhav[1,2]*†, Sushmita Bhattacharya[1,2]*†, Daniel Vogt[1,2], Yaniv Aluma[1], Pernille Tønnesen[1,3], Akarsh Prabhakara[4], Swarun Kumar[4], Shane Gero[1,5], Robert J. Wood[1,2], Stephanie Gil[1,2]

Rendezvous with sperm whales for biological observations is made challenging by their prolonged dive patterns. Here, we propose an algorithmic framework that codevelops multiagent reinforcement learning–based routing (autonomy module) and synthetic aperture radar–based very high frequency (VHF) signal–based bearing estimation (sensing module) for maximizing rendezvous opportunities of autonomous robots with sperm whales. The sensing module is compatible with low-energy VHF tags commonly used for tracking wildlife. The autonomy module leverages in situ noisy bearing measurements of whale vocalizations, VHF tags, and whale dive behaviors to enable time-critical rendezvous of a robot team with multiple whales in simulation. We conducted experiments at sea in the native habitat of sperm whales using an "engineered whale"—a speedboat equipped with a VHF-emitting tag, emulating five distinct whale tracks, with different whale motions. The sensing module shows a median bearing error of 10.55° to the tag. Using bearing measurements to the engineered whale from an acoustic sensor and our sensing module, our autonomy module gives an aggregate rendezvous success rate of 81.31% for a 500-meter rendezvous distance using three robots in postprocessing. A second class of fielded experiments that used acoustic-only bearing measurements to three untagged sperm whales showed an aggregate rendezvous success rate of 68.68% for a 1000-meter rendezvous distance using two robots in postprocessing. We further validated these algorithms with several ablation studies using a sperm whale visual encounter dataset collected by marine biologists.

INTRODUCTION

Sperm whales are known for their multilevel social structure, complex communication system, advanced cognition, rich behavioral diversity, and cultures (*1*–*8*). To support further studies of these whales, detailed data collection is crucial. Over the past several years, innovative technologies such as underwater acoustic sensing to record whale vocalizations and inertial measurement unit–enabled animal-worn tags have been applied to collect increasingly rich continuous datasets from sperm whales (*9*). These tags have enabled researchers to localize and track sperm whales to study their diving behavior, foraging, and communication (*10*–*13*).

Acquiring in situ (at sea) visual observations of whales, henceforth referred to as "rendezvous," is critical for validating tag deployments, acoustic recordings of whale vocalizations underwater, and other remotely sensed data. These observations provide highly valued ground-truth labels, for example, physical proximity and interactions between individuals, photo-identification, and tracking of other contextual or behavioral events, allowing for important ties between vocalizations and behavior (*14*). Specifically, the ability to combine whale-generated acoustic data with visual in situ observations provides valuable information about the function of whale vocalizations (*15*, *16*). This is one of the major goals of several whale observation and science projects, including the Earth Species Project (*15*) and the Cetacean Translation Initiative (CETI) project (*17*, *18*). Project CETI, which this work is a part of, focuses on decoding the acoustic communication of sperm whales in Dominica by applying advanced machine learning and state-of-the-art robotics (*18*).

Unfortunately, collecting in situ observations of whales is very time consuming, even with advancements in animal-worn sensor tags. Researchers still face challenges in deploying these tags and collecting data, primarily because of the behavior of sperm whales, who spend less than 25% of their time at the surface (Fig. 1A) (*19*). This limited surface activity results in frequent missed rendezvous opportunities, further complicating data collection efforts. Consequently, substantial time and effort are required to successfully implement tagging and monitoring procedures. Figure 1B shows the recorded time lag between consecutive visual whale identification times in an example study of social sperm whale behavior in Dominica. Any time lag of more than roughly 1 hour indicates a missed visual rendezvous opportunity, given that the time between consecutive whale surfacings in this community is ~57 min (*20*).

Angle-of-arrival (AOA) tracking of whales can be used to improve rendezvous outcomes, by measuring AOA of animal-worn very high frequency (VHF) signal–emitting tags when they are surfaced (*21*–*24*) and measuring AOA to whales using acoustic tracking of their vocalizations while they are underwater. Using GPS-enabled tags can also provide useful information about whale locations, but the VHF tag alternatives are generally less expensive and more power efficient, making them more compatible for fielded operations (*25*). Sensing the AOA for VHF tags is traditionally done in the field using directional antennas (*26*–*29*), which most often require manual operation and provide coarse AOA accuracy because of their reliance on signal strength.

The application of robotics and autonomy, such as fleets of autonomous unpiloted aerial vehicles (UAVs) or surface vehicles, could mitigate manual effort and costs, improving rendezvous success.

[1]Project CETI, New York, NY, USA. [2]John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. [3]Zoophysiology, Department of Biology, Aarhus University, 8000 Aarhus, Denmark. [4]College of Engineering, Carnegie Mellon University, Pittsburgh, PA, USA. [5]Department of Biology, Carleton University, Ottawa, Ontario, Canada.
*Corresponding author. Email: njadhav@g.harvard.edu (N.J.); sushmita_bhattacharya@g.harvard.edu (S.B.)
†These authors contributed equally to this work.

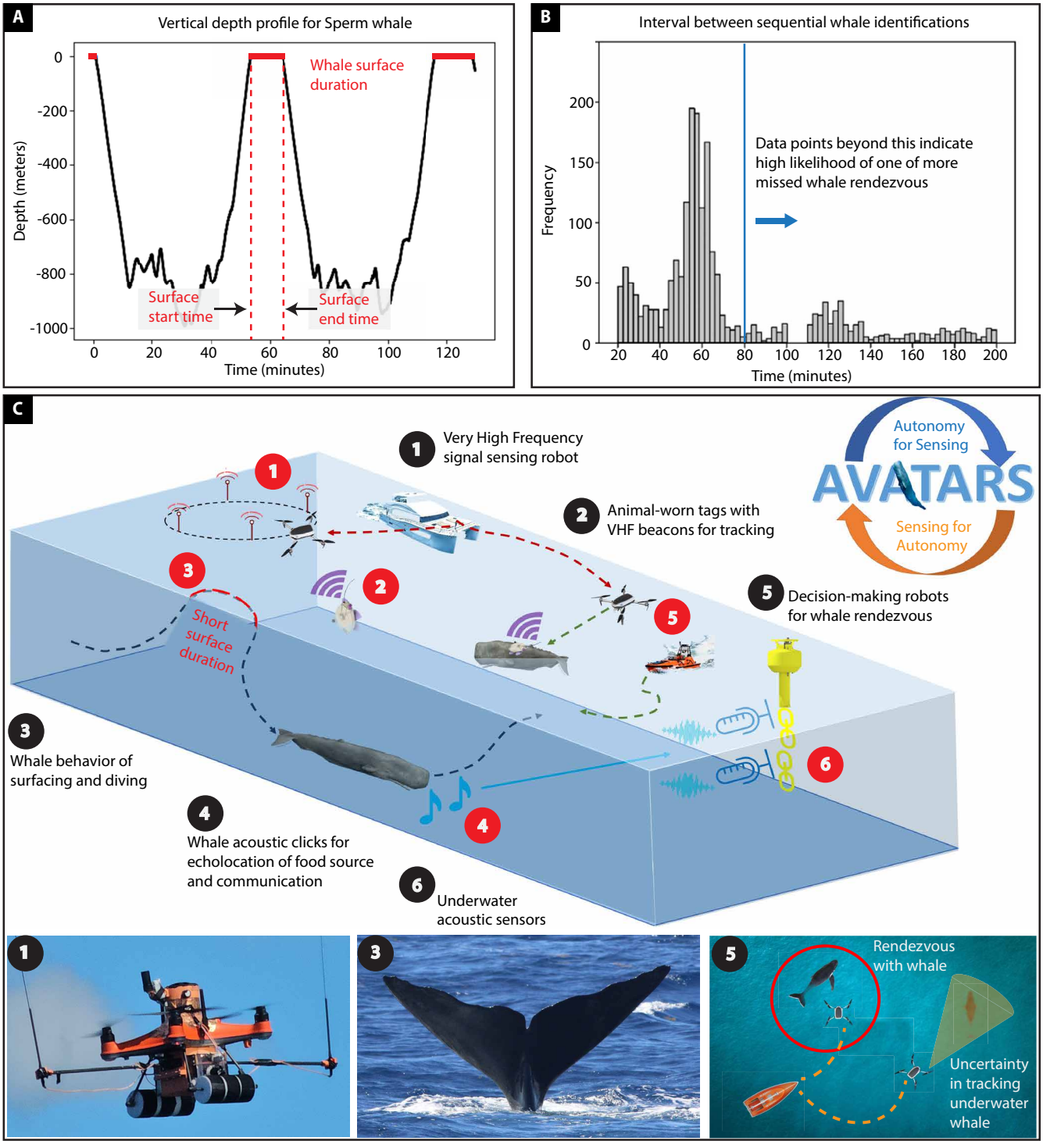

Fig. 1. AVATARS framework for sperm whale rendezvous. (**A**) An instance of a sperm whale's vertical dive profile obtained from pressure sensor data showing very short surfacing duration (*19*). (**B**) Missed sperm whale rendezvous opportunities based on observations collected over 8 years by marine biologists (*20*). (**C**) Our framework, where a UAV with our sensing module obtains AOA to a VHF signal–emitting tag attached to a whale. Acoustic sensors such as hydrophone buoys or towed arrays enable acoustic AOA when whales dive underwater. Robots with our autonomy module use AOA from different sensing modalities while accounting for uncertainty in whale positions due to noisy sensor measurements and stochasticity in surfacings to localize whales and achieve rendezvous with whales when they surface.

This would allow for more efficient and comprehensive data acquisition, both spatially and temporally, thereby improving research outcomes (*16*, *30*, *31*).

At its core, the whale rendezvous problem can be thought of as a variant of an autonomous vehicle routing problem with time window constraints, albeit with some critical differences, which we will outline later. Here, all surfaced whales can be modeled as stochastically appearing requests that must be serviced by a group of autonomous robots and where the locations and servicing time windows of the requests are unknown a priori. Previous work introduced a multiagent rollout–based reinforcement learning (RL) approach in which each agent sequentially minimizes an expected cost by evaluating its available controls, one agent at a time. This process scales linearly with the number of robots (*32*). Here, several possible futures are "rolled out," and the expected cost over these futures is often calculated using Monte Carlo simulations. Thus, the quality of these approaches is often characterized by how well they can approximate the future. Rollout-based RL has shown great promise for learning multiagent policies that take uncertainty into account over long planning horizons, and some variants can also deal with partial observability of the state (*33*–*39*). Rollout-based RL methods have also shown promising results in routing problems, with their ability to adapt to changes in environmental uncertainties by using online replanning with the latest data as they become available, combined with offline learned models that help better estimate future costs (*34*, *35*, *37*, *40*). These attributes also make rollout-based RL appealing for the problem of whale rendezvous.

However, critical differences from the vehicle routing problem must be addressed when planning rendezvous with whales. The whale rendezvous problem is characterized by a hybrid discrete-continuous state, time criticality, and partial observability. First, the hybrid discrete-continuous state refers to the whales' position in three-dimensional space, which represents the continuous component, whereas the discrete component indicates the whales' brief surfacing intervals. This dual nature adds complexity to the planning process. Second, the issue of time criticality arises because sperm whales typically spend less than 10 min at the surface each hour, making timely rendezvous essential. Last, the problem is further complicated by partial observability. While whales are underwater or surfaced but out of visual range, only sparse and highly uncertain measurements of their position or heading are available, either through acoustic AOA estimates of their vocalizations or, in the case of tagged whales, VHF AOA from whale-worn tags during surfacing. This lack of precise positional data introduces challenges to the RL framework. Ideally, a rollout-based RL method could use biological models to accurately capture whale motion while allowing for replanning that is responsive to in situ–sensed data about the whales' whereabouts. For example, sensing AOA to VHF pings from animal-worn tags can enable detection of the surfacing of the whale because VHF signals are usually detectable over several miles. This is needed to refine routes to successfully catch their short rendezvous window and thus highlights the criticality of improving in situ sensing for developing whale-specific RL methods for autonomy.

In this work, we codeveloped the core algorithms of our framework, Autonomous Vehicles for whAle Tracking And Rendezvous by remote Sensing (AVATARS) (Fig. 1C), that combines an RL-based autonomous routing algorithm and VHF signal–based sensing (Fig. 2) to maximize the chances of rendezvous with whales. The two interrelated modules of autonomy and sensing are as follows. The autonomy module comprises a multiagent rollout–based RL algorithm that determines the positioning commands to autonomous robots to maximize whale rendezvous opportunities by fusing data from multiple sensors, such as acoustic and VHF AOA. This is done to determine likely surfacing times of the whales and coordinates the fleet of robots to be in the right place at the right time to catch the surfacing opportunity. The rollout horizon computes the cost of assigning a robot to visit with a particular whale as the expected total time to rendezvous with all whales while accounting for future whale surfacing behaviors. The cost is then minimized over all possible assignments to increase rendezvous success. Key innovations to our approach include resampling future uncertainties during the rollout horizon, considering discrete events like whale surfacing, and dynamically adjusting whale motion predictions and robot assignments by adapting to sensed information. Better sensing is crucial for adjusting future projections of whale surfacings and locations over the rollout horizon, which, in turn, greatly affects successful rendezvous. The sensing module involves the development of a waterproof sensing payload that addresses the size, weight, extreme environments, and power constraints for deployment at sea and a signal processing algorithm based on synthetic aperture radar (SAR) that measures AOA to the VHF tags. SAR methods apply signal processing to phase differences from several pings acquired over the flight of the UAV, thereby emulating the AOA-finding capabilities of a multiple antenna array in software and providing higher accuracy than signal strength–based methods (*41*–*47*). We show that high-accuracy VHF azimuth AOA can be obtained using a UAV via extensive hardware experiments in the field. The sensing module is compatible with off-the-shelf passive VHF tags, which are the most widely used in field operations (*22*–*25*). All core aspects of our sensing and autonomy modules were validated using heterogeneous sensor data collected in the sperm whale environment along the coast of Dominica using a speedboat and sperm whale tracks, as well as through various ablation studies.

RESULTS

We validated the algorithmic developments of the AVATARS framework across three expeditions in 2023 and 2024, both in situ at sea in Dominica and through various ablation studies. We first performed controlled experiments with an engineered whale, an 8-m-long Zodiac speedboat equipped with a passive VHF-emitting tag. We collected real-time acoustic and VHF AOA measurements as the engineered whale mimicked sperm whale movements. Next, we collected the real-time acoustic AOA measurements to three untagged sperm whales. Acoustic AOA to both the engineered whale and sperm whales during their underwater phases was acquired using a towed hydrophone array (fig. S1), whereas the VHF AOA was acquired using our sensing UAV. These field experiments were conducted across several days, covered different whale-surfacing patterns across experiments, and used different sensor configurations (acoustic AOA only for sperm whales and acoustic and VHF AOA for the engineered whale). The in situ AOA measurements were provided to our autonomy module, which postprocessed these measurements and furnished routing commands for autonomous robots in simulation. The robots' routes were verified against the ground-truth whale tracks to report rendezvous success or failure. We report rendezvous success for different distances between the autonomous robots and a surfaced whale, henceforth referred to as the rendezvous radius.

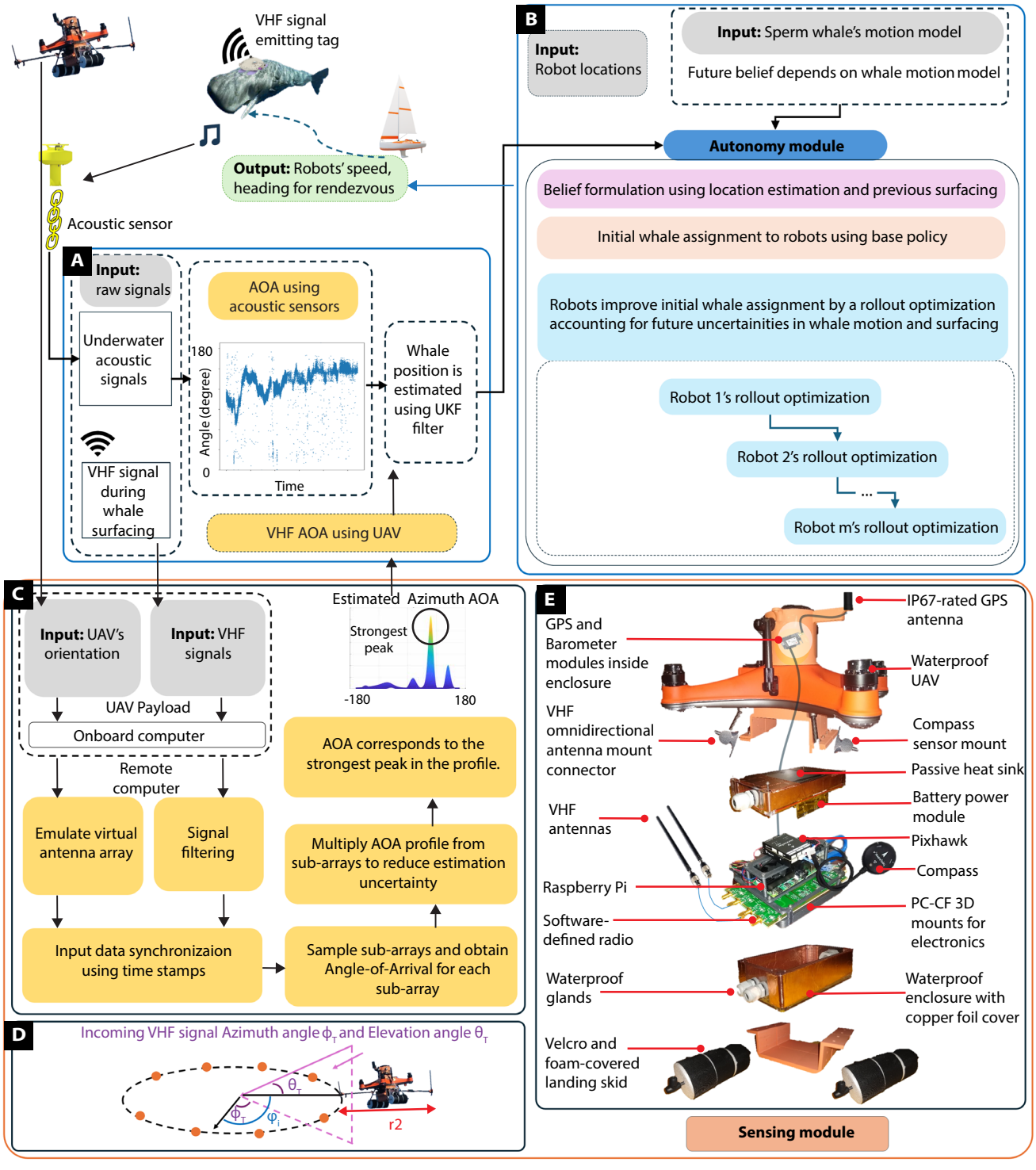

Fig. 2. System design of AVATARS autonomy and sensing modules. (**A**) Whales are localized using a UKF that uses noisy AOA measurements from acoustic and VHF sensors. (**B**) Autonomy module's multiagent rollout architecture. The sperm whale's motion model accounts for future uncertainties by predicting future surfacing times, duration (*19*), and changes in whale heading. The module then performs a series of optimizations to assign whales to robots. Each optimization uses the knowledge of the optimized assignment decision communicated by the previous robot in the series. The module furnishes control commands to enable the rendezvous of robots with whales in minimal time. (**C**) VHF AOA computation flowchart. (**D**) Our formulation has two antennas with a fixed separation. The UAV's orientation is denoted by φ. (**E**) SwellPro SplashDrone 4 UAV and sensing payload hardware. The payload weighs 1.6 kg with dimensions of 18 cm by 12 cm by 8 cm. The UAV dimensions are 40 cm by 40 cm by 30 cm. PC-CF, polycarbonate material with reinforced carbon fiber.

Setup

A whale motion behavior model corresponding to their underwater and surface phases was used by the autonomy module to predict future surface intervals and whale motion. We obtained this information using the Dominica Sperm Whale Project (DSWP) dataset. This dataset contains a rich set of sperm whale sighting data collected over 15 years by marine biologists (*19*, *20*). It also includes information about the duration of sperm whales' surface and underwater phases; we approximated these from empirical data using Gaussian distributions $\mathcal{N}\left(\mu_{\text{surface}} = 9, \sigma^2_{\text{surface}} = 3^2\right)$ and $\mathcal{N}\left(\mu_{\text{underwater}} = 34, \sigma^2_{\text{underwater}} = 19^2\right)$, respectively. We used these values for the fielded experiments with sperm whales and the ablation study. For the experiments with the engineered whale, the model was obtained on the basis of the engineered whale's controlled behavior with the time duration of the underwater and surface phases.

Robots obtained acoustic and VHF AOA observations, whenever available, from sensors whose positions could be calculated. We considered these sensor measurements to have zero-mean Gaussian error, with SDs σ_{acoustic} and σ_{VHF}, respectively. Table S1 shows various parameters used in our experiments.

While running our autonomy algorithm in postprocessing, we considered the maximum speed of the robots to be $v_{\max} = 15$ m/s. The position of each whale was initialized in the horizontal plane using the location at their first surface end time. For the field experiments with the engineered whale, we used a custom VHF beacon that operated at 500 mW. It sent continuous signal pulses with a duration of 80 ms and an interpulse duration of 1000 ms. The engineered whale also relayed its ground-truth GPS locations in real time over a telemetry radio. Ground-truth positions for sperm whales were obtained from visual observations at the whales' surface end times from the catamaran CETI-1. These ground-truth positions were used when evaluating the autonomy and sensing modules.

Metrics for evaluation

A rendezvous was considered successful when a robot was within a predefined distance (rendezvous radius, $\rho_{\text{rendezvous}}$) from a surfaced whale; the robot needed to rendezvous with a whale only once. The performance of the autonomy module was evaluated on the basis of the following metrics. "Successful rendezvous percentage" denotes the ratio of the number of successful rendezvous to the total surfacing events for a group of whales until the end of mission, expressed as a percentage. "Mission time" denotes the total time from the start of the mission until the last whale was rendezvoused during the operation time horizon. "Missed whale frequency" denotes the normalized frequencies of the number of whales that were never visited during the total operation time.

Fielded validation in Dominica

We performed two experiments in the Caribbean Sea along the west coast of Dominica (Fig. 3A). These experiments required addressing challenges such as limited sensor availability and accuracy, tracking sperm whales for data collection, intermittent rain, varying temperatures, and wind speeds of 4 to 5.5 m/s with gusts of 6.5 to 9 m/s.

The first experiment used a speedboat named CETI-2 as an engineered whale with the VHF tag aboard (Fig. 3B). It emulated a sperm whale's motion by moving with fixed headings along a specific direction during each underwater phase. During this time, we collected acoustic AOA to it using a towed linear array of four hydrophones, with a 100-m-long cable connected to the passive acoustic monitoring software PAMGuard, which uses target motion analysis (*48*, *49*).

The VHF AOA was obtained by deploying our sensing UAV to emulate a virtual antenna array in real time (movies S1 and S2); this also validated the practicality of our payload design. Both the towed array and the UAV were deployed from CETI-1.

We evaluated the autonomy module by postprocessing measurements from acoustic AOA collected during the engineered whale's underwater phases and VHF AOA collected during the engineered whale's surface phases. The module then generated actions that would be furnished to a team of robots attempting to rendezvous with the whales. For the second experiment that involved sperm whales, we collected acoustic AOA to underwater whale vocalizations using the towed array for four untagged sperm whales and applied our autonomous routing in postprocessing.

Data collection for the engineered whale

We controlled the engineered whale to closely emulate a sperm whale's underwater and surface phases as observed by biologists (*20*). During the underwater phase, we continuously moved the engineered whale at a constant speed of 2.5 m/s. We collected acoustic AOA of its engine noise in lieu of a sperm whale's underwater vocalization for a total of 15 underwater phases.

The surfacing phase involved two behaviors. "Logging," which refers to surface resting, was achieved by stopping the engineered whale so that it drifted by 15 m on average over the entire duration of this phase. "Traveling" refers to active motion at the slowest speed of 0.75 m/s that resulted in a movement of 50 m on average during its surface phase. During this phase, we launched the UAV that collected data along for 40 s to compute real-time VHF AOA of the engineered whale's tag. The VHF AOA values were estimated for 12 surfacing phases.

We measured acoustic AOA only during the underwater phase and VHF AOA only during the surface phase because sperm whales often do not vocalize on the surface. Across the expeditions, we collected data from the engineered whale emulating five individual sperm whales in an area of 175 km^2 by overlaying the tracks together in postprocessing (Fig. 3C). Each track consisted of three surfacing and three underwater phases with the mean duration being 8.33 ± 8.89 min and 15.06 ± 5.29 min, respectively. During these experiments, the engineered whale's distance to CETI-1 varied between 65 and 285 m.

Evaluating AVATARS performance with the engineered whale data

Figure 4A shows a sample UAV trajectory during the AOA estimation process in varying wind speeds, as evident by the trajectory distortion. The corresponding AOA profile is shown in Fig. 4B; the AOA error was computed using the average ground-truth position of the engineered whale when the UAV was in flight. The AOA estimation (Fig. 4C) was not affected by the UAV's displacement because the SAR formulation uses the relative channel between the two antennas on the UAV, resulting in translation resilience; given a fixed separation between two antennas, the relative position vector only changes with rotation (*43*). For the logging behavior, we observed a median absolute error of 8.49° with an SD of 18.36° across eight samples. For the traveling behavior, we observed a median absolute error of 14.51° with an SD of 7.69° across four samples. Overall, we observed a combined median absolute error of 10.55° with an SD of 15.27°.

We applied our autonomy module by overlaying different tracks of the engineered whale over several days. The error in acoustic AOA (in degrees) was approximated using a Gaussian distribution $\mathcal{N}\left(0, 3.34^2\right)$, estimated using the acoustic sensor and ground-truth data. The error in VHF AOA was based on measurements from the UAV flights. We leveraged our field logs to determine the

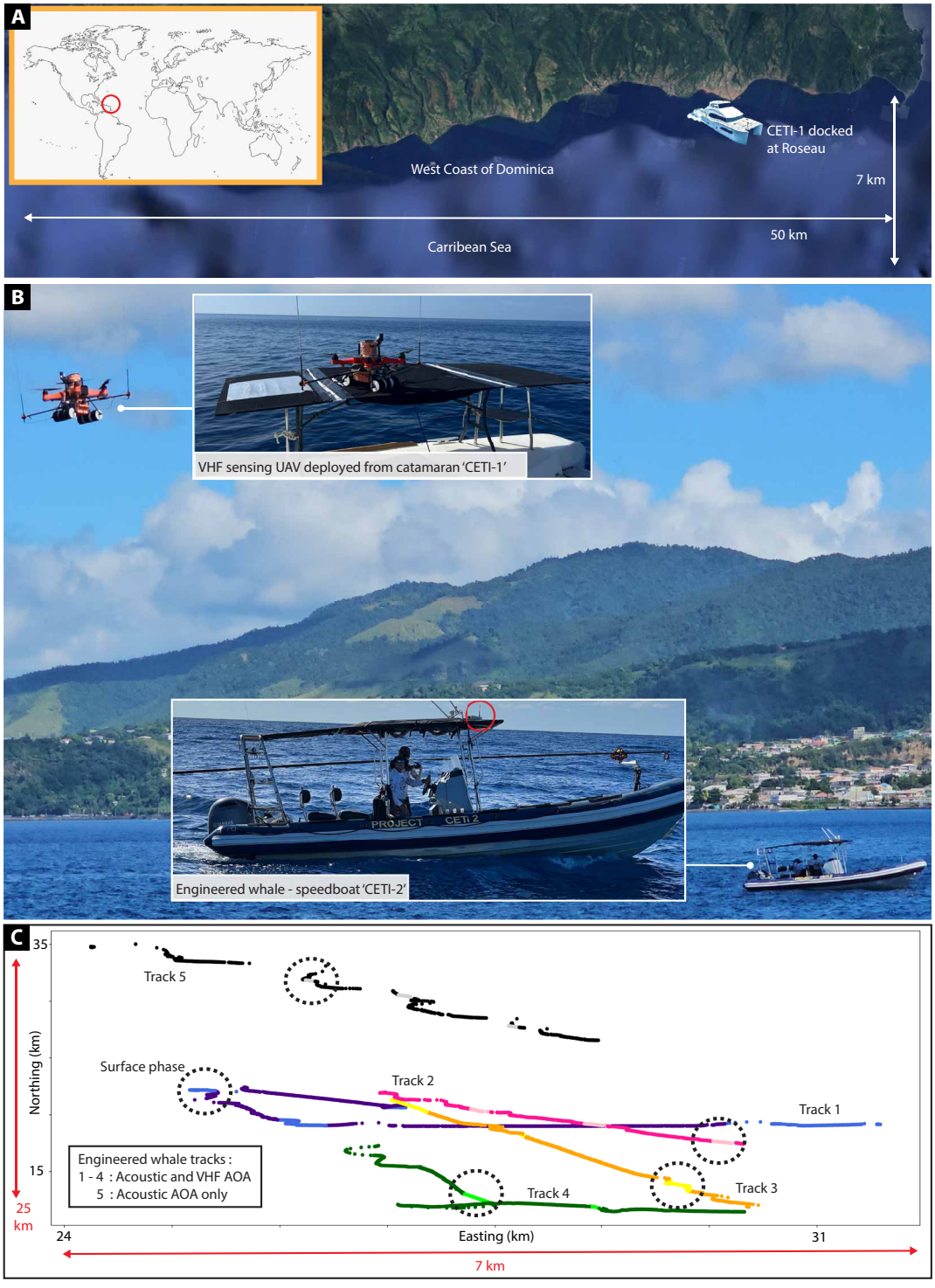

Fig. 3. Field experiment of system validation with the engineered whale. (**A**) The location of data collection and experiments in the Caribbean Sea, 7 km off the west coast of Dominica. (**B**) Instance showing the UAV, deployed from the catamaran CETI-1, estimating AOA of the VHF signal tag on the engineered whale, speedboat CETI-2, emulating a sperm whale's motion. (**C**) Overlaid reconstructed tracks of the engineered whale in postprocessing using acoustic and VHF AOA. A lighter shade denotes the surfacing phase (also indicated by the dashed circles), and a darker shade denotes the underwater phase.

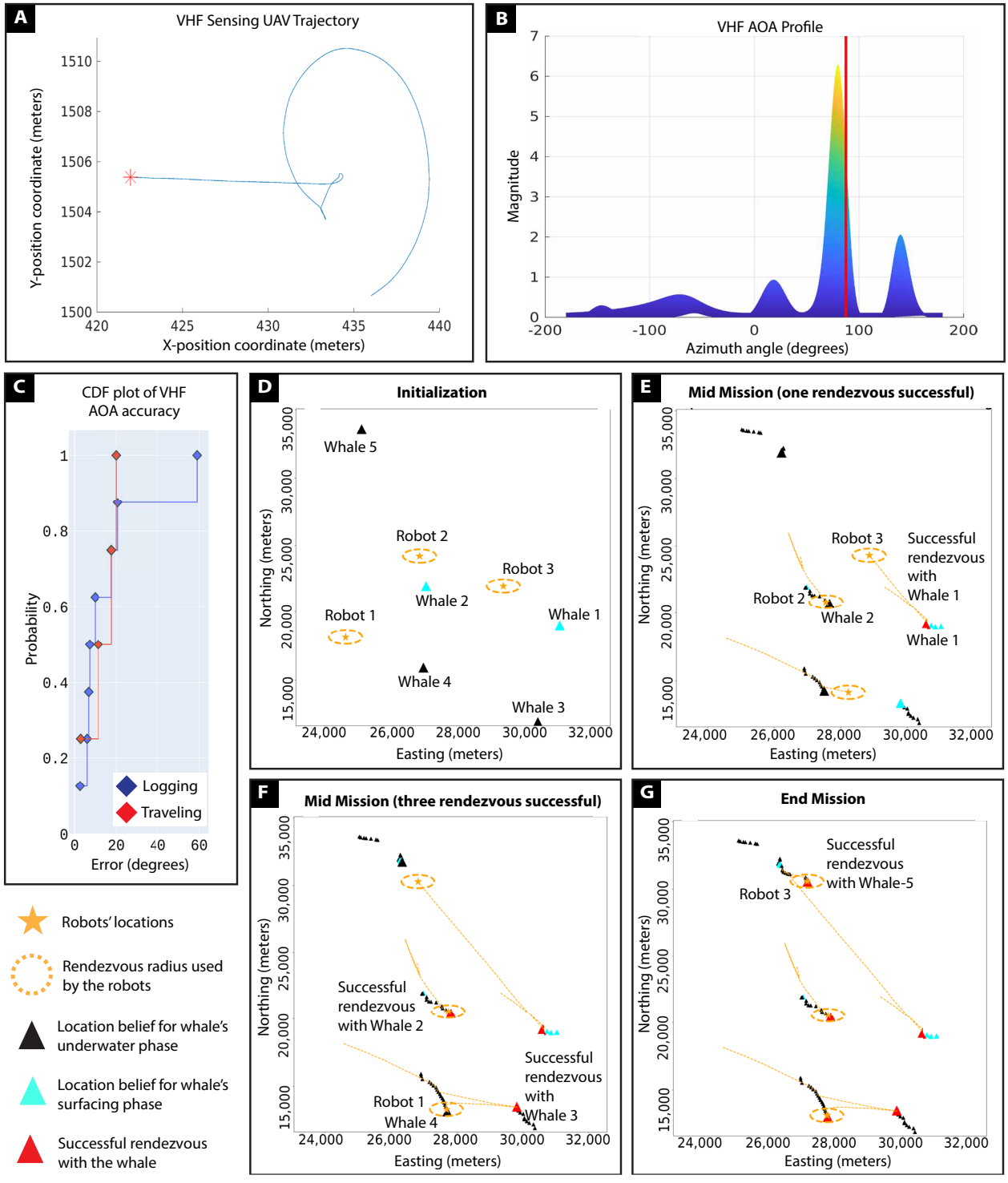

Fig. 4. Results of experiments with the engineered whale. (**A**) The UAV's trajectory sample from field experiments; star indicates start position. (**B**) Corresponding VHF azimuth AOA estimation with ground truth indicated by the red line. The accuracy was preserved despite trajectory distortion due to wind because of the translation-resilient property of our formulation. (**C**) Real-time VHF AOA accuracy for different surface behaviors of the engineered whale. (**D**) An instance of the postprocessing showing an initialization of three robots (stars) and five engineered whales (triangles). Robots used a rendezvous radius of 500 m. (**E**) Robot 3 achieved rendezvous with whale 1 and moved toward whale 5 while robot 2 continued to track whale 2. (**F**) Robot 1 tracked whale 4 after rendezvous with whale 3 and robot 2 rendezvous with whale 2. (**G**) The end of the mission shows a successful rendezvous with all the engineered whales.

engineered whale's transition between underwater and surface phases. At each time step, the localization algorithm estimated a whale's horizontal position using the AOA collected from the acoustic and VHF sensors, with a mean localization error of 224.04 ± 180.95 m. The simulated autonomous robots attempted to rendezvous with the whales while moving at a maximum speed of 15 m/s (velocity of a speedboat used by biologists when tracking sperm whales in the field) for rendezvous radii of 200, 300, and 500 m. We generated aggregate results for different numbers of robots over 50 runs each (see Fig. 6A and table S5). For each run, the initial locations of the autonomous robots were randomized within the operation region and the start time of the individual whale tracks were randomized within their first underwater phases. We observed that a radius of 500 m achieved the highest successful rendezvous percentage of 80.4% ± 10.7% for four engineered whales and two autonomous robots. Figure 4 (D to G) shows another instance where three robots achieved successful rendezvous with five engineered whales. Different science missions may need different rendezvous radii. A smaller rendezvous radius may be needed to support short-range complex operations, such as whale tagging (*50*), and a larger rendezvous radius can be used for long-range tasks such as visual data collection. Our results demonstrated the effectiveness of using the AVATARS sensing module to attain smaller rendezvous radii.

Data collection for sperm whales

We collected acoustic data during the underwater phases from three different sperm whale groups, referred to as units henceforth. Two groups were from well-studied social units unit F and unit R (*20*, *51*). We collected data for two instances from the whale Pinchy (whale #5560) of unit F (3.5 hours with five surfacings and 2 hours with three surfacings), one instance from unit R (2 hours with three surfacings), and one instance of a whale from an unidentified unit (1.5 hours with two surfacings). Although there has been some progress toward identifying individual whales solely on the basis of whale vocalization measurements (*52*), it is still an open challenge. Hence, we only considered instances where it was known with high probability that the acoustic data were captured from a single whale. An exception was made for the data collected for unit R, where four whales were observed socializing in close proximity of one another (within 10 m). Because the whales made several shallow dives while traveling together and maintained proximity throughout their underwater and surfacing phase, we considered the corresponding acoustic AOA measurements to be obtained from a single whale. Details about ground-truth estimation of whale positions are provided in the Supplementary Materials.

Evaluating AVATARS autonomy module with sensor data from sperm whales

Because the sperm whales were untagged, we computed their ground-truth positions during surface end time on the basis of the photos of their fluke (tail) (Fig. 5A). We collected acoustic AOA during their underwater phases and filtered the raw measurements by applying a set of selection criteria (Fig. 5, B and C). Complete details of these steps are provided in the Supplementary Materials.

We collected the four whale tracks in an area of 720 km^2 and overlaid them in postprocessing (Fig. 5D). A mean localization error of 408.12 ± 171.15 m was observed for these tracks (fig. S3 shows individual errors).

We used rendezvous radii of 500 m, 1000 m, and 1500 m when evaluating our autonomy module in postprocessing. We observed a successful rendezvous percentage of 70.4% ± 7.1% for a combination of three whales and two robots with a rendezvous radius of 1500 m, aggregated over 50 runs. Figure 6, A and B, shows the aggregated successful rendezvous percentage and mission time, and Fig. 6C shows the missed whale frequency. The start locations of the rendezvous robots were randomized within the operation region for each run. Results with three whales were aggregated over all possible combinations of three out of four whale tracks.

From these experiments, we observed that the autonomy module explicitly allows for adapting to varying dive duration on different days or even during different times on the same day. For example, the whale from unit F was swimming with prototypical dive duration, closely following the whale dive model (*20*), whereas the one from unit R had longer dive periods.

Ablation study for AVATARS autonomy module

Simulated sensor setup and evaluation

We simulated an acoustic sensor and a VHF sensor that moved northbound with a speed of 2.57 m/s, starting from 3.2 and 6.4 km, respectively, off the shore of the dock located at 15.3092° latitude and −61.3794° longitude. Such a strategy is often exercised when tracking whales using the towed array along Dominica's coastline by project CETI's team. We interpolated fluke sightings in the DSWP dataset to simulate whales' pseudo-tracks (see the Supplementary Materials) and simulated AOA observations using the pseudo-tracks. The VHF and acoustic sensors sampled AOA measurements for a whale's surfacing and underwater phases with Gaussian errors in degrees of $\mathcal{N}(0, 2^2)$ (following best-performing field experiments) and $\mathcal{N}(0, 3^2)$, respectively. We also simulated a more accurate sensing, such as using multiple hydrophone buoys with underwater localization error in meters of $\mathcal{N}(0, 200^2)$ and specialized GPS tags with localization error in meters of $\mathcal{N}(0, 10^2)$ when the whale surfaces. The acoustic sensing errors are based on (*53*). For each run, we uniformly randomly selected 50 sets of four whales out of 334 whale sightings from the DSWP dataset. For each whale, we randomly chose 2 hours' worth of data. We reported successful rendezvous percentage and mission time, aggregated over these 50 runs, each containing four whales with two robots. The start locations of the robots were randomized within the operation region. We report the results of the autonomy module applied over a horizon of 120 min in postprocessing.

To test the effect of sensing modality on the AVATARS autonomy module's performance, we used different combinations of simulated sensors that resulted in varying localization accuracy (Fig. 7A). Incorporating VHF AOA with acoustic AOA improved the results of our approach. We also observed an improvement in performance resulting from reduced localization error due to better sensor configuration. However, we note that the availability of such accurate localization will require higher cost and more power requirements than low-power VHF tags.

Next, we studied the effect of varying rendezvous radii on our autonomy module using the values of 200, 300, and 500 m for the rendezvous radius. These values showed that a higher rendezvous radius is required to improve success when using sensors with high error, such as acoustic AOA-only sensing, as opposed to sensors with low localization error, including GPS localization. For a 500-m rendezvous radius with acoustic and VHF AOA sensing, our approach achieved successful rendezvous with 53.03%.

Last, we performed a comparison study of our approach with other routing algorithms. Because our algorithm is an online approach, we compared it against two other widely used online methods, namely, the instantaneous assignment (IA) method inspired by the multitarget

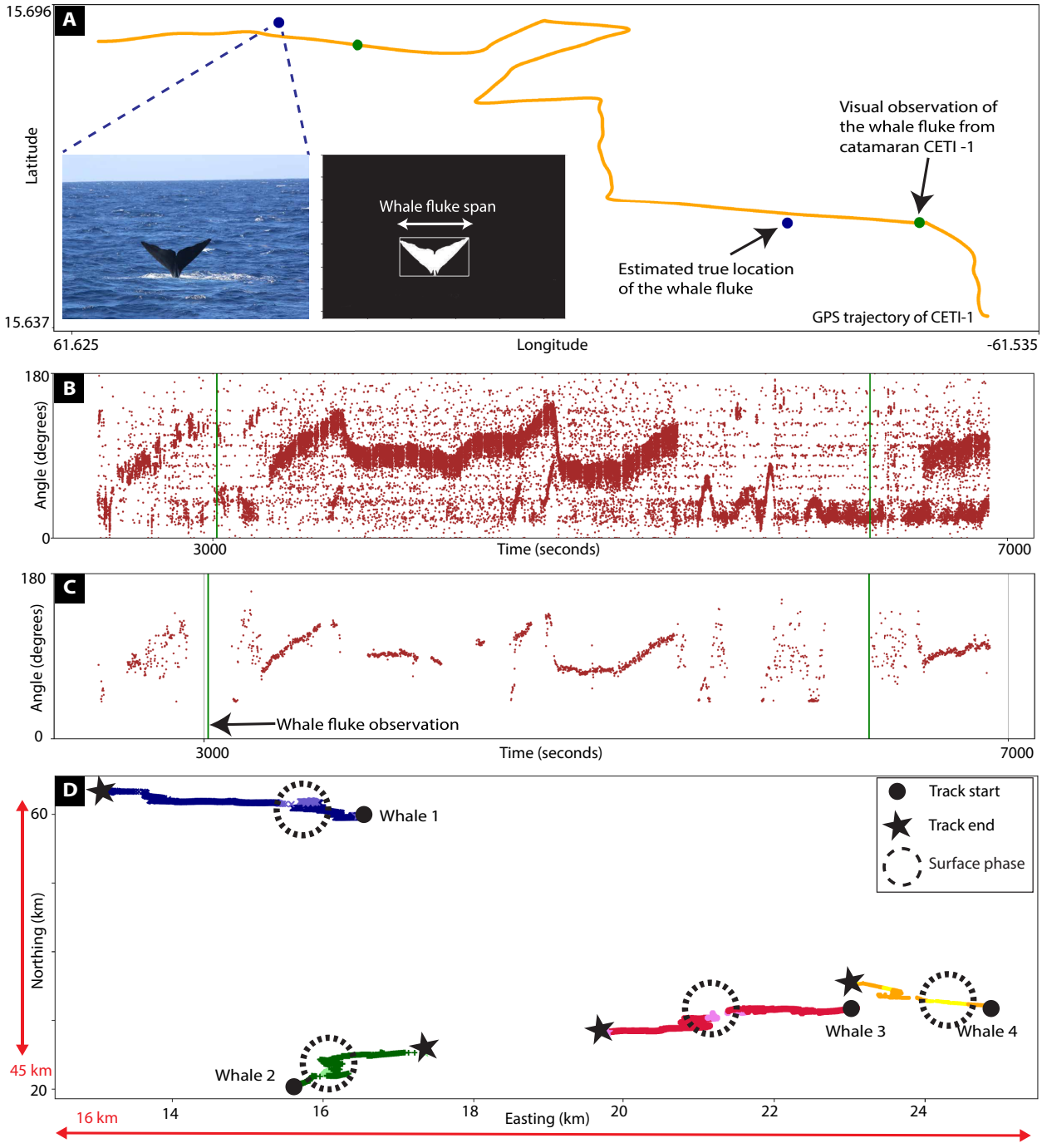

Fig. 5. Data collection for sperm whales. (**A**) A sample GPS trajectory of the catamaran CETI-1 between two consecutive whale surfacings. The ground-truth position of a whale was calculated using range (derived from the fluke/tail photo and the actual tail size) and bearing (from the camera's internal compass). (**B**) Raw acoustic AOA data samples from a sperm whale. (**C**) Cleaned acoustic AOA data samples from a sperm whale after applying the data selection criteria explained in the Supplementary Materials. (**D**) Whale tracks were estimated using UKF-based localization and acoustic AOA measurements. A lighter shade denotes a surfacing phase (also indicated by the dashed circles), and a darker shade denotes an underwater phase.

observation problem (Fig. 7B) (*54*) and the vehicle routing problems with time window constraints (VRP-TW) (Fig. 7C) (*55*, *56*). Given that these methods do not account for future uncertainties, our autonomy module outperformed them across all metrics because we considered the stochastic nature of whale behavior over a future planning horizon. Implementation details of IA and VRP-TW are provided in Materials and Methods.

Ablation study for the sensing module

Experiments with off-the-shelf VHF tags

Off-the-shelf passive VHF tags have been used for wildlife tracking because of their low power usage and long battery life (*57*, *58*). Our results show compatibility of our SAR-based approach for such a tag with a 20-ms pulse duration and a 1100-ms interpulse duration. The

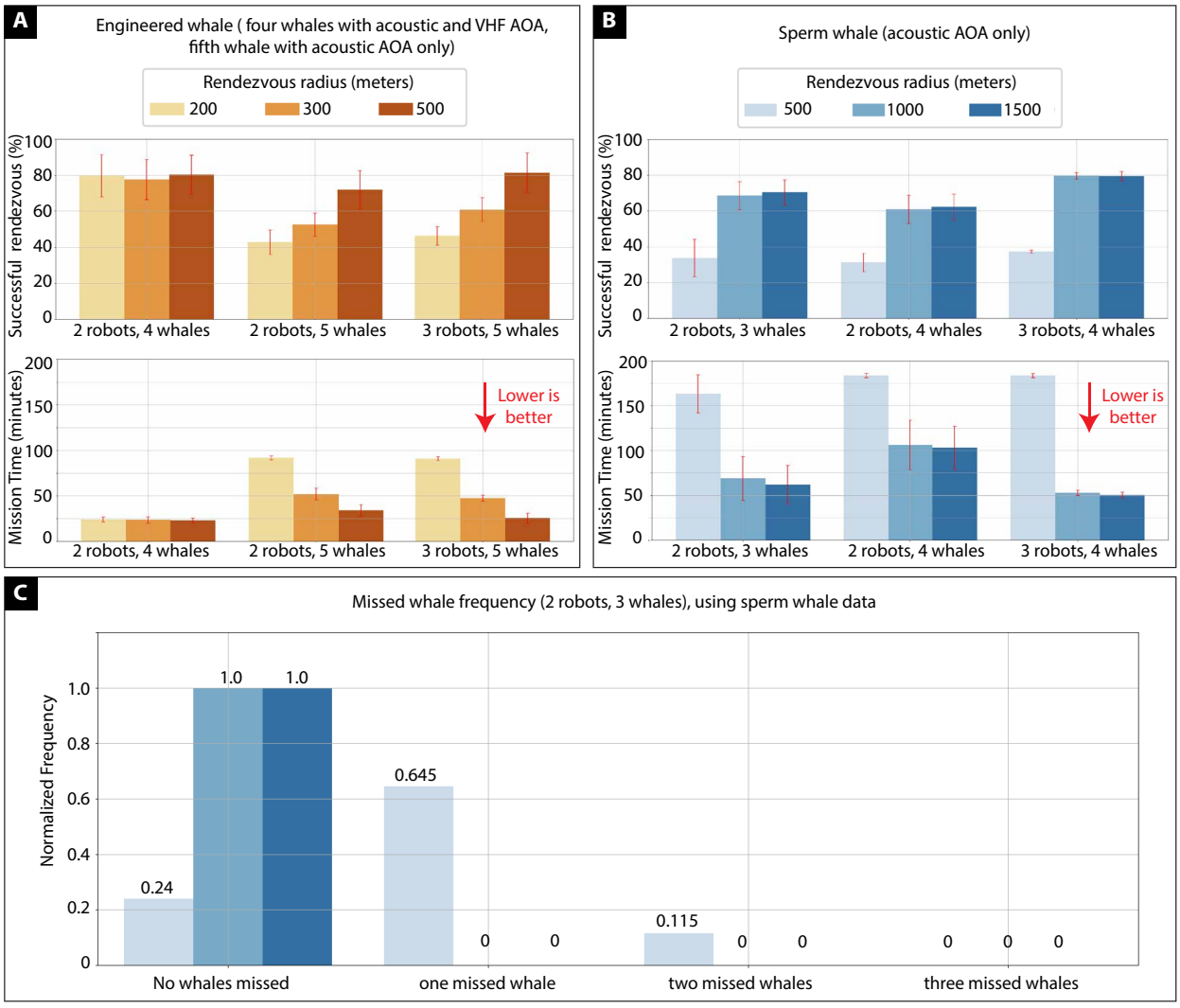

Fig. 6. Aggregated results for postprocessed data collected during field experiments. Results for the three metrics—successful rendezvous, mission time, and missed whale frequency—were generated with randomized initial locations of the robots within the operation region and the start time of individual tracks of the whales within their first underwater phases. The bars in engineered whale results in (**A**) and sperm whale results in (**B**) indicate means, and error bars indicate SD over 50 samples. For results with three sperm whales, we used all possible combinations of the four sperm whale tracks. (**C**) For the sperm whale experiment, the normalized frequency of the number of whales that were missed within the mission time. In the sperm whale experiments, we used the DSWP dataset to estimate the future surface minutes spent $[\mathcal{N}(9, 3^2)]$ and the future underwater minutes spent $[\mathcal{N}(34, 19^2)]$, and, in the engineered whale experiments, we used the engineered whale's engine start and stop times to estimate the future surface minutes spent $[\mathcal{N}(8, 9^2)]$ and the future underwater minutes spent $[\mathcal{N}(15, 5^2)]$.

experiments were performed at Horn Pond, Massachusetts (Fig. 8, A and B). We used an off-the-shelf tag and our custom tag during these experiments and compared their performance. The transmitter antennas of the VHF tags were not more than 40 cm above the water surface. Data to compute VHF AOA were collected for 40 s, during which the tags drifted by 2.5 m on average. During the experiments, the distance between the tag and the receiver was maintained between 100 and 200 m. We collected 17 samples for our custom tag and rejected two samples as per our data selection criteria (see the Supplementary Materials). We observed a median error of 5.59° with an SD of 7.52° for the custom tag. For the off-the-shelf tag, we collected eight samples and rejected three samples and observed a median error of 4.41° with an SD of 7.99° (Fig. 8C).

Experiments with a stationary tag in ideal conditions

We demonstrated VHF AOA accuracy for a stationary VHF tag in the Ohiri field at Harvard University (fig. S5). The custom tag was located ~150 m away and transmitted a signal every second. We chose three random locations for the tag to obtain aggregate results and placed the tag at a height between 0.5 and 1 m to avoid interference from the ground surface. We collected 34 samples in total and rejected two samples as per our data selection criteria. The median error for the 32 samples was 7.11° with an SD of 5.05°. Additional comparison details are provided in the Supplementary Methods.

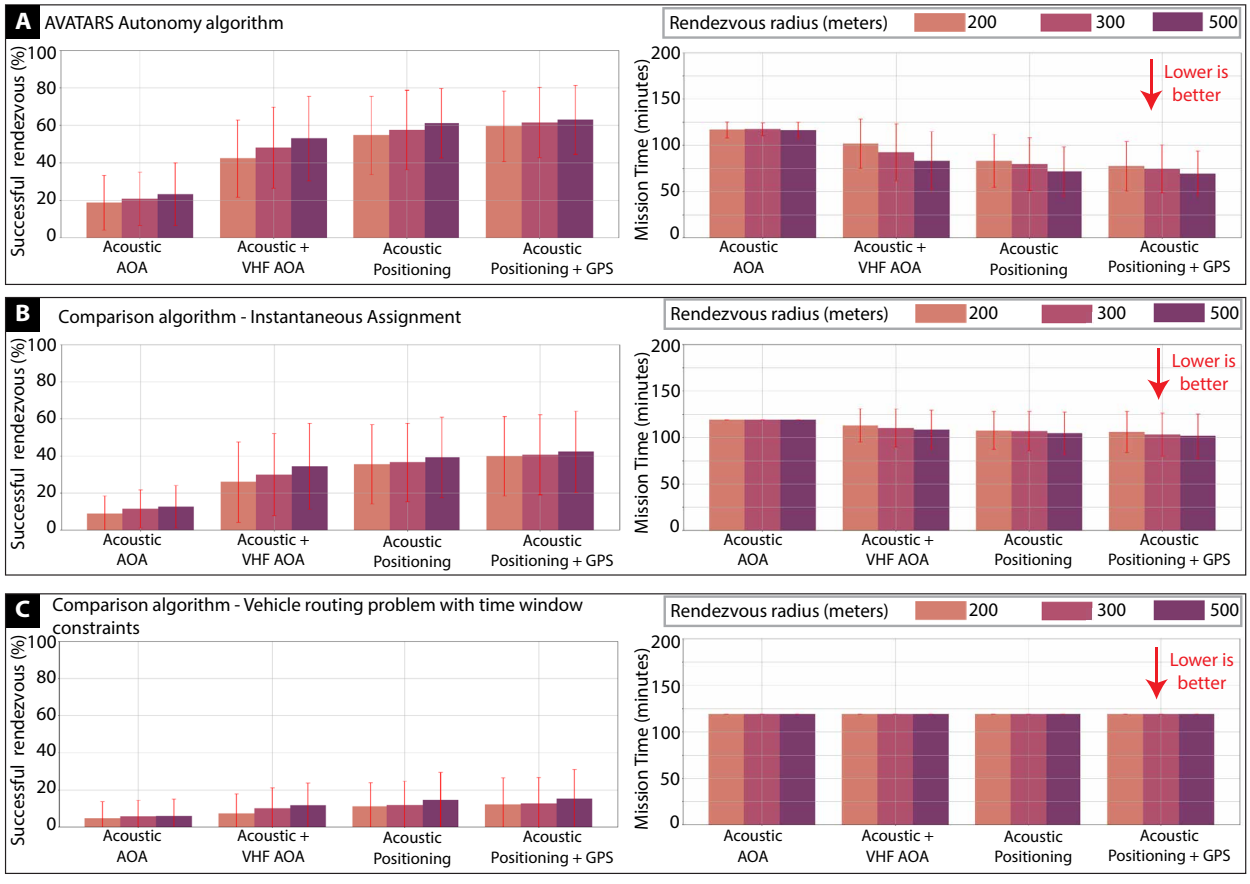

Fig. 7. Results of ablation and comparison study for autonomy module. Results of successful rendezvous percentage and mission time (in minutes) with two autonomous robots and four simulated whales using the DSWP dataset (*20*) for various sensing modalities and rendezvous radii. Results show means (bar plots) and SDs (error bars) for different metrics generated over 50 trials with randomized initial locations of the robots within the operation region and randomized the start time of individual tracks of the whales within their first underwater phases. The figure shows the performance of (**A**) our autonomy algorithm, (**B**) IA, and (**C**) VRP-TW. We used the DSWP dataset to estimate the future surface minutes spent [$\mathcal{N}(9, 3^2)$] and the future underwater minutes spent [$\mathcal{N}(34, 19^2)$].

DISCUSSION

This work codevelops an algorithmic framework that combines an RL-based autonomous routing algorithm and VHF signal–based sensing to maximize the chances of rendezvous with whales at sea. The primary challenges in developing this framework involve sparse and uncertain in situ AOA measurements, as well as learning from and adapting to whale position and surfacing variability. In addition, we developed a UAV sensor payload to enhance AOA measurement quality from passive VHF tags using a signal phase–based sensing algorithm specifically designed for maritime operations. Our experiments were conducted over three separate expeditions in Dominica. The experiments included the real-time deployment of the sensing module, sensor data collection from both engineered and sperm whales, and the postprocess validation of rendezvous performance by the autonomy module. In addition, we conducted various ablation studies. These experiments validated the performance of our framework in different environments, with different numbers of sperm whales, and with different configurations of sensors (using acoustic and VHF AOA versus using acoustic AOA only).

Our results demonstrate 68.68% rendezvous success for the configuration of two robots and three sperm whales, with a rendezvous radius of 1000 m using only acoustic AOA. For the experiments with the engineered whale mimicking different whale surface motions, such as logging and travelling, we show a median accuracy of 10.55° for VHF AOA and an 80.4% rendezvous success rate for a rendezvous radius of 500 m, using AOA from acoustic and our VHF-sensing module for a configuration of two robots and four engineered whales. We report rendezvous success for several different radii. Our field experiments demonstrate the AVATARS ability to obtain accurate real-time AOA to VHF tags. Using this information improves rendezvous success with a tighter rendezvous radius in postprocessing.

Our experiments also reveal several open challenges. The sensing module is limited in detection range for commercial tags, and improvements to signal pulse detection methods will enable low power signal pulse extraction at longer distances (*29*). The real-time deployment of the autonomy module with a robotic team to rendezvous with sperm whales necessitates several infrastructure and operational setups. These include tagging of whales and real-time AOA computation capabilities of underwater acoustic sensors. In addition, the system must support the streaming of acoustic AOA data from underwater sensors to the robots. A mesh communication network among the robots will be necessary to enable real-time coordination. Furthermore, the robots must share a common belief regarding whale positions to facilitate distributed computation. Last, a fleet of physical

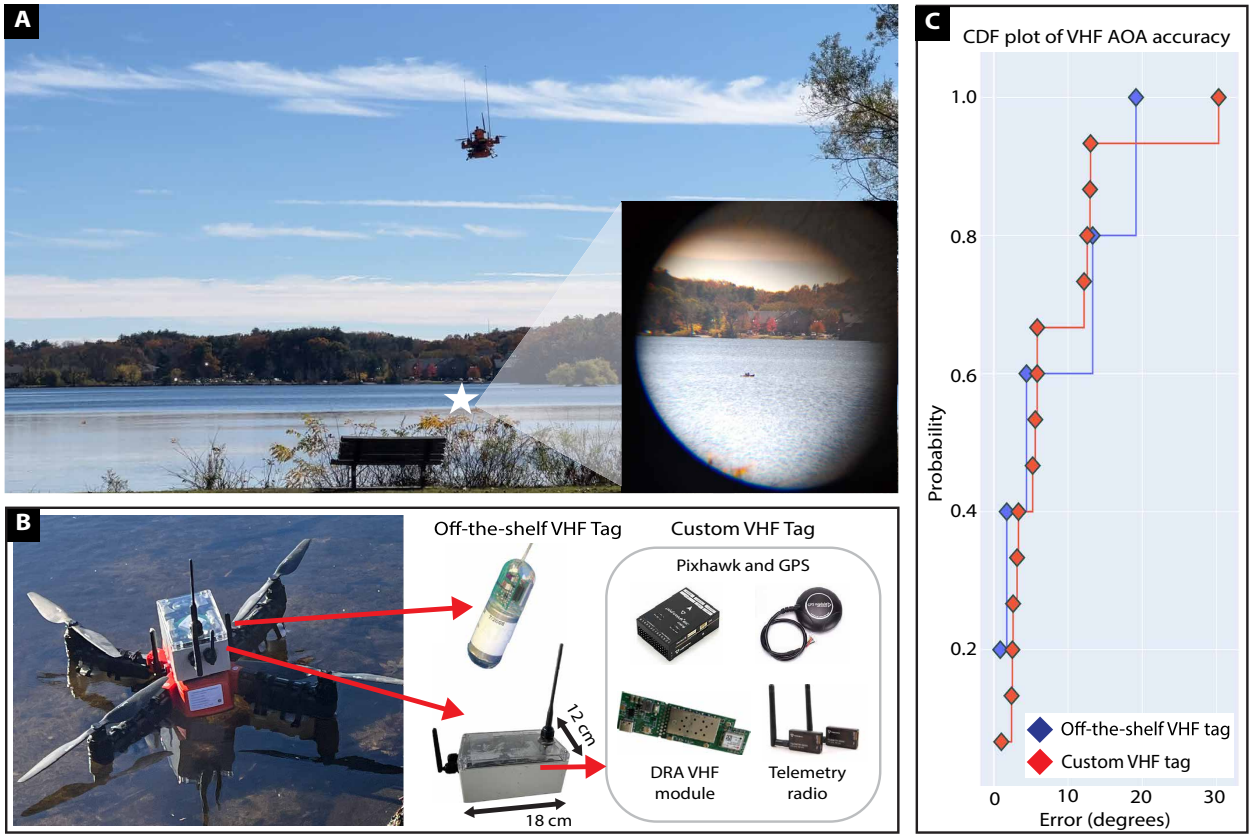

Fig. 8. Results of ablation study for sensing module. (**A**) Hardware experiment showing an instance of AOA estimation for floating VHF signal beacons approximately 150 m away from the bank tested at Horn Pond, Boston, Massachusetts. White star indicates floating UAV. (**B**) Floating UAV SwellPro fisherman max, with our custom tag and off-the-shelf tag from ATSTrack. The custom tag includes a Pixhawk, a GPS module, and a telemetry radio for obtaining position of the VHF tags to enable real-time ground-truth data collection. (**C**) Cumulative distribution function (CDF) plot comparing VHF AOA estimation performance for different tags used in this experiment.

robots certified for maritime operations and capable of long operational duration would be essential for successful deployment. Some of these efforts are already underway as part of a larger effort by Project CETI (*10*, *11*, *16*, *59*) but are out of scope for the current work. The focus of this paper was to develop and validate the critical algorithmic pieces of the AVATARS framework, thus taking the initial steps toward achieving autonomous data collection for Project CETI.

We also learned critical lessons when transitioning from the lab to the field. We incorporated multiple stages of outdoor experiments with increasing complexity, including experiments at a local pond and flight tests in the Boston harbor. This resulted in efficient preparation for field experiments regarding issues such as equipment setup and maintaining hardware structural integrity, among others, as given in table 1 of (*60*). A unique challenge that we encountered was acoustic AOA acquisition. The towed array used for the experiments lacks an inertial or positional sensor and accumulated more noise than a dedicated underwater buoy for acoustic tracking because of its drag in the ocean and the engine noise of the towing boat. These issues made the acoustic AOA measurements challenging and required that we apply the data selection criteria. The length of the towed array made maneuvering the boat challenging, and the availability of CETI's marine operations and biology team, with prior experience in whale science, was instrumental in the process of data acquisition. We also had to ensure that the boat captain was on a constant lookout for neighboring fishing or tourist boats and proactively communicated with them about our use of the array when in their proximity. Simultaneous deployment of the towed array and the UAV required assigning specific roles to all crew members, with everyone maintaining situational awareness. Additional details about how we addressed these unique challenges are provided in the Supplementary Methods.

The AVATARS framework presented here provides a critical step along the path of enabling robots to assist in capturing visual data for whale science at sea. A long-term vision for this work is one in which teams of UAVs can be continually and autonomously deployed from depots stationed at sea whenever the likelihood of a whale surfacing and rendezvous is high. This will enable whale tagging and/or capturing visual data fully autonomously and across spatially distant areas of the environment. In this way, we hope that this work will help our efforts to arrive at a future where robots are our science "avatars" at sea, allowing for scientific discovery with less manual effort than what is possible today.

MATERIALS AND METHODS

AVATARS autonomy module

We formulated a finite horizon partially observable Markov decision process (POMDP) with a hybrid discrete-continuous belief space that included discrete whale surfacing phases and continuous whale locations. The whale locations were estimated using a Gaussian

distribution, obtained from the output of an unscented Kalman filter (UKF) inspired from (*61*) (see details in the Supplementary Materials). The UKF used the latest AOA observations from heterogeneous sensors to update its predictions at each time step. To solve the POMDP, we implemented a model-based RL approach, namely, multiagent rollout [inspired by (*32*, *37*)], that performed a sequence of optimizations, one per robot, in a given order, scaling linearly in the number of agents. In each one-step lookahead optimization (Eq. 1), a robot was assigned to a whale for rendezvous, such that the expected cost for each possible whale assignment was minimized, thereby increasing the rendezvous success percentage. The expected cost for an assignment (Eq. 2) includes both the time to rendezvous with the currently assigned whales and the time to rendezvous with whales in subsequent assignments while accounting for future whale location and surfacing uncertainties. The expectation was calculated via several Monte Carlo simulations, sampling future whale locations and surfacing times from the whale dive model estimated from the DSWP dataset (*20*). After a successful rendezvous with a whale in the current assignment, we applied binary integer programming (BIP; Eq. 4) at a future time step in the planning horizon to assign robots to the subset of the remaining whales without considering subsequent assignments. In other words, the BIP acted as a myopic policy that is easily obtainable for a given surfacing interval, which our rollout approach further improves by looking into future uncertainties using the look-ahead optimization. Figure 2 (A and B) shows various components of our autonomy module. Figure S2 shows a detailed view of the rollout optimization for a robot in the multirobot setup.

Input and assumptions

Our robot model assumed that each robot knows all robots' positions. We considered homogeneous robots with the same maximum speed $v_{\max}$, operating time as long as the total operation time, and a finite rendezvous radius $\rho_{\text{rendezvous}}$. A successful rendezvous happens when a whale is at the surface and within $\rho_{\text{rendezvous}}$ distance from a robot. Robots could receive the initial position of a whale in the horizontal plane with a zero mean Gaussian error, which is used to initialize the UKF for each whale's location. Robots could determine whether a whale is on the surface, either from acoustic data and if the whale is tagged, from the VHF signal.

We used Gaussian distributions to estimate future uncertainties in surfacing times in the planning horizon of the rollout optimization. In particular, the time spent by a whale during its surface and underwater phases were estimated using Gaussian distributions $\mathcal{N}\left(\mu_{\text{surface}}, \sigma^2_{\text{surface}}\right)$ and $\mathcal{N}\left(\mu_{\text{underwater}}, \sigma^2_{\text{underwater}}\right)$, respectively. Robots obtained acoustic and VHF AOA observations, whenever available, from sensors whose positions could be calculated. We considered these sensor measurements to have zero-mean Gaussian error, with SDs σ_{acoustic} and σ_{VHF}, respectively.

Belief state formulation

The belief state of m autonomous robots and n whales at time k is denoted by x_k, which is represented as a set of S belief particles. The sth belief particle at time k is denoted by $x_{k,s} = [k, \mathbf{d}_s, \mathbf{w}_s, \mathbf{z}_s, \mathbf{I}_s, \mathbf{r}_s]^T$, where $\mathbf{d}_s = \left[d^1_s, \ldots, d^m_s\right]$, where $d^\ell_s \in \mathbb{R}^2$ is robot ℓ's position in the horizontal plane; $\mathbf{w}_s = \left[w^1_s, \ldots, w^n_s\right]$, where $w^q_s \in \mathbb{R}^4$ is whale q's location and velocity in the horizontal plane obtained from the output of the UKF-based localization using AOA measurements; $\mathbf{z}_s = \left[z^1_s, \ldots, z^n_s\right]$, where z^q_s is whale q's surface indicator for belief particle s and is set to 1 if whale q is on the surface at time k and is set to 0 otherwise; $\mathbf{I}_s = \left[I^1_s, \ldots, I^n_s\right]$, where $I^q_s = \left\{\left[a_1, b_1\right], \left[a_2, b_2\right], \ldots, \left[a_T, b_T\right]\right\}$, where a_i and b_i are the start and end of the ith surface interval sampled at time k and T is the number of intervals sampled over the planning horizon for whale q's sth belief particle. Here, $r_s = [\mathrm{r}^1{}_s, \ldots, \mathrm{r}^n{}_s]$, where $\mathrm{r}^q{}_s$ is whale q's rendezvous status for belief particle s and is set to 1 if a robot has successfully rendezvoused with whale q at any time between 0 to k; otherwise, $\mathrm{r}^q{}_s$ is 0.

Multiagent rollout policy

The multiagent rollout policy (outlined in Fig. 2 and fig. S2) finds an assignment $\boldsymbol{\Psi}_{\text{rollout}}(x_k) = \left[\Psi^1_{\text{rollout}}, \cdots, \Psi^m_{\text{rollout}}\right]$ for a belief x_k by solving m optimization problems in a sequence. Here, $\Psi^\ell_{\text{rollout}} \in \{1, \ldots, n\}$ is the whale assigned to robot ℓ, which is obtained by applying a look-ahead minimization over all possible whales. Robot ℓ leverages the knowledge of the whales assigned to robots 1, …, $\ell-1$ using rollout optimization earlier in the sequence. Robot ℓ assumes that robots $\ell+1$, …, m that come later in the sequence are assigned to whales using a base policy given by a BIP. More formally, starting from robot $\ell = \{1, \ldots, m\}$, the rollout assignment is obtained as follows

$$\Psi^\ell_{\text{rollout}} \in \arg\min_{q\in\{1,\ldots,n\}} \mathbb{E}_s \left[J\left(x_{k,s}, \left[\Psi^1_{\text{rollout}}, \ldots, \Psi^{\ell-1}_{\text{rollout}}, q, \Psi^{\ell+1}_{\text{BIP}}(x_k), \ldots, \Psi^m_{\text{BIP}}(x_k)\right]\right)\right] \quad (1)$$

Here, $\Psi^1_{\text{rollout}}, \cdots, \Psi^{\ell-1}_{\text{rollout}}$ are the whales assigned to robots 1, …, $\ell-1$ by their corresponding rollout optimizations, respectively, and $\Psi^\ell_{\text{BIP}}(x_k) \in \{1, \ldots, n\}$ is the whale assigned to robot $\bar{\ell} = \{\ell+1, \ldots, m\}$ using the BIP policy (Eq. 6).

The cost (in Eq. 1), denoted by $J(x_{k,s}, \boldsymbol{\Psi})$, represents the time taken to rendezvous with the first whale in the current assignment $\boldsymbol{\Psi} = [\Psi^1, \ldots, \Psi^m]$, starting at belief particle $x_{k,s}$, followed by the time taken to rendezvous with whales in future assignments. The future assignments for a given belief particle are given by BIP policy (Eq. 4). Mathematically,

$$J(x_{k,s}, \boldsymbol{\Psi}) \begin{cases} \overset{(1)}{=} 0 & \text{if all whales are rendezvoused or the planning time horizon has elapsed,} \\ \overset{(2)}{=} \sum_q \left(1 - r^q_s\right) + J\left(f\left(x_{k,s}, \boldsymbol{\Psi}\right), \boldsymbol{\Psi}\right) & \text{if no rendezvous happens at time } k, \\ \overset{(3)}{=} J\left(x_{k,s}, \boldsymbol{\Psi}_{\text{BIP}}(x_{k,s})\right) & \text{otherwise.} \end{cases} \quad (2)$$

In the cost function (Eq. 2), case (2) adds a unit cost for each whale that has not been rendezvoused yet and a recursive cost associated with unvisited whales in the planning horizon. Case (3) expresses the cost of the BIP assignment (Eq. 4) once at least one whale in the current assignment $\boldsymbol{\Psi}$ is successfully rendezvoused.

The state transition function f used in the rollout optimization (Eq. 2) evolves belief particle $x_{k,s}$ to the next belief particle over the planning horizon by applying the whale assignments for all robots denoted by $\boldsymbol{\Psi} = [\Psi^1, \ldots, \Psi^m]$ is

$$f\left(x_{k,s}, \boldsymbol{\Psi}\right) = f\left(\begin{bmatrix} k \\ \mathbf{d}_s \\ \mathbf{w}_s \\ \mathbf{z}_s \\ \mathbf{I}_s \\ \mathbf{r}_s \end{bmatrix}, \boldsymbol{\Psi}\right) = \begin{pmatrix} k+1 \\ \mathbf{d}_s + \left[\mathbf{v}\cos(\boldsymbol{\phi}), \mathbf{v}\sin(\boldsymbol{\phi})\right]^T \\ A\mathbf{w}_s \\ \left[f^1_z\left(x_{k,s}, \boldsymbol{\Psi}\right), \ldots, f^n_z\left(x_{k,s}, \boldsymbol{\Psi}\right)\right] \\ \mathbf{I}_s \\ \left[f^1_r\left(x_{k,s}, \boldsymbol{\Psi}\right), \ldots, f^n_r\left(x_{k,s}, \boldsymbol{\Psi}\right)\right] \end{pmatrix} \quad (3)$$

where the distance moved by robots in unit time $\mathbf{v} = [v^1, \ldots, v^m]$ and heading angle $\boldsymbol{\phi} = [\phi^1, \ldots, \phi^m]$ were obtained from assignment $\boldsymbol{\psi}$. Robot ℓ's assignment ψ^ℓ determines the heading from the robot's location d_s^ℓ to the whale's belief location $Lw_s^{\psi^\ell}$, denoted by $\phi^\ell = \angle\left(d_s^\ell, Lw_s^{\psi^\ell}\right)$, and velocity $v^\ell = \min\left\{v_{\max}, \|Lw_s^{\psi^\ell} - d_s^\ell\|\right\}$. Here, matrix $L = [1,0,0,0; 0,1,0,0]$, and $Lw \in \mathbb{R}^2$ extracts the location in the horizontal plane from the whale's location and velocity vector $w \in \mathbb{R}^4$. The bearing $\angle(s, d)$ gives the angle between the location d as calculated from location s. Here, whale motion matrix in the horizontal plane is $\boldsymbol{A} = [1,0,1,0; 0,1,0,1; 0,0,1,0; 0,0,0,1]$ assuming constant velocity and a unit time step. Here, $f_z^q(x_{k,s}, \boldsymbol{\psi})$ is an indicator, set to 1, if the planning time step $k+1$ belongs to an interval in the set of surface intervals I_s^q, or $\exists\,[a, b] \in I_s^q, k+1 \in [a, b]$, and it is set to 0 otherwise. Notably, in Eq. 3, the set of surface intervals $\boldsymbol{I}_s$ for belief particle $x_{k,s}$ does not change in the planning horizon. The interval set $\boldsymbol{I}_s$ only gets updated during the policy execution, after an observation regarding a whale surfacing phase is made (see the Supplementary Materials). In Eq. 3, $f_r^q(x_{k,s}, \boldsymbol{\psi})$ is the rendezvous status of whale q at planning time step $k+1$.

$$f_r^q(x_{k,s}, \boldsymbol{\psi}) = \begin{cases} 1 & \text{if whale } q \text{ is already visited } (r_s^q = 1), \\ 1 & \text{if whale } q \text{ at time } k+1 \text{ is at the surface } (f_s^q(x_{k,s}, \boldsymbol{\psi}) = 1) \\ & \text{and the distance between a robot location } d_s^\ell \text{ is less than } \rho_{\text{rendezvous}} \\ & \text{meter of whale } q\text{'s location } Aw_s^q \text{ at time } k+1\ (\exists \ell \in \{1, \ldots, m\} \text{ such that} \\ & \left\| d_s^\ell + \left[v^\ell \cos\left(\phi^\ell\right), v^\ell \sin\left(\phi^\ell\right)\right]^T - LAw_s^q \right\| \le \rho_{\text{rendezvous}}\Big), \\ 0 & \text{otherwise.} \end{cases}$$

The whale assignments to robots at a future planning time step κ for belief particle $x_{\kappa,s}$ used in Eq. 2 are obtained by $\boldsymbol{\Psi}_{\text{BIP}}(x_{\kappa,s}) = \left[\psi^1_{\text{BIP}}(x_{\kappa,s}), \ldots, \psi^m_{\text{BIP}}(x_{\kappa,s})\right]$, where robot ℓ is assigned the whale $\psi^\ell_{\text{BIP}}(x_{\kappa,s})$.

$$\begin{aligned} &\psi^\ell_{\text{BIP}}(x_{\kappa,s}) \leftarrow q, \text{ if } \widetilde{y}_{\ell,q} = 1, \text{ where } \widetilde{y} \text{ solves the optimization} \\ &\widetilde{y} \in \operatorname{argmin}_{y \in \{0,1\}^{m \times n}} \sum_{\substack{\ell \in \{1, \ldots, m\}, \\ q \in \{1, \ldots, n\}, r_s^q = 0}} y_{\ell,q} \cdot \tau(x_{\kappa,s}, \ell, q) \\ &\text{Subject to constraints in expression (5)} \end{aligned} \tag{4}$$

$$\begin{cases} \text{Each robot } \ell \text{ is assigned to exactly one unvisited whale:} \\ \quad \sum_{q \in \{1,\ldots,n\}, r_s^q = 0} y_{\ell,q} = 1, \text{ and} \\ \text{Each unvisited whale } q \text{ with } r_s^q = 0 \text{ is assigned to at most one robot:} \\ \quad \sum_{\ell \in \{1,\ldots,m\}} y_{\ell,q} \le 1, \end{cases} \quad \left\{ \begin{array}{l} \text{if there are} \\ \text{fewer robots} \\ \text{than unvisited} \\ \text{whales, or } m \le \\ \sum_q \left(1 - r_s^q\right), \end{array} \right.$$

$$\begin{cases} \text{Each unvisited whale } q \text{ with } r_s^q = 0 \text{ is assigned to exactly one robot:} \\ \quad \sum_{\ell \in \{1,\ldots,m\}} y_{\ell,q} = 1, \text{ and} \\ \text{Each robot } \ell \text{ is assigned to at most one unvisited whale:} \\ \quad \sum_{q \in \{1,\ldots,n\}, r_s^q = 0} y_{\ell,q} \le 1, \end{cases} \quad \left\{ \begin{array}{l} \\ \\ \text{otherwise.} \\ \\ \\ \end{array} \right. \tag{5}$$

Here, $\tau(x_{\kappa,s}, \ell, q)$ is the time taken by robot ℓ to rendezvous with whale q, starting from a belief particle $x_{\kappa,s}$ at time step κ in the rollout planning horizon. Rendezvous is only possible with whale q when it is at the surface, which is determined by I_s^q.

The whale assignments to robots at the current belief x_k used in Eq. 1 are obtained by $\boldsymbol{\Psi}_{\text{BIP}}(x_k) = \left[\psi^1_{\text{BIP}}(x_k), \ldots, \psi^m_{\text{BIP}}(x_k)\right]$, where robot ℓ is assigned the whale $\psi^\ell_{\text{BIP}}(x_k)$.

$$\begin{aligned} &\psi^\ell_{\text{BIP}}(x_k) \leftarrow q, \text{ if } \widetilde{y}_{\ell,q} = 1, \text{ where } \widetilde{y} \text{ solves the optimization} \\ &\widetilde{y} \in \arg\min_{y \in \{0,1\}^{m \times n}} \sum_{\substack{\ell \in \{1, \ldots, m\}, \\ q \in \{1, \ldots, n\}, r_s^q = 0}} y_{\ell,q} \cdot \mathbb{E}_s\left[\tau(x_{k,s}, \ell, q)\right] \\ &\text{Subject to constraints in expression (5)} \end{aligned} \tag{6}$$

Here, $\mathbb{E}_s\left[\tau(x_{k,s}, \ell, q)\right]$ is the expected time taken by robot ℓ to rendezvous with whale q starting from the current belief x_k, considering the uncertainties involved in the whale locations and future surface intervals associated with all belief particles.

At each time step of policy execution, robots are assigned to whales $\left[\psi^1_{\text{rollout}}, \ldots, \psi^m_{\text{rollout}}\right]$ using m rollout optimizations using Eq. 1. Robot ℓ moves toward whale $\psi^\ell_{\text{rollout}}$, with a heading angle $\phi^\ell_{\text{rollout}}$ and velocity v^ℓ_{rollout}

$$\begin{aligned} \phi^\ell_{\text{rollout}} &= \angle\left(d^\ell, L\, w^{\psi^\ell_{\text{rollout}}}\right), \\ v^\ell_{\text{rollout}} &= \min\left\{v_{\max}, \left\| L\, w^{\psi^\ell_{\text{rollout}}} - d^\ell \right\|\right\} \end{aligned} \tag{7}$$

where d^ℓ is the current robot location and $L\, w^{\psi^\ell_{\text{rollout}}}$ is the current location of whale $\psi^\ell_{\text{rollout}}$, assigned to robot ℓ.

In principle, each robot can execute its rollout optimization in a decentralized fashion by sharing its beliefs and controls. We consider an implementation where a central server/computer collects all AOA observations of whales, assumes access to all agent beliefs and controls, and dictates the order of rollout optimization. Previous work (*33*) has investigated multiagent rollout with communication constraints and randomized agent order; we leave its integration as a part of future work. Details on the localization algorithm, whale surface interval sampling, data preprocessing, and system implementation are provided in the Supplementary Materials.

Autonomy module comparison

We compared the performance of our autonomy module with two online routing methods in the comparison study with the DSWP dataset presented in Fig. 7 in the "AVATARS autonomy module" section. The VRP-TW algorithm found routes for a team of robots to rendezvous with whales with predefined surfacing windows for the whales. For each whale, we used a time window interval $(0, \mu_{\text{surface}})$ if the whale was at the surface, and $(\mu_{\text{underwater}}, \mu_{\text{underwater}} + \mu_{\text{surface}})$ otherwise. VRP-TW minimized the total time to visit all whales, each only once, with time window constraints. For this comparison study, we used the VRP-TW implementation (*56*) where robots were not reassigned other than the initial route.

The IA algorithm involved a deterministic greedy assignment of unvisited whales to robots by minimizing the time taken by a robot to reach the whale at its upcoming surfacing. In contrast with our rollout-based approach, both VRP-TW and IA methods do not use stochastic optimization and cannot consider the stochastic nature of whales' behaviors in their framework without changing their approach. This resulted in the myopic actions of the robots.

AVATARS sensing module

Mathematical formulation

The SAR formulation used for emulating a virtual planar circular antenna array uses relative signal phase between the two fixed omnidirectional antennas to cancel channel phase noise. Thus, it allows for compatibility with passive signal transmitters such as off-the-shelf VHF fish tags. This capability also enables obtaining VHF AOA for UAVs that have small form factor and/or constrained dynamics, such as fixed-wing UAVs. We note that achieving sufficient phase change between the two antennas requires that the separation between them (denoted by r_2) to be $\geq\lambda/2$, where λ is the signal wavelength (*46*). This relative channel $\hat{h}_i$ between the physical antennas on the UAV at its ith position is obtained as $\hat{h}_i = \hat{h}_i^2 * \mathrm{conj}\left(\hat{h}_i^1\right)$, where $\hat{h}_i^1, \hat{h}_i^2$ denotes the signal collected at the receiver antennas and conj(.) denotes complex conjugate (*43*). For a signal arriving from direction (ϕ_T, θ_T), it is given by

$$\hat{h}_i = \frac{1}{d^2}\exp\left(\frac{-j2\pi r_2}{\lambda}\cos\left(\phi_T - \varphi_i\right)\sin\theta_T\right) \quad (8)$$

Thus, the relative channel is obtained as the product of the complex signal received at the first receiving antenna with the complex conjugate of the signal received by the second antenna. This removes signal noise, referred to as carrier frequency offset, in passive VHF signals while also maintaining translation resilience to mitigate the effect of trajectory distortion. The signal phase delays across the positions of the UAV (virtual antenna elements) are represented by the steering vector $\mathbf{a}(\phi, \theta)$ for the planar circular array (*62*). $a_i(\phi, \theta)$ is an element of the steering vector corresponding to this ith position of the UAV and is given by

$$a_i(\phi, \theta) = \exp\left(\frac{j2\pi r_2}{\lambda}\cos\left(\phi - \varphi_i\right)\sin\theta\right) \quad (9)$$

Last, we used the Bartlett equation for computing an AOA profile $f(\phi, \theta)$ for all pairs of (ϕ, θ) as follows (*63*, *64*)

$$f(\phi, \theta) = \left|\frac{1}{n}\sum_{i=1}^{n}\hat{h}_i\, a_i(\phi, \theta)\right|^2 \quad (10)$$

Our system returned the estimated AOA as arg max ($f(\phi, \theta)$). However, prior work has shown that emulating a circular antenna array can lead to inaccurate estimation of elevation angle θ (*42*). Hence, for practical purposes, only the azimuth angle ϕ is used as an input to the autonomy module.

System implementation

Real-time estimation of VHF AOA occurred in two stages. First, the UAV's pose and the received signal data were collected on board by Raspberry Pi (RPi) using the Pixhawk and the software-defined radio (SDR), respectively. We used MAVROS (*65*) for collecting data from sensors connected to Pixhawk and the SoapySDR framework (*66*) to collect data from the SDR. The SDR data were collected for the VHF signal frequency of a tag at a sampling rate of 71.428 kHz and bandwidth of 200 kHz at the two receiver channels of the SDR, each connected to an omnidirectional antenna. For the experiments with the engineered whale, we used a custom VHF beacon that uses a DRA818V VHF module. The off-the-shelf passive VHF tag used was F1840B from ATSTrack. Next, the RPi transmitted these collected data over Wi-Fi to a remote computer. The AOA algorithm, implemented in MATLAB filters for VHF signal pings using the pulse finding algorithm from (*21*), used Eq. 10 to compute the AOA profile $f(\phi, \theta)$ within 3 s on average. Additional details are provided in the Supplementary Materials.

UAV payload development

The payload consisted of a USRP B210 SDR, with the antennas connected to its two phase-coherent receivers. To collect UAV pose data, we used a Pixhawk 6C flight controller for easy integration of compass and GPS sensors. We note that the Pixhawk is optional if the data can be collected directly via the UAV's telemetry. The SDR and Pixhawk were connected serially over USB to the RPi. The Pixhawk and RPi were powered with two pairs of 3800 mAh LiPO cells; the SDR was powered by the serial connection. To minimize electromagnetic interference from the SDR, the payload was covered in a copper foil.

To enable operation in extreme environments, the SDR, RPi, and Pixhawk were housed inside a waterproof enclosure. However, complete sealing of the enclosure leads to heat buildup. Hence, the 3D-printed mounts for the electronics used polycarbonate material with reinforced carbon fiber because of its higher heat resistance. The RPi also had a separate active heat sink, and an aluminum sheet was placed inside for passive cooling. The GPS and compass sensors were mounted outside the payload. Waterproof glands with O-rings allowed wiring between the external sensors and the Pixhawk. Two additional glands allowed for connecting omnidirectional antennas to the SDR via cables with subminiature version A connectors. The contact surface between the enclosure and the glands was covered with epoxy to prevent water ingress. A temperature sensor monitored and displayed the payload's inner temperature and humidity.

Our waterproof UAV was a SwellPro SplashDrone 4 with custom modifications to carry the payload and the antennas. The payload was attached to the base of the UAV using 3D-printed supports. Carbon fiber rods extended out from the supports on both sides that enabled mounting of the antennas. The UAV's GPS enclosure was extended out and covered with copper and tin foils to eliminate interference from the SDR. Floaters, covered in Velcro tape, were added to the UAV's landing skid to enable buoyancy in water and landing on the platform installed on our catamaran (movies S3 and S4). The landing platform was a foldable table (1 m by 2.5 m) affixed with 5-cm-thick plywood covered with Velcro.

Statistical analysis

To evaluate our autonomy module, we used 50 initial states with randomized robot locations within the operation region and randomized start times of the whales within their first underwater phases. We report mean and SDs for the successful rendezvous percentages and mission times and normalized frequencies for missed whales for each configuration, including various rendezvous radii and sensing modalities. Errors in acoustic AOA measurements from the towed array were estimated using the engineered whales' AOA and the GPS ground-truth locations. Error for VHF AOA is reported with median and SD for 12 runs during the engineered whale experiments.

Supplementary Materials

The PDF file includes:

Supplementary Methods
Figs. S1 to S8
Tables S1 to S5
Legends for movies S1 to S4
References (*67*–*73*)

Other Supplementary Material for this manuscript includes the following:
Movies S1 to S4

REFERENCES AND NOTES

1. S. Gero, J. Gordon, H. Whitehead, Individualized social preferences and long-term social fidelity between social units of sperm whales. *Anim. Behav.* **102**, 15–23 (2015).
2. H. Whitehead, R. Antunes, S. Gero, S. N. Wong, D. Engelhaupt, L. Rendell, Multilevel societies of female Sperm Whales (*Physeter macrocephalus*) in the Atlantic and Pacific: Why are they so different? *Int. J. Primatol.* **33**, 1142–1164 (2012).
3. P. Sharma, S. Gero, R. Payne, D. F. Gruber, D. Rus, A. Torralba, J. Andreas, Contextual and combinatorial structure in sperm whale vocalisations. *Nat. Commun.* **15**, 3617 (2024).
4. S. Gero, H. Whitehead, L. Rendell, Individual, unit and vocal clan level identity cues in sperm whale codas. *R. Soc. Open Sci.* **3**, 150372 (2016).
5. L. Marino, "Cetacean cognition" in *The Oxford Handbook of Animal Studies*, L. Kalof, Ed. (Oxford Univ. Press, 2017), pp. 227–239.
6. L. E. Rendell, H. Whitehead, Vocal clans in sperm whales (*Physeter macrocephalus*). *Proc. R. Soc. London Ser. B Biol. Sci.* **270**, 225–231 (2003).
7. T. A. Hersh, S. Gero, L. Rendell, M. Cantor, L. Weilgart, M. Amano, S. M. Dawson, E. Slooten, C. M. Johnson, I. Kerr, R. Payne, A. Rogan, R. Antunes, O. Andrews, E. L. Ferguson, C. A. Hom-Weaver, T. F. Norris, Y. M. Barkley, K. P. Merkens, E. M. Oleson, T. Doniol-Valcroze, J. F. Pilkington, J. Gordon, M. Fernandes, M. Guerra, L. Hickmott, H. Whitehead, Evidence from sperm whale clans of symbolic marking in non-human cultures. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2201692119 (2022).
8. M. Cantor, S. Gero, H. Whitehead, L. Rendell, "Sperm whale: The largest toothed creature on Earth" in *Ethology and Behavioral Ecology of Odontocetes*, C. W. Clark, E. C. Garland, Eds. (Springer, 2019), pp. 261–280.
9. M. P. Johnson, P. L. Tyack, A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE J. Oceanic Eng.* **28**, 3–12 (2003).
10. S. L. Watwood, P. J. O. Miller, M. Johnson, P. T. Madsen, P. L. Tyack, Deep-diving foraging behaviour of sperm whales (*Physeter macrocephalus*). *J. Anim. Ecol.* **75**, 814–825 (2006).
11. P. Tonnesen, S. Gero, M. Ladegaard, M. Hohnson, P. Madsen, First-year sperm whale calves echolocate and perform long, deep dives. *Behav. Ecol. Sociobiol.* **72**, 165 (2018).
12. P. Tønnesen, C. Oliveira, M. Johnson, P. T. Madsen, The long-range echo scene of the sperm whale biosonar. *Biol. Lett.* **16**, 20200134 (2020).
13. C. Oliveira, M. Wahlberg, M. A. Silva, M. Johnson, R. Antunes, D. M. Wisniewska, A. Fais, J. Gonçalves, P. T. Madsen, Sperm whale codas may encode individuality as well as clan identity. *J. Acoust. Soc. Am.* **139**, 2860–2869 (2016).
14. T. Clutton-Brock, B. C. Sheldon, Individuals and populations: The role of long-term, individual-based studies of animals in ecology and evolutionary biology. *Trends Ecol. Evol.* **25**, 562–573 (2010).
15. C. Rutz, M. Bronstein, A. Raskin, S. C. Vernes, K. Zacarian, D. E. Blasi, Using machine learning to decode animal communication. *Science* **381**, 152–155 (2023).
16. J. Andreas, G. Beguš, M. M. Bronstein, R. Diamant, D. Delaney, S. Gero, S. Goldwasser, D. F. Gruber, S. de Haas, P. Malkin, N. Pavlov, R. Payne, G. Petri, D. Rus, P. Sharma, D. Tchernov, P. Tønnesen, A. Torralba, D. Vogt, R. J. Wood, Toward understanding the communication in sperm whales. *iScience* **25**, 104393 (2022).
17. R. Payne, "I spent my life saving the whales. Now they might save us," *Time*, 5 June 2023; https://time.com/6284884/whale-scientist-last-please-save-the-species/.
18. J. Andreas, G. Beguš, M. M. Bronstein, R. Diamant, D. Delaney, S. Gero, S. Goldwasser, D. F. Gruber, S. de Haas, P. Malkin, R. Payne, G. Petri, D. Rus, P. Sharma, D. Tchernov, P. Tønnesen, A. Torralba, D. Vogt, R. J. Wood, Cetacean translation initiative: a roadmap to deciphering the communication of sperm whales. arXiv:2104.08614 [cs.SD] (2021).
19. The Dominica Sperm Whale Project; www.thespermwhaleproject.org.
20. S. Gero, M. Milligan, C. Rinaldi, P. Francis, J. Gordon, C. Carlson, A. Steffen, P. Tyack, P. Evans, H. Whitehead, Behavior and social structure of the sperm whales of Dominica, West Indies. *Mar. Mammal Sci.* **30**, 905–922 (2014).
21. A. Torabi, M. W. Shafer, G. S. Vega, K. M. Rothfus, "UAV-RT: An SDR based aerial platform for wildlife tracking" in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)* (IEEE, 2018), pp. 1–6.
22. H. V. Nguyen, M. Chesser, L. P. Koh, S. H. Rezatofighi, D. C. Ranasinghe, Trackerbots: Autonomous unmanned aerial vehicle for real-time localization and tracking of multiple radio-tagged animals. *J. Field Robot.* **36**, 617–635 (2017).
23. H. Bayram, N. Stefas, K. S. Engin, V. Isler, "Tracking wildlife with multiple UAVs: System design, safety and field experiments" in *2017 International Symposium on Multi-Robot and Multi-Agent Systems (MRS)* (IEEE, 2017), pp. 97–103.
24. H. V. Nguyen, F. Chen, J. A. Chesser, H. Rezatofighi, D. C. Ranasinghe, "Lavapilot: Lightweight UAV trajectory planner with situational awareness for embedded autonomy to track and locate radio-tags" in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2020), pp. 2488–2495.
25. B. Thomas, J. D. Holland, E. O. Minot, Wildlife tracking technology options and cost considerations. *Wildlife Res.* **38**, 653–663 (2011).
26. O. M. Cliff, R. C. Fitch, S. Sukkarieh, D. Saunders, R. Heinsohn, "Online localization of radio-tagged wildlife with an autonomous aerial robot system" in *Proceedings of Robotics: Science and Systems* (RSS Foundation, 2015), 10.15607/RSS.2015.XI.042.
27. M. W. Shafer, G. S. Vega, K. M. Rothfus, P. G. Flikkema, UAV wildlife radiotelemetry: System and methods of localization. *Methods Ecol. Evol.* **10**, 1783–1795 (2019).
28. F. Körner, R. Speck, A. Göktogan, S. Sukkarieh, "Autonomous airborne wildlife tracking using radio signal strength" in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2010), pp. 107–112.
29. M. W. Shafer, P. G. Flikkema, Tracking small wildlife with minimal-complexity radio frequency transmitters: Near-optimal detection. *IEEE Access* **11**, 40029–40037 (2023).
30. S. Goldwasser, D. F. Gruber, A. T. Kalai, O. Paradise, "A theory of unsupervised translation motivated by understanding animal communication" in *Advances in Neural Information Processing Systems 36 (NeurIPS 2023)*, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, S. Levine, Eds. (Curran Associates, 2023), pp. 37286–37320.
31. S. L. King, F. H. Jensen, Rise of the machines: Integrating technology with playback experiments to study cetacean social cognition in the wild. *Methods Ecol. Evol.* **14**, 1873–1886 (2023).
32. D. Bertsekas, Multiagent reinforcement learning: Rollout and policy iteration. *IEEE/CAA J. Autom. Sin.* **8**, 249–272 (2021).
33. S. Bhattacharya, S. Kailas, S. Badyal, S. Gil, D. Bertsekas, Multiagent reinforcement learning: Rollout and policy iteration for pomdp with application to multirobot problems. *IEEE Trans. Robot.* **40**, 2003–2023 (2024).
34. D. Garces, S. Bhattacharya, S. Gil, D. P. Bertsekas, "Multiagent reinforcement learning for autonomous routing and pickup problem with adaptation to variable demand" in *2023 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2023), pp. 3524–3531.
35. D. Garces, S. Bhattacharya, D. Bertsekas, S. Gil, "Approximate multiagent reinforcement learning for on-demand urban mobility problem on a large map," in *2024 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2024), pp. 6843–6849.
36. S. Bhattacharya, S. Kailas, S. Badyal, S. Gil, D. P. Bertsekas, "Multiagent rollout and policy iteration for pomdp with application to multi-robot repair problems" in *Proceedings of the 2020 Conference on Robot Learning*, J. Kober, F. Ramos, C. Tomlin, Eds. (MLResearchPress, 2021), pp. 1814–1828.
37. D. Bertsekas, *Lessons from AlphaZero for Optimal, Model Predictive, and Adaptive Control* (Athena Scientific, 2022).
38. S. Bhattacharya, S. Badyal, T. Wheeler, S. Gil, D. P. Bertsekas, Reinforcement learning for POMDP: Partitioned rollout and policy iteration with application to autonomous sequential repair problems. *IEEE Robot. Autom. Lett.* **5**, 3967–3974 (2020).
39. D. Silver, J. Veness, "Monte-Carlo planning in large POMDPs," *Advances in Neural Information Processing Systems 23 (NeurIPS 2010)*, J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, A. Culotta, Eds. (Curran Associates, 2010).
40. D. Garces, S. Gil, "Surge routing: Event-informed multiagent reinforcement learning for autonomous rideshare" in *AAMAS '24:Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems* (ACM, 2024), pp. 641–650.
41. S. Gil, S. Kumar, D. Katabi, D. Rus, Adaptive communication in multi-robot systems using directionality of signal strength. *Int. J. Robot. Res.* **34**, 946–968 (2015).
42. N. Jadhav, W. Wang, D. Zhang, O. Khatib, S. Kumar, S. Gil, A wireless signal-based sensing framework for robotics. *Int. J. Robot. Res.* **41**, 955–992 (2022).
43. S. Kumar, S. Gil, D. Katabi, D. Rus, "Accurate indoor localization with zero start-up cost" in *MobiCom '14: Proceedings of the 20th Annual International Conference on Mobile Computing and Networking* (ACM, 2014), pp. 483–494.
44. S. Gil, S. Kumar, M. Mazumder, D. Katabi, D. Rus, Guaranteeing spoof-resilient multi-robot networks. *Autonom. Robot.* **41**, 1383–1400 (2015).
45. W. Wang, N. Jadhav, P. A. Vohs, N. Hughes, M. Mazumder, S. Gil, "Active rendezvous for multi-robot pose graph optimization using sensing over Wi-Fi" in *Robotics Research: The 19th International Symposium ISR*, T. Asfour, E. Yoshida, J. Park, H. Christensen, O. Khatib, Eds., vol. 20 of *Springer Proceedings in Advanced Robotics* (Springer, 2022), pp. 832–849.
46. D. E. Davies, "Circular arrays" in *The Handbook of Antenna Design Volume 2*, A. W. Rudge, K. Milne, A. D. Olver, P. Knight, Eds., vol. 28 of *Electromagnetic Waves Series* (London Peregrinus on behalf of the Institution of Electrical Engineers, 1983), chap. 12.
47. D. Tse, P. Viswanath, *Fundamentals of Wireless Communication* (Cambridge Univ. Press, 2012).
48. J. D. J. Macaulay, D. Gillespie, Pamguard: Open-source detection, classification, and localization software. *J. Acoust. Soc. Am.* **151**, A27–A28 (2022).
49. D. Gillespie, J. Gordon, M. Caillat, D. Claridge, D. Moretti, I. Boyd, Detection of beaked whales using near surface towed hydrophones: Prospects for survey and mitigation. *J. Acoust. Soc. Am.* **123**, 3774 (2008).
50. D. N. Wiley, C. J. Zadra, A. S. Friedlaender, S. E. Parks, A. Pensarosa, A. Rogan, K. A. Shorter, J. Urbán, I. Kerr, Deployment of biologging tags on free swimming large whales using uncrewed aerial systems. *R. Soc. Open Sci.* **10**, 221376 (2023).
51. D. Blount, S. Gero, J. V. Oast, J. Parham, C. Kingen, B. Scheiner, T. Stere, M. Fisher, G. Minton, C. Khan, V. Dulau, J. Thompson, O. Moskvyak, T. Berger-Wolf, C. V. Stewart, J. Holmberg,

J. J. Levenson, Flukebook: An open-source AI platform for cetacean photo identification. *Mammal. Biol.* **102**, 1005–1023 (2022).
52. G. Gubnitky, R. Diamant, Detecting the presence of sperm whales echolocation clicks in noisy environments. *IEEE/ACM Trans. Audio Speech Langu. Process.* **32**, 2050–2061 (2024).
53. E. Dubrovinskaya, R. Diamant, P. Casari, "Anchorless underwater acoustic localization" in *2017 14th Workshop on Positioning, Navigation and Communications (WPNC)* (IEEE, 2017), pp. 1–6.
54. B. B. Werger, M. J. Matarić, "Broadcast of local eligibility for multi-target observation" in *Distributed Autonomous Robotic Systems 4*, L. E. Parker, G. Bekey, J. Barhen, Eds. (Springer, 2000), pp. 347–356.
55. G. B. Dantzig, J. H. Ramser, The truck dispatching problem. *Manag. Sci.* **6**, 80–91 (1959).
56. V. Furnon, L. Perron, Or-Tools Routing Library; https://developers.google.com/optimization/routing/.
57. F. Dressler, S. P. Ripperger, M. Hierold, T. Nowak, C. Eibel, B. Cassens, F. Mayer, K. Meyer-Wegener, A. Koelpin, From radio telemetry to ultra-low-power sensor networks: tracking bats in the wild. *IEEE Commun. Mag.* **54**, 129–135 (2016).
58. H. Bayram, N. Stefas, V. Isler, "Aerial radio-based telemetry for tracking wildlife" in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2018), pp. 4723–4728.
59. A. Maalouf, N. Jadhav, K. M. Jatavallabhula, M. Chahine, D. M. Vogt, R. J. Wood, A. Torralba, D. Rus, Follow anything: Open-set detection, tracking, and following in real-time. *IEEE Robot. Autom. Lett.* **9**, 3283–3290 (2023).
60. K. Melo, T. Horvat, A. J. Ijspeert, Animal robots in the African wilderness: Lessons learned and outlook for field robotics. *Sci. Robot.* **8**, eadd8662 (2023).
61. E. A. Wan, R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation" in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)* (IEEE, 2000), pp. 153–158.
62. P. Ioannides, C. Balanis, Uniform circular arrays for smart antennas. *IEEE Antennas Propag. Mag.* **47**, 192–206 (2005).
63. H. Krim, M. Viberg, Two decades of array signal processing research: The parametric approach. *IEEE Signal Process. Mag.* **13**, 67–94 (1996).
64. P. Stoica, R. L. Moses, *Spectral Analysis of Signals* (Prentice Hall, 2005).
65. Mavros, https://github.com/mavlink/mavros.
66. SoapySDR Project, https://github.com/pothosware/SoapySDR/wiki.
67. N. Jaquet, A simple photogrammetric technique to measure sperm whales at sea. *Mar. Mammal Sci.* **22**, 862–879 (2006).
68. M. Nishiwaki, S. Ohsumi, Y. Maeda, Change of form in the sperm whale accompanied with growth. *Sci. Rep. Whales Res. Inst. Tokyo* **17**, 1–14 (1963).
69. R. Clarke, O. Paliza, Sperm whales of the southeast Pacific part III: Morphometry. *Hvalrad. Skrift. Sci. Results Mar. Biol. Res.* **53**, 1–106 (1972).
70. J. Gordon, Evaluation of a method for determining the length of sperm whales (*Physeter catodon*) from their vocalizations. *J. Zool.* **224**, 301–314 (1991).
71. D. Gillespie, D. K. Mellinger, J. Gordon, D. McLaren, P. Redmond, R. McHugh, P. Trinder, X.-Y. Deng, A. Thode, PAMGUARD: Semiautomated, open source software for real-time acoustic detection and localization of cetaceans. *J. Acoust Soc. Am.* **125**, 2547 (2009).
72. B. Miller, S. Dawson, R. Vennell, Underwater behavior of sperm whales off Kaikoura, New Zealand, as revealed by a three-dimensional hydrophone array. *J. Acoust Soc. Am.* **134**, 2690–2700 (2013).
73. D. Zhang, B. Zhao, K. Zhu, H. Jiang, Dynamic analysis of towed cable with variable length during turning maneuvers. *Sci. Rep.* **13**, 3525 (2023).

Acknowledgments: We thank the Chief Fisheries Officers and the Dominica Fisheries Division officers for research permits and collaboration in data collection; all of the crews of R/V Balaena and the DSWP team for data collection, curation, and annotation of photo-identification and Dtag datasets; as well as Dive Dominica, Al Dive, and W.E.T. Dominica for logistical support while in Dominica. We are grateful to M. Johnson and P. L. Tyack for in-kind contribution of DTAGs and associated code during DSWP field research. We thank field biologists Y. Mevorach and C. Baumgartner and the CETI marine operations team of K. George, D. Gibbons, O. Harve, and D. Gibbons for support during the Dominica expedition. We are grateful to M. Salino-Hugg for supporting custom tag development and A. Kenny, H. Zink, M. Awah, and A. Wright for supporting UAV modifications. We also thank the office of Harvard Athletics for providing access to outdoor drone testing space and D. Reeves, P. Choi Redfern, M. MacLeod, and M. Borton for help in coordinating. We thank D. Garces for thoughtful feedback on visualizations. **Funding:** This work was funded by Project CETI via grants from Dalio Philanthropies and Ocean X; Sea Grape Foundation; Rosamund Zander/Hansjorg Wyss, Chris Anderson/Jacqueline Novogratz through the Audacious Project: a collaborative funding initiative housed at TED. Further support was from National Geographic Society Grant (no. NGS72337T-20) and Lyda Hill Philanthropies. Fieldwork for DSWP, which produced the photo-identification (2005–2019) and Dtag (2014–2018) datasets used in this work, was supported by an FNU fellowship for the Danish Council for Independent Research supplemented by a Sapere Aude Research Talent Award, a Carlsberg Foundation expedition grant, a grant from Focused on Nature, two Explorer Grants from the National Geographic Society, and supplementary grants from the Arizona Center for Nature Conservation, Quarters For Conservation to S. Gero, with supplemental funding from the Dansk Akustisks Selskab, Oticon Foundation, and the Dansk Tennis Fond to P.T. Further funding was provided by Discovery and Equipment grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) to H. Whitehead of Dalhousie University and an FNU large frame grant and a Villum Foundation Grant to P. Madsen of the Marine Bioacoustics Lab at Aarhus University. We thank National Science Foundation for providing partial funding through the CAREER Award (CNS-2114733). We acknowledge support from the NSF (2106921, 2030154, 2007786, 1942902, and 2111751), ONR YIP N00014-21-1-2714, ONR, AFRETEC, MFI, CISCO, Safety21, and CyLab-Enterprise. **Author contributions:** N.J. and S.B.: Conceptualization, data curation, formal analysis, investigation, methodology, resources, software, validation, visualization, writing, and expedition. D.V. and Y.A.: Investigation, resources, and expedition. P.T.: Data curation. A.P. and S.K.: Formal analysis, methodology, and validation. S. Gero, R.J.W., and S. Gil: Conceptualization, investigation, project administration, resources, supervision, visualization, writing, funding acquisition, and expedition. **Competing interests:** The authors declare that they have no competing interests. **Data and material availability:** All data and code associated with this study can be found at https://doi.org/10.5061/dryad.05qfttfc9 and https://github.com/Project-CETI/avatars-code.

Submitted 24 December 2023
Accepted 7 October 2024
Published 30 October 2024
10.1126/scirobotics.adn7299