Privacy-Preserving Absence Confirmation in Sensitive Nuclear Facilities

Eric Lepowsky[1,2,†] David Snyder[1,3,†] Alexander Glaser[1,2] Anirudha Majumdar[1,3]

***Abstract*— Radioactive source term estimation is applicable to domains ranging from disaster response to nuclear monitoring. While robotic inspectors have been proposed, and in some cases deployed, for this task in certain nuclear applications, deployment in verification contexts – where privacy concerns are of the utmost importance – has not been addressed. Such applications impose strict constraints on sensors and stored information due to the potentially secret nature of observable features in the inspected site. One verification scheme, which avoids direct observation of sensitive items, is to confirm the absence of radioactive sources; however, this has not yet been adapted to the field of robotics. We propose a minimally-intrusive verification procedure that confirms the absence of sources without requiring, nor providing, any information about the search environment. The privacy and correctness of our random walk-based approach are validated by extensive simulated and experimental demonstrations.**

I. INTRODUCTION

Nuclear safeguards and arms control are cornerstones of the broader global security mission. Safeguards, which are measures to verify that nuclear facilities are not misused and nuclear material is not diverted from peaceful uses, require verification throughout the nuclear fuel cycle [1], [2]. Arms control, which is concerned with limiting arms competition and regulating arsenals, is made possible through declarations, data exchange, and inspections to verify compliance with agreed upon limits [3]–[5].

Onsite inspections play an important role in safeguards and arms control verification. Inspection tasks may include a variety of typically human-based measures, including verification of tags and seals, counting of objects, and, of principle interest here, radiation measurements. In scenarios where no radioactive sources (e.g., nuclear weapons or enriched uranium) are declared, the ability to identify anomalies could be crucial for verifying compliance with safeguards and arms control agreements. Furthermore, future agreements will likely require new verification approaches that minimize the need for human access to sites, such as storage or dismantlement facilities, and treaty accountable items [6].

The introduction of robotics to the field of nuclear verification has the potential to fundamentally improve and transform the efficiency, effectiveness, and capability of relevant inspection approaches [7]–[10]. A "robotic inspector" enables remote inspections so that human inspectors need not be physically present onsite, thereby reducing safety and security concerns while improving costs and timeliness.

Corresponding Author: `lepowsky@princeton.edu`. [1]Mechanical and Aerospace Engineering, Princeton University, NJ, USA. [2]Program on Science and Global Security (SGS), Princeton, NJ, USA. [3]Intelligent Robot Motion (IRoM) Lab, Princeton, NJ, USA. [†]E. L. and D. S. contributed equally; other authors listed alphabetically.

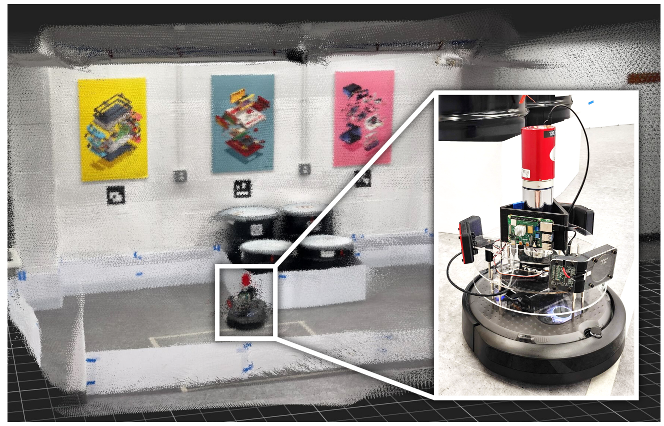


Fig. 1: Robotic inspector in a representative laboratory search environment. The test environment shown is approximately 15 m^2 with steel drum obstacles; dividers are used to reconfigure the space for varied test environments which are unknown to the inspector. The robotic inspector is comprised of a Create 3 platform fitted with a sodium iodide scintillator detector and Geiger counters. The inspector performs a heterogeneous random walk to explore the space while confirming the absence or presence of a radioactive source using only non-sensitive information.

Although robotic systems have been proposed, and in some cases deployed, for applications in nuclear disaster response and facility monitoring [11]–[13], there remain significant hurdles for deployment in safeguards and treaty verification. Perhaps most prominently, no sensitive information from the host site shall be unnecessarily or inadvertently revealed to the inspector. We assume that the environment layout and other observable features in the inspected site – which may include photos, dimensions, radiation measurements, etc. – are considered sensitive. Many robot-compatible methods are in contention with this requirement by either using *a priori* knowledge of the environment or, by consequence of the algorithmic design, revealing the site's configuration or radiation field *a posteriori* [14]–[21].

We also note that current approaches do not consider the inverse problem of confirming the *absence* of sources. Rather than localizing or identifying the presence of sources, which inherently requires acquiring and storing radiation measurements, confirming the absence of sources can be non-intrusive by design and avoids radiation measurements on potentially sensitive items [22]. In bringing absence confirmation to the robotics domain, we also take inspiration from the concept of "forgetting," often cited in human-robot interaction contexts [23], by implementing a more extreme alternative: never learning or remembering in the first place. With this motivation, we seek a high-confidence approach to confirm the absence of radioactive sources that neither reveals nor stores any sensitive information.

Statement of contributions. The primary contribution of this work is the development of a minimally-intrusive algorithm for confirming the absence of sources. Although the use case motivating this work and demonstrated here is specific to nuclear verification, the proposed approach can be adapted to secure inspections of other scalar fields, such as sound sources, gas leaks, aerosols, or any other emissions of interest where the signal drops with distance from the source. Guarantees on privacy preservation and bounds on the false positive rate are proven theoretically. Further, the false negative rate is characterized in terms of several parameters fundamental to the inspection task. Validation of the algorithm is undertaken in simulation and on extensive hardware experiments.

II. RELATED WORK

Radioactive source term estimation. Several methods, both for discrete source localization and distributed field mapping, have been reported which would be compatible with robotic radiation detection, including Bayesian methods [24]–[27], maximum likelihood estimation [28], [29], and machine learning [30], [31]. Mapping inherently generates a potentially sensitive result, since the inspecting party may not be privy to the configuration or layout of the search environment. A map representation of the environment can also be deduced from certain source localization methods. In the event that a source is present, the location and identifiable characteristics may be revealed. While leveraging scene understanding may improve accuracy or speed [32], [33], we do not see commensurate attention afforded to the opposite problem of completing the task with minimal information, which may be critical for verification applications.

Minimal-information decision making. Techniques like dimensionality reduction (e.g., via encoders) [34]–[36] and control-theoretic methods [37], [38] promote robustness via (respectively) lossy, compressed representation of the sensory feedback and regularization of information usage. Additionally, explicit analytical frameworks of "available information" [39], [40] quantify and formalize a relation between available and utilized information and performance. Our task falls on the extreme of this spectrum; concretely: can the source detection task be solved (in a predictable, theoretically-validated manner) with minimal (zero) information leakage? In this vein, the literature on differential privacy [41]–[43] is closely related in providing sufficient notions of information "security." However, it cannot account for the *performance* requirements demanded by our application.

Random walks. The algorithm that we propose guarantees exploration via random walk processes, for which the literature is exceedingly rich. Worst-case results for the expected cover time of walks on undirected graphs have been given in [44], [45]. General properties for finite graphs with a multitude of structures have been shown in [46]–[50], with classical Perron-Frobenius theory summarized in [51], [52]. More recent interest has been toward developing approximations to expected cover times, as in [53]–[55].

III. PROBLEM FORMULATION

Assume that a site is declared (by the host) to contain no radioactive sources. The inspection task is to verify the declaration, which requires confirming the absence of sources (or to confirm their presence in a non-compliance situation) by traversing the free space and measuring the observed radiation field. Complicating this is a crucial security constraint: the capacity of the robot inspector to gather and retain information (any observable characteristics or features of the environment) must be kept to a minimum. Ideal verification methods must allow for *both* calibrated correctness (being able to choose the probability that the robot will return the correct inspection result) and provable privacy (minimizing the robot's capacity to "leak" information).

A. Environment Definition

Define an environment $E(\mathcal{I}, s, M)$, in which a robotic inspector $\mathcal{I}$ is tasked with determining the presence or absence of a radioactive source of strength $s \geq 0$ in the environment map M. Key physical parameters are defined as follows: the robotic inspector $\mathcal{I}$ has a fundamental length $r_I > 0$ (e.g., its diameter); source strength is non-negative, with $s = 0$ corresponding to the case of "no source"; map $M(l_x, l_y, B)$ is physically bounded by positive length constants l_x, l_y, has Poisson-distributed (i.i.d.) background radiation of mean $B \geq 0$, and has an (unknown) occupancy function defining the free space.

B. Map Compression

For this work, we must restrict ourselves to the class of maps with a single, traversable (i.e., contiguous) region of free space. Additionally, we regularize the set of valid maps by discretizing the problem into a directed graph representation, where each bin (node) is a region of space. Specifically, we will need the inequalities of Eq. 1 to hold for the discretization length ϵ_M of map M:

$$r_I \leq \epsilon_M \leq \frac{r_D}{\sqrt{2}}. \tag{1}$$

The left-hand inequality ensures that traversability is preserved in the discretization process, while the latter ensures that if the robot enters a particular bin, it can detect a source anywhere else in that bin. r_D is the detector range, or the distance from which a source is readily detectable above background (we take this to be the distance at which the signal-to-background ratio reaches unity). Consequently, each time the discretized space is covered, there must be at least one potential anomalous measurement if a source is present. We note that r_I and r_D are non-sensitive and known prior to the inspection.[1]

Henceforth, we will refer to a *compressed map* as $M(l_x, l_y, B, \epsilon_M)$ and define a class of compressed maps as $\mathbb{M}(l_x, l_y, B, \epsilon_{\underline{M}}) = \{M(l_x, l_y, B, \epsilon_M) : \epsilon_M \geq \epsilon_{\underline{M}}\}$. The

[1] Further details are deferred to an extended version of this work. Where necessary, for all subsequent theoretical claims throughout the paper, the reader is directed to the extended version, available here: https://github.com/elepowsky/verification.

property $\epsilon' \geq \epsilon \implies \mathbb{M}(l_x, l_y, B, \epsilon') \subseteq \mathbb{M}(l_x, l_y, B, \epsilon)$ follows directly, yielding a clean subset relation. Incorporating Eq. 1, the valid set of maps for a given inspector is the set difference $\mathbb{M}_\mathcal{I} = \{\mathbb{M}(l_x, l_y, B, r_I) \setminus \mathbb{M}(l_x, l_y, B, \frac{r_D}{\sqrt{2}})\}$.

Importantly, although the inspector acts in $M \in \mathbb{M}_\mathcal{I}$, it does not see nor construct a representation of the underlying map; furthermore, the inspector does not collect or store information (e.g., a state history) which would be sufficient to deduce the map. We emphasize that all map-dependent results presented in this paper (including coverage times) are from an omniscient view *unavailable to the inspector*.

C. Source Detection with Limited Information

The verification task considered in this work has two distinct failure modes. To minimize the false negative rate (FNR) – i.e., returning "no source" when a source is present – the robot needs to guarantee exploration of the space. To control the false positive rate (FPR), each behavior indicative of source-presence must individually have a guaranteed FPR. This is similar to a standard suite of problems in robotics, ranging in application from out-of-distribution (OOD) and anomaly detection [60], [61] to failure prediction [62], among others. What distinguishes our setting, however, is the additional fundamental constraint that the stored information $\mathcal{G}_t$ be exclusively non-sensitive, that it not allow for reconstruction of the underlying map, and that this property hold uniformly across all $t \in \mathbb{N}$.

To formalize the constraint, we specify the information *available* to the robotic inspector. First, given the known r_I and r_D, it is assumed that the map M is indeed drawn from the class of valid maps $\mathbb{M}_\mathcal{I}$.[2] Second, the measurement model is Poisson-distributed: $h \sim \mathcal{P}(B + g(s, x, y))$. The non-negative function g is 0 if $s = 0$ or the robot position (x, y) does not have line-of-sight to the source; otherwise, $g \propto \frac{s}{r^2}$, where r is the Euclidean distance from the inspector to the source. Critically, the actual measurements h_t acquired are considered sensitive and cannot be directly stored; any trace of their history must be filtered in a sufficiently lossy fashion. Further, the robot's position cannot be known during operation, as this would reveal information about the map.

Therefore, the verification algorithm $\mathcal{A}$ must choose physical actions a_t and decisions d_t ("absence confirmed," "anomaly detected," or "continue") that rely only on non-sensitive accumulated information $\mathcal{G}_t$. We present Alg. 1 to accomplish this, and characterize both theoretical and empirical properties of its operation, spanning guarantees of minimal information leakage, performance (time to completion), and robustness (high-probability bounds on termination).

IV. METHODOLOGY

The algorithm we propose takes inspiration from randomized, sampling-based motion planners [63], [64] and out-of-distribution detection [65], with an emphasis on scalar task-relevant detection [66]. When measurements are consistent with source-absence, the robot moves according to a "reference" random walk that explores the space; otherwise, if consistent with source-presence, it moves according to an "out-of-distribution" random walk. Detection of a shift in the composite (realized) action distribution is accomplished by Kolmogorov-Smirnov (KS) testing [67]. Since the actions depend only on the detected counts, the resulting distribution over actions (step sizes) for any source-free map is theoretically identical (proven in Sec. V-B); we refer to this source-free action distribution as the reference, V_r.

In practice, we assume that a robot has the capacity to rotate and translate in a controllable manner (e.g., via encoders on a wheeled system), to detect imminent collisions in a non-destructive fashion (e.g., via very "myopic" distance sensors), and to accurately acquire radiation measurements; such a system can run Alg. 1. Intuitively, the robot "slows down" (takes smaller step sizes) when anomalously high counts (statistically above background) are detected. In this way, legitimate evidence of source presence is self-reinforcing, i.e., the robot gets "stuck." We set the reduced step size to $c_L = \frac{c_U}{10}$; for $c_L \lesssim \epsilon_{\underline{M}}$, the inspector is likely to stay in the vicinity of an anomalous source. Note that the step size recorded is the randomly-selected distance between measurement points, and *not* the distance between turns; when the robot encounters an obstacle, it randomly redirects and continues to travel the remaining distance.

Algorithm 1 Random walk absence confirmation.

Input: Background count rate B, outer dimensions L_x, L_y, confidence parameter p^*, run time T, test count n, threshold level z, constants $0 \leq c_L < c_U$, reference V_r
Initialize $\underline{p} = 1.0$, $x_0, y_0, \theta_0 \sim \mathcal{U}(L_x, L_y, 2\pi)$, $t = 1$, $V_e = \{\emptyset\}$
while $t \leq T$ **do**
  $N_t \sim h(x_t, y_t; E)$ {Measurement}
  $c \leftarrow c_L + (c_U - c_L)\mathbb{1}[N_t \leq B + z\sqrt{B}]$
  $ds, d\theta \sim \mathcal{U}(c, 2\pi)$ {Step Length, Rotation}
  Rotate by $d\theta$ rad. and move forward ds distance
  Append ds to memory V_e
  if $t \equiv 0 \pmod{T/n}$ **then**
    $\underline{p} = \min\{\underline{p}, \mathbf{KS}(V_e, V_r)\}$
  end if
  if $\underline{p} \leq p^*/n$ **then**
    return 1 {Anomaly detected}
  end if
end while
return 0 {Absence confirmed}

Note that Alg. 1 requires an estimate of the background radiation. In this work, we assume that the background has been previously characterized (e.g., when the environment was initialized); it could also be learned in a pseudo-online fashion using confidence intervals of the Poisson-distributed counts. We defer this problem to future work.

[2]This assumption is not too onerous given that human inspectors also have non-zero extent, and therefore would struggle to explore an overly obstacle-dense map. However, for now, we cannot rival humans in being able to flag certain "adversarial" maps as requiring alternative verification – though they of course would use (sensitive) sensory information to do so.

V. CORRECTNESS AND PRIVACY

Correctness requires guaranteeing a low FNR without compromising the FPR, and vice versa. We begin with an immediate characterization of the false-positive calibration.

Remark 1 (Calibrated False Positive Rate): The FPR of Alg. 1 (i.e., the probability of incorrectly detecting an out-of-distribution anomaly) is less than or equal to p^*. This follows from a union bound applied to the outcomes of n pre-specified KS tests at significance p^*/n.

To address the FNR, full coverage of the environment is necessary to correctly eliminate the possibility of source-presence with high confidence. Unfortunately, standard coverage algorithms typically rely on detailed knowledge of the environment [68]–[70]. Even planners which don't require the environment *a priori* typically maintain a state history, forming a representation of the space (e.g., a rudimentary occupancy map) that is incompatible with the information constraint. Therefore, while less time-efficient, random walk processes are needed to provide necessary exploration while not requiring any environmental information. However, using random walks implies that, to calibrate the FNR, we must be able to validate the walks' coverage time properties.

A. Coverage Properties

In essence, our absence confirmation algorithm is a random walk in continuous space reduced to a discrete graph, as described in Sec. III-B. We require a bound on coverage time (T) that guarantees, with high probability, that the entire accessible environment has been explored. Formally, we desire a function $\mathcal{T}(N) = \max_M T_M \; \forall M \in \mathbb{M}(l_x, l_y, B, \epsilon_{\underline{M}})$ providing an upper confidence bound on coverage time for a given class of maps. The following lemma characterizes the tail behavior of the distribution governing this quantity.

Lemma 2 (Passage Times in Exponential Family [47]): Consider any compressed map with N nodes. The distribution of first passage times to a node i from any other node $j \neq i$ is a member of the exponential family; namely, the first passage time from node $j^* \neq i$ to node i is distributed geometrically in non-dimensionalized time $\tau = \frac{t}{r_G(i)}$, where $r_G(i) \leq D_G \leq N$. This ensures that relatively tight high-probability bounds on coverage can be obtained; it also reflects how the particular structure of the physical random walk overcomes several worst-case coverage time results for directed graphs.

To achieve a calibrated FNR, the conditional probability of source detection *given* full graph coverage must also be quantified. This is a multi-faceted (but analytical) problem [71]. In general, utilizing higher efficiency detectors and sampling for longer periods improves the detectable range of source strengths. Henceforth, we assume that sources are sufficiently strong (or, that the detector is sufficiently sensitive within the range r_D) that the conditional probability is essentially equal to one. The results of the hardware experiments in Sec. VI-B are consistent with this assumption.

For now, we assess the coverage time empirically on a diverse set of 10 simulated environments (described further in Sec. VI-A), several of which reflect known worst-case configurations for undirected graphs. We simulate 50 independent trials for each 10×10 m environment and each of 5 different maximum step sizes c_U = (2, 4, 6, 8, 10 m). From this, we determine the coverage versus time for a range of discretization sizes (25, 100, or 400 bins of corresponding side length 2, 1, or 0.5 m). We can use these results (summarized in Table I) to approximate $\mathcal{T}(N)$. Figure 2 visualizes the empirical coverage over time for the 5×5 binning; the 25-bin compressed maps are most closely aligned with the hardware experiments of Sec. VI-B.

	2 m	4 m	6 m	8 m	10 m
5×5 bins	810 (3529)	305 (1861)	200 (1721)	159 (818)	145 (699)
10×10 bins	1481 (5285)	741 (2614)	611 (1932)	547 (1444)	547 (1369)
25×25 bins	3422 (12161)	2502 (6710)	2300 (6369)	2256 (6295)	2232 (5022)

TABLE I Empirical coverage time for varied maximum step size and discretization. The mean number of steps, averaged over all 10 environments and all 50 trials, is reported; the maximum over all rooms and trials is reported in parenthesis.

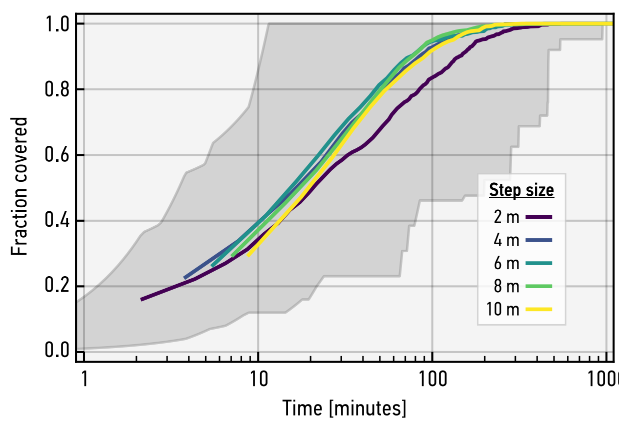


Fig. 2: Empirical coverage versus time for a range of maximum step sizes for the 5×5 binning. For each step size, the average over all 10 environments and all 50 trials is shown; the curves start after 10 initial time steps. The shaded region represents the full range of possible values, evaluated over all step sizes. To convert from step number to real-world time, we assumed 3-second measurements, travel speed of 10 cm/s, and neglect the time spent avoiding obstacles.

B. Information Privacy

In addition to the high-probability coverage and calibrated correctness of our proposed algorithm, a satisfactory verification approach must also be private. Theorem 3 formalizes the privacy of our proposed methodology.

Theorem 3 (Information Privacy of Compliant Hosts): Consider the class $\mathbb{M}^-(l_x, l_y, B, \epsilon_{\underline{M}})$ of compliant (source-free) maps. Then Alg. 1 is private, for all time, with respect to any map $M \in \mathbb{M}^-$, in that the mutual information [72] between any stored data point (namely, the step size between measurements) and the particular compliant (source-free) map is zero (Eq. 2).

$$\mathcal{MI}(\{\mathcal{G}_t \setminus \mathcal{G}_{t-1}\}, \mathbb{M}^-) = \mathcal{MI}(ds_t, \mathbb{M}^-) = 0 \; \forall t \geq 1 \quad (2)$$

In other words, at *no time* in its operation can Alg. 1 distinguish *any pair* of compliant maps.

To demonstrate this result, Figure 3 shows (in simulation) that our proposed algorithm, which stores only the step size between measurements, yields an inspection result that depends *only* on the absence or presence of a source, consistent with Thm. 3.[3] Conversely, a seemingly similar information storage scheme which stores the step size between turns (instead of between measurements), "leaks" enough information to differentiate between environments – the stored information is not environment-independent.

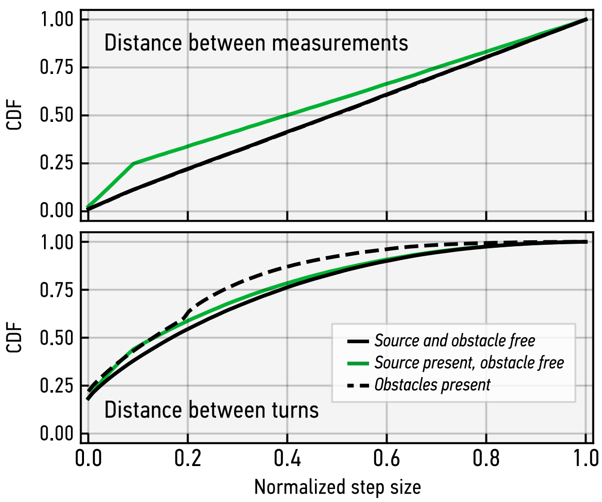


Fig. 3: Cumulative density functions over step size for our algorithm (distance between measurements) and a "leaky" alternative (distance between turns). Our algorithm (above) is only dependent on the presence/absence of a source, whereas the seemingly similar information storage scheme (below) leaks information which can differentiate between environments of differing occupancy. Note that the solid and dashed black lines in the upper plot are overlapped; this particular curve is equivalent to the reference distribution, V_r, which is independent of the environment.

VI. EXPERIMENTS

In this section, we demonstrate the correctness of our proposed algorithm, i.e., the ability to correctly identify the absence or presence of a source, both in diverse simulated environments and on hardware in various laboratory settings. Between the simulated and laboratory environments, we test our algorithm on a wide range of scales and configurations.

A. Simulation in PyBullet

Our simulation environment uses PyBullet [73], [74] and is based on the environment setup from [75], which provides an appropriate framework and robot model. A variety of environments were constructed (30 in total, not including the 10 rooms used for empirical results), each with different occupancy functions from several distinct "families." All environments were designed as a 10 m square (outer dimensions), but the open space inside ranges from the full 100 m^2 (empty room) down to 20 m^2; the assortment of environments includes a mix of predominantly open maps and obstacle-dense maps (see extended version for examples).

[3]Even in the source-presence case, the stored information is secure since the action distribution is essentially a lossy, non-unique convolution of unknown factors (source strength, distance, number of measurements, etc.).

For each map, 100 independent trials were conducted: 10 with and 10 without a source present, for 5 different maximum step sizes. For each trial, the robot and source (if present) were initialized in random positions within the open space. Ray-tracing provided a realistic, albeit simplified, radiation measurement model, where obstacles are assumed to be fully attenuating and the spatial (distance) dependence for non-attenuated counts was experimentally-based.

The evolution of the KS test significance for Alg. 1, averaged over all similar trials (i.e., 1,500 trials each for source absence and presence) across all environments, is shown in Figure 4. We apply the KS test after every 100 measurements, using $p^* = 0.005$ and $n = 500$, assuming a conservative upper coverage time of 50,000 steps. This yields an overall confidence of 99.5% (0.5% FPR). For omniscient reference, the corresponding average coverage is included; this information is *not* acquired, or even able to be inferred, given the data storage of Alg. 1.

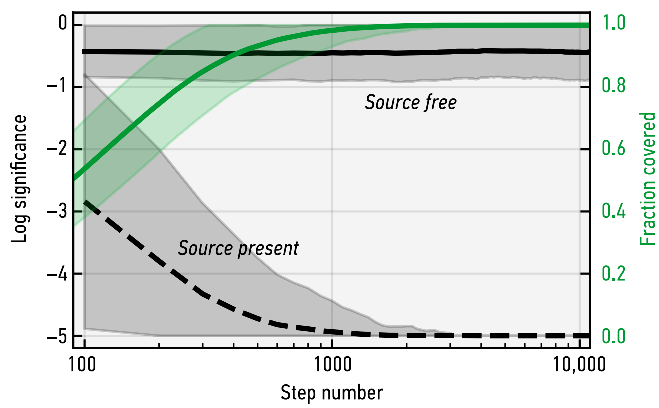


Fig. 4: Evolution of KS test significance and spatial coverage versus iteration for simulated trials. The results for source-absence and source-presence are averaged over several all trials and simulated rooms. Coverage is shown for the source-absence case; coverage for the source-presence case is omitted since Alg. 1 terminates once a source is confirmed (i.e., coverage is not guaranteed, nor is it necessary, if a source is detected). The log-significance, represented in $\log_{10}$-space, is floored at the KS test trigger threshold, which was -5 for the simulated trials. For each curve, the shaded region represents one standard deviation of the full range of values.

B. Hardware Demonstration

The prototype system developed for this work is built on the iRobot Create 3 platform [76]. Collision avoidance is accomplished using the onboard infrared sensors. The radiation detection unit includes two sensing modalities. For near-range measurements, with the potential for directional sensitivity,[4] three Geiger counters are evenly distributed around the circumference of the robot. We utilize the LND 7314 2in Pancake Geiger detector that is used by the Safecast bGeigie Nano [79]. An Adafruit ESP32 Feather V2 is used to count the pulses from each detector individually, and relay the counts to the controller via Bluetooth. We take the detector range of the Geiger counters to be $r_D = 30\ cm$.

[4]Directional radiation detectors [77], [78] enable a pseudo-random walk policy. In the absence of a prevailing source, the direction is random due to the Poisson counting statistics. If a source is present, the robot more expeditiously switches to the "slowed down" policy, manifesting as noisy gradient ascent, with similar functionality as model-free infotaxis [80].

For higher efficiency, long-range measurements, we use a 2-inch Mirion/Canberra NaI scintillator (Model 802) connected to an Osprey Digital MCA Tube Base [81], [82]. A Raspberry Pi 4 Model B reads the data over Ethernet, and similarly relays the counts to the controller. For rudimentary filtering to improve the signal-to-background ratio by excluding low energy noise, channels ≥ 400 of the 2048-channel spectrum are summed to yield gross counts. The NaI detector has a much larger detector range of conservatively $r_D = 1\ m$. The assembly of the unit (pictured in Figure 1) is designed such that the ring of Geiger counters does not obstruct the solid angle of the NaI crystal.

For source-presence environments, a set of gamma-ray check sources were used (Cs-137, Ba-133, and Co-60, among other isotopes) totaling to around 9 μCi of activity. The experimental analog to Figure 4 is shown in Figure 5 for two full-scale environments (20 m^2 rectangle and 22 m^2 L-shape) using the NaI detector and 1 m^2 bins. We apply the KS test after every 20 measurements, using $p^* = 0.005$ and $n = 50$, assuming an upper coverage time of 1000 steps. As before, this yields an overall confidence of 99.5%.

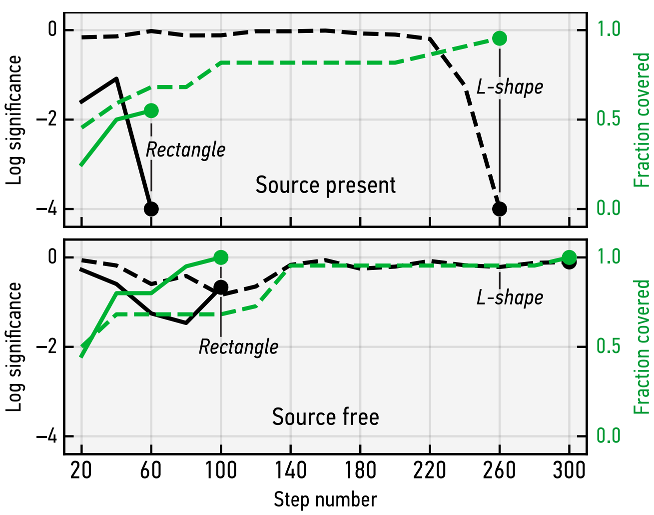


Fig. 5: Evolution of KS test significance and spatial coverage for two representative full-scale experimental environments. For the source-presence cases (above), the time step when Alg. 1 terminates is indicated (i.e., once the KS test log-significance, represented in $\log_{10}$-space, reaches -4). For the source-absence cases (below), the time step when coverage is first reached is indicated; note that the algorithm would *not* terminate in this case, since coverage is not known to the inspector.

In addition to the large-scale environments, we conducted a series of experiments in rooms of varying complexity and connectivity. Four laboratory environments are summarized in Table II, all using the NaI detector and 1 m^2 bins, each with a different open area and structure. For all trials in all environments, Alg. 1 yields the correct result: when no source was present, the robot covered the environment without returning a significant KS test; when a source was present, the robot more expeditiously returned a significant KS test result, indicative of an anomalous source. As seen in Figure 5 and Table II, if a source was present, the KS test causes Alg. 1 to terminate in fewer steps than necessary to achieve full coverage of the search environment.

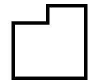
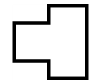
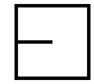
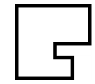
	 14 m^2	 12 m^2	 16 m^2	 13 m^2
Simulated cover time	166 $\pm$ 47	125 $\pm$ 59	265 $\pm$ 131	171 $\pm$ 63
Experiment cover time	76 $\pm$ 22	82 $\pm$ 36	94 $\pm$ 22	128 $\pm$ 42
KS time for source	39 $\pm$ 15 (52 $\pm$ 16)	50 $\pm$ 17 (56 $\pm$ 20)	60 $\pm$ 37 (72 $\pm$ 35)	34 $\pm$ 6 (40 $\pm$ 0)†

TABLE II Key metrics for laboratory trials. For the four environments, the time to reach full coverage for the no-source case and the time to the first instance of $\log_{10}(P) = -4$ for the source case is reported; each table entry includes the average and standard deviation over 5 trials. Note that for our chosen test parameters (1000 maximum steps, 50 KS tests), Alg. 1 can only return a significant KS test result after every 20 steps; the corresponding time to the algorithm terminating is in parenthesis. For reference, the experimental configurations were reproduced in PyBullet; the corresponding coverage time for 10 trials is reported. †For all 5 source trials in the 13 m^2 environment, the KS test triggered after 40 measurements.

To briefly demonstrate the utility of a detector with larger range r_D, we compare the performance of the two sensing modalities. We caution that the conservative obstacle avoidance of the Create 3 compounds with the limited range of the Geiger counters, thereby making it very difficult (and unlikely) that the robot remains in the vicinity of the source for long enough to trigger the KS test. Nonetheless, in a relatively small space measuring $2{\times}4$ m, the Geiger counters required 3.33-times more steps than the NaI-equipped robotic inspector for Alg. 1 to correctly terminate with $d = 1$ (anomaly detected). Similarly, coverage with the NaI detector was achieved 2.54-times faster than with the Geiger counters.

VII. DISCUSSION

This work underscores the prospect and fundamental challenge of deploying robotic radiation detectors for nuclear verification, where security is imperative and information leakage must be minimized. We established a privacy-preserving approach which demonstrates the ability to perform high-confidence source verification tasks without requiring nor providing any information about the search environment.

The constraints of finite time, complex or large environments, and detector efficiency may be addressed by limited, rather than minimal, information approaches. Quantifying this trade-off between permissible information and inspection efficiency remains an interesting research topic. As a future extension, the KS testing may also be used to confirm that an environment is unchanged by redefining the reference distribution (i.e, template matching). Alternatively, by increasing the situational awareness of the robot, limited contextual knowledge can afford the intuition that a radioactive source must be bound to a sufficiently substantial physical feature.

ACKNOWLEDGMENTS

This work has been supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-2039656, and the Consortium for Monitoring, Technology, and Verification under Department of Energy National Nuclear Security Administration award number DE-NA0003920.

REFERENCES

[1] *IAEA Safeguards: Serving Non-Proliferation*, International Atomic Energy Agency, 2018.

[2] J. Carlson, V. Kuchinov, and T. Shea, *The IAEA's Safeguards System as the Non-Proliferation Treaty's Verification Mechanism*, May 2020.

[3] J. Fuller, "Verification on the Road to Zero: Issues for Nuclear Warhead Dismantlement," Arms Control Today, December 2010.

[4] C. Comley, M. Comley, P. Eggins, G. George, S. Holloway, M. Ley, P. Thompson, and K. Warburton, *Confidence, Security & Verification, The Challenge of Global Nuclear Weapons Arms Control*, AWE/TR/2000/001, Atomic Weapons Establishment, Aldermaston, United Kingdom, 2000.

[5] *Radiation Detection Equipment: An Arms Control Verification Tool*, Product No. 211P, Defense Threat Reduction Agency, Fort Belvoir, VA, October 2011.

[6] National Academies of Sciences, Engineering, and Medicine, *Nuclear Proliferation and Arms Control Monitoring, Detection, and Verification: A National Security Priority: Interim Report*, The National Academies Press, 2021.

[7] F. F. Dean, *ROBIN: A Way to Collect In-Plant Safeguards Data with Minimal Inspector Access*, SAND82-1588C, Sandia National Laboratories, 1982.

[8] K. Robertson, R. Stohr, A. Elfes, P. Flick, A. Sokolov, D. Finker, and C. Everton, "The IAEA Robotics Challenge – Demonstrating Robots for Safeguards Inspections," IAEA Symposium on International Safeguards: Building Future Safeguards Capabilities, IAEA-CN-267/215, 2018.

[9] F. E. Schneider and D. Wildermuth, "Real-World Robotic Competitions for Radiological and Nuclear Inspection Tasks," *20th International Carpathian Control Conference (ICCC)*, pp. 1-6, 2019.

[10] B. Bird, A. Griffiths, H. Martin, E. Codres, J. Jones, A. Stancu, B. Lennox, S. Watson, and X. Poteau, "A Robot to Monitor Nuclear Facilities: Using Autonomous Radiation-Monitoring Assistance to Reduce Risk and Cost," *IEEE Robotics & Automation Magazine*, vol. 26, no. 1, pp. 35-43, 2019.

[11] R. Smith, E. Cucco, and C. Fairbairn, "Robotic Development for the Nuclear Environment: Challenges and Strategy," *Robotics*, vol. 9 no. 4, 94, 2020.

[12] M. Chiou, G. T. Epsimos, G. Nikolaou, P. Pappas, G. Petousakis, S. Mühl, and R. Stolkin, "Robot-Assisted Nuclear Disaster Response: Report and Insights from a Field Exercise," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022.

[13] I. Tsitsimpelis, C. J. Taylor, B. Lennox, and M. J. Joyce, "A Review of Ground-Based Robotic Systems for the Characterization of Nuclear Environments," *Progress in Nuclear Energy*, vol. 111, pp. 109-124, 2019.

[14] F. Gagliardi, "Integration of Independent NDA Techniques within a SLAM-based Robotic System for Improving Safeguards Standard Routines: A Review of the Current Status and Possible Future Developments," *ESARDA Bulletin*, vol. 64, no. 2, pp. 10-21, 2022.

[15] D. Hellfeld, T. H. Y. Joshi, M. S. Bandstra, R. J. Cooper, B. J. Quiterm, and K. Vetter, "Gamma-Ray Point-Source Localization and Sparse Image Reconstruction Using Poisson Likelihood," *IEEE Transactions on Nuclear Science*, vol. 66, no. 9, pp. 2088-2099, 2019.

[16] J. R. Vavrek, D. Hellfeld, M. S. Bandstra, V. Negut, K. Meehan, W. J. Vanderlip, J. W. Cates, R. Pavlovsky, B. J. Quiter, R. J. Cooper, and T. H. Y. Joshi, "Reconstructing the Position and Intensity of Multiple Gamma-Ray Point Sources with a Sparse Parametric Algorithm," *IEEE Transactions on Nuclear Science*, vol. 67, no. 11, pp. 2421-2430, 2020.

[17] F. Mascarich, P. D. Petris, H. Nguyen, N. Khedekar, and K. Alexis, "Autonomous Distributed 3D Radiation Field Estimation for Nuclear Environment Characterization," *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2163-2169, 2021.

[18] F. Mascarich, M. Kulkarni, P. De Petris, T. Wilson, and K. Alexis, "Autonomous mapping and spectroscopic analysis of distributed radiation fields using aerial robots," *Autonomous Robots*, vol. 47, pp. 139-160, 2023.

[19] A. West, I. Tsitsimpelis, M. Licata, A. Jazbec, L. Snoj, M. J. Joyce, and B. Lennox, "Use of Gaussian process regression for radiation mapping of a nuclear reactor with a mobile robot," *Scientific Reports*, vol. 11, 13975, 2021.

[20] N. A. Abd Rahman, K. S. M. Sahari, N. A. Hamid, and Y. C. Hou, "A coverage path planning approach for autonomous radiation mapping with a mobile robot," *International Journal of Advanced Robotic Systems*, vol. 19, no. 4, 2022.

[21] K. Groves, E. Hernandez, A. West, T. Wright, and B. Lennox, "Robotic Exploration of an Unknown Nuclear Environment Using Radiation Informed Autonomous Navigation," *Robotics*, vol. 10, no. 2, 78, 2021.

[22] E. Lepowsky, J. Jeon, and A. Glaser, "Confirming the Absence of Nuclear Warheads via Passive Gamma-Ray Measurements," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 990, 164983, 2021.

[23] R. Aylett and P. A. Vargas, "11. Social Interaction: Pets, Butlers, or Companions?" In *Living with robots: What every anxious human needs to know*, MIT Press, 2023.

[24] B. Ristic, M. Morelande, and A. Gunatilaka, "Information driven search for point sources of gamma radiation," *Signal Processing*, vol. 90, no. 4, pp. 1225-1239, 2010.

[25] E. A. Miller, S. M. Robinson, K. K. Anderson, J. D. McCall, A. M. Prinke, J. B. Webster, and C. E. Seifert, "Adaptively Reevaluated Bayesian Localization (ARBL): A novel technique for radiological source localization," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 784, pp. 332-338, 2015.

[26] R. B. Anderson and M. Pryor, "Mobile Robotic Radiation Surverying with Recursive Bayesian Estimation and Attenuation Modeling," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 1, 410-424, 2022.

[27] R. B. Anderson, C. Pehlivantürk, and M. Pryor, "Optimization Strategies for Bayesian Source Localization Algorithms," *IEEE Transactions on Automation Science and Engineering*, vol. 20, no. 1, pp. 394-403, 2023.

[28] G. Cordone, R. R. Brooks, S. Sen, N. S.V. Rao, C. Q. Wu, M. K. Berry, and K. M. Grieme, "Improved multi-resolution method for MLE-based localization of radiation sources," *20th International Conference on Information Fusion*, 2017.

[29] E.W. Bai, K. Yosief, S. Dasgupta, and R. Madumbai, "The maximum likelihood estimate for radiation source localization: Initializing an iterative search," *53rd IEEE Conference on Decision and Control*, pp. 277-282, 2014.

[30] P. Proctor, C. Teuscher, A. Hecht, and M. Osiński, "Proximal Policy Optimization for Radiation Source Search," *Journal of Nuclear Engineering*, vol. 2, no. 4, pp. 368-397, 2021.

[31] Z. Liu and S. Abbaszadeh, "Double Q-Learning for Radiation Source Detection," *Sensors*, vol. 19, no. 4, 960, 2019.

[32] M. S. Bandstra, D. Hellfeld, J. R. Vavrek, B. J. Quiter, K. Meehan, P. J. Barton, J. W. Cates, A. Moran, V. Negut, R. Pavlovsky, and T. H. Y. Joshi, "Improved Gamma-Ray Point Source Quantification in Three Dimensions by Modeling Attenuation in the Scene," *IEEE Transactions on Nuclear Science*, vol. 68, no. 11, pp. 2637-2646, 2021.

[33] G. Christie, A. Shoemaker, K. Kochersberger, P. Tokekar, L. McLean, and A. Leonessa, "Radiation search operations using scene understanding with autonomous UAV and UGV," *Journal of Field Robotics*, vol. 34, pp. 1450–1468, 2017.

[34] D. P. Kingma and M. Welling, "An Introduction to Variational Autoencoders," *Foundations and Trends in Machine Learning*, vol. 12, no. 4, pp. 307-392, 2019.

[35] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.A. Manzagol, "Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion," *Journal of Machine Learning Research*, vol. 11, no. 110, pp. 3371-3408, 2010.

[36] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," arXiv:1312.6114, 2022.

[37] V. Pacelli and A. Majumdar, "Robust Control Under Uncertainty via Bounded Rationality and Differential Privacy," *International Conference on Robotics and Automation (ICRA)*, pp. 3467-3474, 2022.

[38] M. Booker and A. Majumdar, "Learning to Actively Reduce Memory Requirements for Robot Control Tasks," *Proceedings of the 3rd Conference on Learning for Dynamics and Control*, PMLR, vol. 144, 2021.

[39] Y. Xu, S. Zhao, J. Song, R. Stewart, and S. Ermon, "A Theory of Usable Information under Computational Constraints," *International Conference on Learning Representations*, 2020.

[40] A. Majumdar, Z. Mei , and V. Pacelli, "Fundamental limits for sensor-based robot control," *The International Journal of Robotics Research*, 2023.

[41] C. Dwork and A. Roth, "The Algorithmic Foundations of Differential Privacy," *Foundations and Trends in Theoretical Computer Science*, vol. 9, no. 3-4, pp. 211-407, 2013.

[42] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating Noise to Sensitivity in Private Data Analysis," S. Halevi and T. Rabin, (Eds.), *Theory of Cryptography*, TCC 2006, pp. 265-284, "Lecture Notes in Computer Science," vol. 3876, Springer, Berlin, Heidelberg, 2006.

[43] F. McSherry and K. Talwar, "Mechanism Design via Differential Privacy," *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pp. 94-103, 2007.

[44] D. Aldous, "An Introduction to Covering Problems for Random Walks on Graphs," *Journal of Theoretical Probability*, vol. 2, pp. 87-99, 1989.

[45] R. Aleliunas, R. M. Karp, R. J. Lipton, L. Lovasz, and C. Rackoff, "Random Walks, Universal Traversal Sequences, and the Complexity of Maze Problems," *20th Annual Symposium on Foundations of Computer Science (SFCS)*, pp. 218-223, 1979.

[46] D. Aldous and P. Diaconis, "Strong uniform times and finite random walks," *Advances in Applied Mathematics*, vol. 8, no. 1, pp. 69-97, 1987.

[47] M. Mihail, "Conductance and convergence of Markov chains-a combinatorial treatment of expanders," *30th Annual Symposium on Foundations of Computer Science*, pp. 526-531, 1989.

[48] A. Sinclair and M. Jerrum, "Approximate counting, uniform generation and rapidly mixing markov chains extended abstract," H. Göttler and H.J. Schneider (Eds.), *Graph-Theoretic Concepts in Computer Science*, WG 1987, pp. 134-148, "Lecture Notes in Computer Science," vol. 314, Springer, Berlin, Heidelberg, 1988.

[49] F. Ball, B. Dunham, and A. Hirschowitz, "On the Mean and Variance of Cover Times for Random Walks on Graphs," *Journal of Mathematical Analysis and Applications*, vol. 207, no. 2, pp. 506-514, 1997.

[50] S. M. Ross, *Introduction to Probability Models*, 10th ed., Elsevier, Amsterdam, 2010.

[51] E. Seneta, *Non-negative Matrices and Markov Chains*, "Springer Series in Statistics," Springer, New York, 1981.

[52] D. Serre, *Matrices: Theory and Applications*, "Graduate Texts in Mathematics," vol. 216, Springer, New York, 2010.

[53] J.-Q. Dong, W.-H. Han, Y. Wang, X.-S. Chen, and L. Huang, "Universal cover-time distribution of heterogeneous random walks," *Physical Review E.*, vol. 107, no. 2, 024128, 2023.

[54] M. Chupeau, O. Bénichou, and R. Voituriez, "Cover times of random searches," *Nature Physics*, vol. 11, no. 10, pp. 844-847, 2015.

[55] L. Régnier, M. Dolgushev, S. Redner, and O. Bénichou, "Universal exploration dynamics of random walks," *Nature Communications*, vol. 14, no. 1, 618, 2023.

[56] G. M. Viswanathan, V. Afanasyev, S. V. Buldyrev, S. Havlin, M. G. E. da Luz, E. P. Raposo, and H. E. Stanley, "Lévy Flights in Random Searches," *Physica A: Statistical Mechanics and its Applications*, vol. 282, no. 1–2, pp. 1-12, 2000.

[57] A. M. Reynolds and C. J. Rhodes, "The Lévy flight paradigm: random search patterns and mechanisms," *Ecology*, vol. 90, pp. 877-887, 2009.

[58] J. Wei, Y.Q. Chen, Y. Yu, and Y. Chen, "Optimal Randomness in Swarm-Based Search," *Mathematics*, vol. 7, no. 9, 828, 2019.

[59] B. Pang, Y. Song, C. Zhang, and R. Yang, "Effect of random walk methods on searching efficiency in swarm robots for area exploration," *Applied Intelligence*, vol. 15, pp. 5189-5199, 2021.

[60] A. Sharma, N. Azizan, and M. Pavone, "Sketching curvature for efficient out-of-distribution detection for deep neural networks," *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, PMLR, pp. 1958–1967, 2021.

[61] R. Sinha, A. Sharma, S. Banerjee, T. Lew, R. Luo, S. M. Richards, Y. Sun, E. Schmerling, and M. Pavone, "A System-Level View on Out-of-Distribution Data in Robotics," arXiv:2212.14020, 2022.

[62] A. Farid, D. Snyder, A. Z. Ren, and A. Majumdar, "Failure Prediction with Statistical Guarantees for Vision-Based Robot Control," arXiv:2202.05894, 2022.

[63] S. M. LaValle and J. J. Kuffner, "Randomized Kinodynamic Planning," *International Journal of Robotics Research*, vol. 20, no. 5, pp. 378-400, 2001.

[64] J. J. Kuffner and S. M. LaValle, "RRT-Connect: An Efficient Approach to Single-Query Path Planning," *IEEE International Conference on Robotics and Automation (ICRA)*, vol. 2, pp. 995-1001, 2000.

[65] M. Basseville, "Detecting Changes in Signals and Systems – A Survey," *Automatica*, vol. 24, no. 3, pp. 309-326, 1988.

[66] A. Farid, S. Veer, and A. Majumdar, "Task-Driven Out-of-Distribution Detection with Statistical Guarantees for Robot Learning," *Proceedings of the 5th Conference on Robot Learning (PMLR)*, vol. 164, pp. 970-980, 2022.

[67] A. N. Kolmogorov, "Sulla Determinazione Empirica di Una Legge di Distribuzione," *Giornale dell'Istituto Italiano degli Attuari*, vol. 4, pp. 83-91. 1933

[68] Y. Cao, Y. Han, J. Chen, X. Liu, Z. Zhang, and K. Zhang, "Optimal Coverage Path Planning Algorithm of the Tractor-formation Based on Probabilistic Roadmaps," *IEEE International Conference on Unmanned Systems and Artificial Intelligence (ICUSAI)*, pp. 27-32, 2019.

[69] Z. Khanam, S. Saha, D. Ognibene, K. McDonald-Maier, and S. Ehsan, "An Offline-Online Strategy for Goal-Oriented Coverage Path Planning using A Priori Information," *14th IEEE International Conference on Industry Applications (INDUSCON)*, pp. 874-881, 2021.

[70] S. A. Sadat, J. Wawerla, and R. Vaughan, "Fractal trajectories for online non-uniform aerial coverage," *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2971-2976, 2015.

[71] G. F. Knoll, *Radiation Detection and Measurement*, 4th ed., Wiley, New York, 2010.

[72] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 1991.

[73] E. Coumans and Y. Bai, "PyBullet, a Python module for physics simulation for games, robotics and machine learning," `http://pybullet.org`, 2016–2021.

[74] B. Ellenberger, "PyBullet Gymperium," `https://github.com/benelot/pybullet-gym`, 2018–2019.

[75] Y. Kadhi, V. Lim, D. Zheng, and S. Doncieux, "Learning and generalization on a navigation task of a wheeled robot", `https://github.com/Yurof/WheeledRobotSimulations/`, 2021.

[76] iRobot, "Create 3 Docs," `https://iroboteducation.github.io/create3_docs/`, 2021–2023.

[77] F. Mascarich, C. Papachristos, T. Wilsonm, and K. Alexis, "Distributed Radiation Field Estimation and Informative Path Planning for Nuclear Environment Characterization," *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2318-2324, 2019.

[78] E. Lepowsky, M. Kütt, S. Aslam, H. Fetsch, S. Snell, A. Glaser, and R. J. Goldston, "Experimental Demonstration and Modeling of a Robotic Neutron Detector with Spectral and Directional Sensitivity for Treaty Verification," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 1041, 167362, 2022.

[79] A. Brown, P. Franken, S. Bonner, N. Dolezal, and J. Moross. "Safecast: successful citizen-science for radiation measurement and communication after Fukushima," *Journal of Radiological Protection*, vol. 36, no. 2, pp. S82-S101, 2016.

[80] M. Vergassola, E. Villermaux, and B. Shraiman, "'Infotaxis' as a strategy for searching without gradients," *Nature*, vol. 445, pp. 406-409, 2007.

[81] Mirion Technologies, "802 Scintillation Detectors Data Sheet," 2017.

[82] Mirion Technologies, "Osprey: Universal Digital MCA Tube Base for Scintillation Spectrometry Data Sheet," 2017.