2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Intermittent adaptive spatial field estimation and concurrent evacuation planning using field-dependent evacuee guidance

Michael A. Demetriou

Abstract— **This work combines a path planning scheme used for the evacuation of humans in indoor environments with the real-time estimation of spatially varying fields using adaptive methods. When a hazardous environment is known, then one possible trajectory planning scheme uses level-set methods to guide a human to safety (escape exit) while at the same time minimizes the accumulated amount of the hazardous field modelled as the hazardous substance inhaled. When the field representing the spatial distribution of the hazardous substance in unknown, then an arrested adaptive estimate of the spatial field is proposed in the level-set guidance. The human evacuee viewed as a mobile agent, obtains spatial field measurements and process them in an adaptive learning scheme to obtain an estimate of the spatial field. When a planning period is added to the traveling period, the adaptive scheme obtains the most recent spatial field estimate (arrested adaptation) and uses it as a time-invariant spatial field for trajectory planning.**

I. Introduction

The problem at hand is to guide a human starting from the interior of an indoor environment to the boundary of this environment. While this problem can leverage established results on trajectory planning with collision avoidance, the problem is exacerbated by the accumulated effects of a hazardous spatial field in the indoor environment. Such accumulated effects represent the total amount of a harmful substance, for example carbon monoxide, found in a human's lungs. While both the instantaneous levels of this harmful substance and the accumulated amount greatly affect a human's health, we concentrate on the accumulated amount as the singular feature for escape to safety.

Due to human inhaling of harmful substances during the inhale-exhale cycle, the amount inhaled at each time is added to the amount already present in the lungs. If the total amount exceeds a prescribed safety level, the escaping human may faint or, worse yet, may expire. If the field representing the spatial distribution of the harmful substance in the indoor environment is known, then an optimal solution to trajectory planning with guaranteed levels of the accumulated amount be below certain levels has been addressed in [1], [2]. This is an extension of the Zermelo navigation problem [3] with additional constraints. However, such an optimal escape policy is open-loop. As an alternate, an exhaustive search over all paths to safety examining also the accumulated amount associated with them may provide the set of acceptable paths, but that too may prove computationally prohibitive. A combination of earlier works on navigation functions via artificial potential fields may partially address this problem by using the spatial field as the potential field. A form of reactive navigation for avoiding regions of large instantaneous values of a spatial field has been presented in [4]. This of course would only minimize the instantaneous exposure to the harmful field and may not ensure that the accumulated effects fall below a prescribed threshold.

A somewhat suboptimal but feasible policy leverages trajectory planning using level-sets [5], [6], [7], [8], [9] to ensure that a human evacuee will reach safety and at the safety, the accumulated effects fall below the harmful threshold [10], [11]. The problem in this case is that the field is no longer known.

Estimating unknown spatial fields requires multiple static sensors. The number of measurements n required depends on the number of unknown parameters that are assumed in a series expansion of the unknown spatial field parametrization. Further, to ensure a certain identifiability condition is ensured, the spatial location of the sensing devices must not coincide with the zeros of the regressor (spatial) functions. Using a mobile sensor with prescribed path and leveraging observability properties with subspace identification-based eigenvalue estimation of a regressor matrix has been examined in detail in [12], [13]. An alternative to this is to employ adaptive estimation of the unknown parameters combined with a single mobile measurement [14]. The sensor motion, as demonstrated in [15], ensures parameter convergence, and hence functional convergence of the adaptive learning scheme. This relies on the fact that the outer product of the regressor vector evaluated at a single spatial point is rank-one matrix, but this rank-one matrix with mobile measurements can result in the sought-after *persistence of excitation* condition [16] when integrated over a time interval. Thus, the motion of *a single* measurement guarantees functional convergence for *any* spatial field of approximation order n!

Attempting to combine the level-set guidance over a known spatial field with guaranteed accumulated levels below a threshold with the adaptive learning using mobile measurements leads to a bottleneck since the level-set guidance requires known time-invariant fields. This paper provides a solution to this difficulty by employing *arrested adaptation* of the spatial field. This is implemented as follows. In a given cycle, one has a *planning stage* where the agent uses the most recent adaptive estimate of the spatial field, which is time-invariant, and implements the level-set guidance to define the path to take. During this time, the agent does not move but continues to obtain spatial field measurements. At the next stage, which is termed the *travelling stage*, the agent

M. A. Demetriou is with Worcester Polytechnic Institute, Aerospace Engineering Dept, Worcester, MA 01609, USA, mdemetri@wpi.edu. The author gratefully acknowledges financial support from the Office of Naval Research Grant N00174-22-1-0004.

implements the navigation based on the arrested adaptive estimate of the spatial field. If computing is negligible meaning that there is no need for he planning stage, then the cycle duration consists of only the traveling stage. Spatial information is based on the arrested adaptive estimate of the spatial field that is obtained at the beginning of the cycle. In such a case, one has a hybrid adaptation-plus-guidance whereby the mobile sensor is obtaining information throughout the duration of a given cycle, but uses the arrested estimate of the field at the beginning of the cycle.

The problem is formulated with the parametrization of the unknown spatial field given in Section II. The adaptive estimation of spatial fields with a single mobile agent is summarized in Section III. The inclusion of planning and travelling stages over a given time cycle for the human evacuation using arrested adaptive estimates of the unknown spatial field combined with the earlier development of a level-set path planning are given in Section IV. Demonstration of the adaptive learning scheme over a large domain representing a typical indoor environment is presented in Section V with conclusions following in Section VI.

II. Mathematical Formulation

The unknown spatial field, assumed for now time-invariant, is denoted by $x(\theta)$, where the spatial variable belongs to a bounded domain $\theta \in \Omega \subset \mathbb{R}^d$, $d = 1,2,3$. The one-dimensional case with $d = 1$ does not provide any interesting attributes for the field-dependent guidance over spatial fields. The two-dimensional spatial field with $d = 2$ represents the most interesting case as most of the cases of evacuation in enclosed domains assume a 2D spatial field representing hazardous fields such as carbon monoxide concentrations. This is because the 3D spatial field exhibits axisymmetry around the face of an evacuee; spatial gradients in the vertical direction are negligible around an evacuee's face with the length of a face being around 20cm compared to a typical room height of about 3m.

The spatial field measurements provided by a mobile sensor are modelled as a pointwise-in-space sensor distribution with a time-varying centroid $\theta_s(t)$.[1] This is assumed to be the inhaled amount of the hazardous substance by a human evacuee. Thus the measurement readout from the mobile sensor is precisely the value of the spatial function $x(\theta)$ at the sensor location and is given by

$$y(t;\theta_s(t)) = \int_\Omega \delta(\theta - \theta_s(t)) x(\theta)\, \mathrm{d}\theta = x(\theta_s(t)). \qquad (1)$$

As was noted in [15], the sensor readouts are time varying, despite the fact that the spatial field $x(\theta)$ is time-invariant. This is due to the time-varying sensor centroid $\theta_s(t)$.

Given an agent initial position $\theta_s(t_0)$ in the spatial domain Ω, the problem at hand is to

1) move to a desired location θ_d at the boundary of Ω representing safety in the smallest possible time, and
2) reach the desired location θ_d by minimizing the *accumulated measurements*.

The first goal can easily be achieved by harnessing established results on path-planning guidance over known environments, even by incorporating obstacle avoidance. If the field has no effects on the ability of an agent to move within the spatial domain, then a simple path from $\theta_s(t_0)$ to θ_d is easily obtained. Figure 1 depicts an indoor environment and for a constant speed, various paths result in different escape times that are proportional to the distance travelled.

The difficulty arises when exposure to the field has negative effects on the ability of the mobile agent to traverse within Ω. In particular, if the *instantaneous measurements* (1) have negative effects on the agent, and if the spatial field is known, then an obstacle-avoidance scheme can be used, where the large values of the field to be avoided are viewed as the obstacles to be avoided. The spatial field assumes the role of the artificial potential (navigation function) and the navigation schemes in [17], [18] address this problem. However, when the *accumulated measurements* have negative effects on the agent, then a new guidance must be used.

The accumulated amount of the hazardous material is given by the sum of all measured values (1) along a specific path to safety. It represents for example the amount of the hazardous material inhaled by an agent during an indoor evacuation. The accumulated amount up to time t is

$$z(t) = \int_0^t y(\tau;\theta_s(\tau))\, \mathrm{d}\tau. \qquad (2)$$

The above equation is derived from the line integral proposed in [10], [11] which uses a given path $\mathbf{r}$

$$J(\mathbf{r}) = \frac{1}{2}\int_{\mathbf{r}} x(\mathbf{r})\, \mathrm{d}s$$

and includes the 50% *scaling factor* to account for the inhale-exhale cycle. For a constant speed υ, the line integral [19] simplifies to

$$J(t) = \frac{\upsilon}{2}\int_0^t y(\tau;\theta_s(\tau))\, \mathrm{d}\tau.$$

Absorbing the scaling factor and the speed to a constant, it simplifies to the proposed cost in (2).

The second problem of minimizing the accumulated cost at the free final time t_f is equivalent to selecting a path within the indoor environment Ω to reach the desired safety exit in minimum time t_f while ensuring that the accumulated amount at the final time t_f is below a prescribed threshold with $z(t_f) < z_{\text{threshold}}$.

The optimal guidance problem when the field is known and time-invariant has been solved in [1] and subsequently extended to known time-varying spatial fields in [2]. In both cases the navigation solution resulted in an open-loop controller which, in evacuation cases, makes it difficult to implement. A closed-loop approach utilizing level-set path-planning was presented for known time-invariant and known time-varying spatial fields in [10], [11]. An extension incorporated non-zero accumulated amounts (i.e., $z(0) \neq 0$) due to delays in decision making in [20] and which accounted for the psychological effects of freezing before making a

[1] The assumption that the human barycenter coincides with the sensing information is not conservative despite the fact that the nose is about $20-30$ cm from the barycenter.

decision to escape the contaminated indoor environments.

In the remainder of the paper, we make some assumptions on agent speed and the desired position.

Assumption 1 (indoor environment geometry): For the evacuation problem of a mobile agent in an indoor environment, it is assumed that

- The agent can move freely within the spatial domain Ω with a constant speed υ.
- The initial position of an agent $\theta_s(0)$ can be anywhere in the interior of the spatial domain, i.e., $\theta_s(0) \in \Omega \backslash \partial\Omega$.
- There is only one safety exit at the boundary $\partial\Omega$ with known coordinates $\theta_d = (\xi_d, \zeta_d) \in \partial\Omega$, see Figure 1.

A. Unknown spatial field modeling and parametrization

The spatial field is assumed to admit the parametrization

$$x(\theta) = \sum_{i=1}^{n} \alpha_i \phi_i(\theta), \quad \theta \in \Omega, \tag{3}$$

where $\alpha_i > 0$, $i = 1, \ldots, n$ are unknown constant coefficients and $\phi_i(\theta)$ are known spatial functions defined over the domain Ω. The summation limit n is not known but an upper bound is available. In such a case, one may have an overparameterized expansion

$$x(\theta) = \underbrace{\sum_{i=1}^{n_p} \alpha_i \phi_i(\theta)}_{\text{true}} + \underbrace{\sum_{i=n_p+1}^{n} \alpha_i \phi_i(\theta)}_{\text{overparameterization}} .$$

The true values $\alpha_i = 0$, for $i = n_p + 1, \ldots, n$ and the adaptive learning scheme ought to produce the estimates $\widehat{\alpha}_i = \alpha_i$, for $i = 1, \ldots, n_p$ and $\widehat{\alpha}_i = 0$, for $i = n_p + 1, \ldots, n$.

Using the parametrization (3), the spatial field measurements (1) obtained by the mobile agent are also given by

$$y(t; \theta_s(t)) = x(\theta)\Big|_{\theta=\theta_s(t)} = \sum_{i=1}^{n} \alpha_i \phi_i(\theta_s(t)),$$

written compactly as

$$y(t; \theta_s(t)) = \Phi^T(\theta_s(t))\alpha, \tag{4}$$

where

$$\alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \qquad \Phi(\theta_s(t)) = \begin{bmatrix} \phi_1(\theta_s(t)) \\ \vdots \\ \phi_n(\theta_s(t)) \end{bmatrix}.$$

The observations from the mobile agent (4) must be provided to a learning scheme to estimate the unknowns α_i, $i = 1, \ldots, n$. When static sensors are used, identifiability conditions require a minimum of n pointwise measurements to uniquely determine the unknown coefficients. Use of an adaptive scheme that produces the *adaptive estimates* $\widehat{\alpha}_i(t)$, of the coefficients α_i, $i = 1, \ldots, n$, imposes a *persistence of excitation* condition [16], which requires an associated integral involving the outer product of regressor vectors $\Phi(\theta_s(t))$ evaluated at the current position $\theta_s(t)$ of the mobile agent to have rank n. However, a single mobile agent, obtaining measurements (4), *can* induce the sought-after persistence of excitation by the appropriate motion within the domain Ω!

When a spatial gradient-based guidance scheme is used, the motion of the mobile agent within the spatial domain

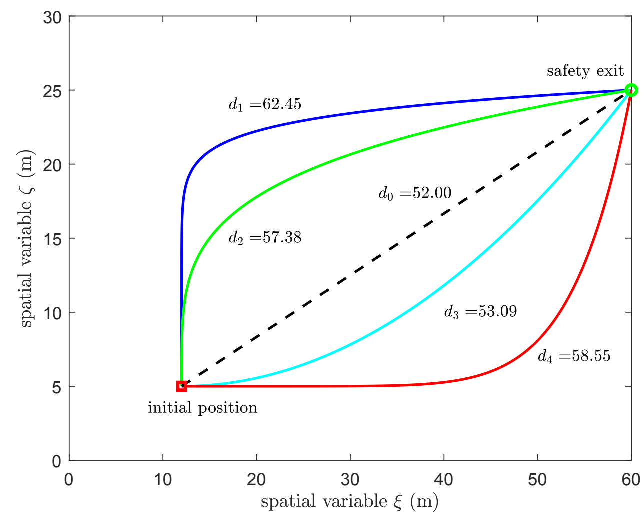

Fig. 1: Different escape paths result in different flight distances and associated escape times.

Ω provides a necessary condition for persistence of excitation (and hence parameter convergence in the sense $\lim_{t\to\infty} \widehat{\alpha}_i(t) = \alpha_i$, $\forall i$). When agent dynamics are used, then *any* agent guidance that satisfies the persistence of excitation becomes a necessary and sufficient condition for adaptive parameter convergence, see [15].

Problem statement: The guidance problem is to start at a given initial position $\theta_s(t_0)$ within the spatial domain Ω in an unknown spatial field $x(\theta)$ and use the mobile agent measurements (4) to estimate the unknown spatial field $x(\theta)$ and subsequently use the adaptive estimate $\widehat{x}(t, \theta)$ of the spatial field at each time to guide the agent towards the safety exit while minimizing the accumulated exposure to the unknown hazardous field with $z(t_f) < z_{\text{threshold}}$.

The above problem produces nested problems with increasing complexity. Parts of these tasks were addressed in earlier works in some form. The first one implements an adaptive learning scheme to estimate the coefficients α_i in (3) using any *admissible* guidance. Admissible guidance is any guidance that ensures adaptive learning in the sense $\lim_{t\to} \widehat{x}(t, \theta) = x(\theta)$ for all $\theta \in \Omega$. This was addressed in [15]. The next one builds upon the first one and additionally requires that one selects from the set of admissible guidances the one that guarantees that at some future time the agent reaches a desired position. The next level imposes the upper limit on the accumulated amount and selects from the admissible guidance set the one that ensures adaptive learning with convergence in the desired position *and* arriving at the accumulated amount falling below the threshold. The most complex case is when the guidance is selected to ensure adaptive learning and the adaptive learning is used to select the guidance that places the mobile agent at the desired location and ensuring $z(t_f) < z_{\text{threshold}}$.

III. Adaptive Estimation of Spatial Fields via Mobile Sensing Agents

The adaptive scheme (learning) when no desired final position is imposed and the mobile agent is free to roam Ω

simply to be able to learn the spatial field is studied in [15]. This is summarized here to make the paper self-contained.

The adaptive estimate of the spatial field is given by

$$\widehat{x}(t,\theta)=\sum_{i=1}^{n}\widehat{\alpha}_i(t)\phi_i(\theta), \quad \theta\in\Omega, \tag{5}$$

where $\widehat{\alpha}_i(t)$ are the adaptive estimates of the unknown α_i, $i=1,\ldots,n$. The associated *state estimation error* is given by

$$\begin{aligned} e(t,\theta) &= \widehat{x}(t,\theta)-x(\theta) \\ &= \sum_{i=1}^{n}\widehat{\alpha}_i(t)\phi_i(\theta)-\sum_{i=1}^{n}\alpha_i\phi_i(\theta) \\ &= \sum_{i=1}^{n}\widetilde{\alpha}_i(t)\phi_i(\theta), \end{aligned} \tag{6}$$

where $\widetilde{\alpha}_i(t)=\widehat{\alpha}_i(t)-\alpha_i$, $i=1,\ldots,n$ are the parameter errors.

To generate the update laws for the adaptive estimates $\widehat{\alpha}_i(t)$, one defines the *output estimation error* as the difference of the measured output (4) and the estimated state at the current sensor location $\theta_s(t)$

$$\begin{aligned} \varepsilon(t;\theta_s(t)) &= e(t,\theta_s(t))=\widehat{x}(t,\theta_s(t))-x(\theta_s(t)) \\ &= \widehat{y}(t;\theta_s(t))-y(t;\theta_s(t)) \\ &= \Phi^T(\theta_s(t))\widehat{\alpha}(t)-\Phi^T(\theta_s(t))\alpha \\ &= \Phi^T(\theta_s(t))\widetilde{\alpha}(t), \end{aligned} \tag{7}$$

where the parameter vector is

$$\widetilde{\alpha}(t)=\widehat{\alpha}(t)-\alpha=\left[\begin{array}{ccc}\widehat{\alpha}_1(t)-\alpha_1 & \ldots & \widehat{\alpha}_n(t)-\alpha_n\end{array}\right]^T.$$

A Lyapunov-redesign approach can be used here to obtain the adaptive laws for the adaptive parameter vector $\widehat{\alpha}(t)$. The motion (guidance) of the mobile agent that would provide the needed persistence of excitation requires additional process information, in addition to (4).

Assumption 2: The mobile agent can also provide the spatial gradients of the spatial field at its current sensor location $\theta_s(t)=(\xi_s(t),\zeta_s(t))$, given by

$$\begin{aligned} y_\xi(t;\theta_s(t)) &= \left.\frac{\partial x(\theta)}{\partial\xi}\right|_{\xi=\xi_s(t)}, \\ y_\zeta(t;\theta_s(t)) &= \left.\frac{\partial x(\theta)}{\partial\zeta}\right|_{\zeta=\zeta_s(t)}. \end{aligned} \tag{8}$$

Lemma 1 ([15]): Assume that the unknown spatial field $x(\theta)$ admits the expansion (3) and further assume that the sensor can provide spatial gradient information as given in Assumption 2. The update laws for the adaptive estimate of $x(\theta)$ in (5) are given by

$$\dot{\widetilde{\alpha}}(t)=\dot{\widehat{\alpha}}(t)=-\gamma\varepsilon(t)\Phi(\theta_s(t)), \tag{9}$$

where $\gamma>0$ is a user-defined adaptive gain, and the associated mobile sensor guidance is

$$\left\{\begin{array}{l} \dot{\xi}_s(t)=-\upsilon\,\mathrm{sign}\left(\varepsilon(t)\varepsilon_\xi(t)\right), \\ \dot{\zeta}_s(t)=-\upsilon\,\mathrm{sign}\left(\varepsilon(t)\varepsilon_\zeta(t)\right). \end{array}\right. \tag{10}$$

A compact form of the guidance (10) is

$$\dot{\theta}_s(t)=-\upsilon\,\mathrm{sign}\left(\varepsilon(t)\nabla\varepsilon(t)\right), \quad \nabla\varepsilon(t)=\left.\nabla e(t,\theta)\right|_{\theta=\theta_s(t)}.$$

The above adaptive scheme is realizable since it requires:

(i) the scalar output estimation error $\varepsilon(t;\theta_s(t))$
(ii) the regressor functions evaluated at the current sensor location $\phi_i(\theta_s(t))$, $i=1,\ldots,n$
(iii) the scalar gradients of the output estimation error $\varepsilon(t;\theta_s(t))$ which are the spatial gradients of the estimation error $e(t,\theta)$ evaluated at the current location

$$\left.\frac{\partial e(t,\theta)}{\partial\xi}\right|_{\xi=\xi_s(t)}, \quad \left.\frac{\partial e(t,\theta)}{\partial\zeta}\right|_{\zeta=\zeta_s(t)}.$$

The convergence results are as follows:

- If the following inner product of the regressor vector
$$\Phi^T(\theta_s(t))\Phi(\theta_s(t))\geq\beta>0, \tag{11}$$
is satisfied uniformly in time, then $\lim_{t\to\infty}\varepsilon(t)=0$.
- Parameter convergence in the sense $\lim_{t\to\infty}\widehat{\alpha}(t)=\alpha$ can only be established when a persistence of excitation condition is satisfied. It requires the time integral of the matrix resulting from the outer product of the regressor vector $\Phi(\theta_s(t))\Phi^T(\theta_s(t))$ be uniformly positive definite over any interval $[t,t+T_0]$ despite the fact that the $n\times n$ matrix defined by the outer product $\Phi(\theta_s(t))\Phi^T(\theta_s(t))$ is singular for each t. The PE condition is
$$c_1\mathbf{I}_n\geq\frac{1}{T_0}\int_t^{t+T_0}\Phi(\theta_s(\tau))\Phi^T(\theta_s(\tau))\,\mathrm{d}\tau\geq c_0\mathbf{I}_n, \tag{12}$$
for some positive scalars c_0,c_1,T_0.

The condition (11) required for the convergence of the output estimation error $\varepsilon(t;\theta_s(t))$ follows from the Lyapunov analysis with the Lyapunov functional selected as

$$V=\varepsilon^2(t)/(2\gamma).$$

Its time derivative along (7) is

$$\begin{aligned} \dot{V} &= \frac{1}{\gamma}\varepsilon(t)\frac{\mathrm{d}}{\mathrm{d}t}\left(\Phi^T(\theta_s(t))\widetilde{\alpha}(t)\right) \\ &= \frac{1}{\gamma}\varepsilon(t)\left[\Phi^T(\theta_s(t))\dot{\widetilde{\alpha}}(t)+\left(\frac{\mathrm{d}}{\mathrm{d}t}\Phi^T(\theta_s(t))\right)\widetilde{\alpha}(t)\right] \\ &= -\left(\Phi^T(\theta_s(t))\Phi(\theta_s(t))\right)\varepsilon^2(t)+\frac{\varepsilon(t)}{\gamma}\nabla^T\varepsilon(t)\dot{\theta}_s(t) \\ &= -\left(\Phi^T(\theta_s(t))\Phi(\theta_s(t))\right)\varepsilon^2(t)-\upsilon|\varepsilon(t)|\|\nabla\varepsilon(t)\| \\ &\leq -\left(\Phi^T(\theta_s(t))\Phi(\theta_s(t))\right)\varepsilon^2(t). \end{aligned}$$

In order to have $\dot{V}\leq-cV$ for some $c>0$ and subsequently ensure $\lim_{t\to\infty}\varepsilon(t)=0$, one must impose (11).

When human dynamics are used [21], [22], similar to vehicle dynamics for mobile agents, and have the form

$$\dot{\theta}_s(t)=F(\theta_s(t),u(t)), \tag{13}$$

then the guidance, expressed via the controller design $u(t)$, can ensure that the PE condition (13) is satisfied.

Lemma 2: When the mobile agent motion is dictated by the dynamics (13), then the guidance $u(t)$ selected to satisfy (12) results in a necessary *and* sufficient condition for the functional convergence

$$\lim_{t\to\infty}\widehat{x}(t,\theta)=x(\theta), \quad \forall\theta\in\Omega,$$

of the adaptive estimate with adaptation (9).

Remark 1: The guidance based on the control $u(t)$ requires only the regressor functions $\phi_i(\theta)$ to search for any path that ensures (12) is satisfied over any time interval. In this case the regressor vector $\Phi(\theta_s(t))$ parameterized by the admissible paths $\theta_s(t)$ is selected such that (12) is satisfied. This in turn provides the required control signal $u(t)$ for (13).

Remark 2: It is observed that when the guidance (10) is not used, the adaptive scheme (9) with the controller in (13) selected to satisfy (12) does not require the availability of the spatial gradients at the current sensor location in (8).

Now, when a desired final position $\theta_d = (\xi_d, \zeta_d)$ is imposed for the mobile agent, then the guidance resulting from the integrated learning and evacuation planning can be achieved in one of two different ways:

1) assume the availability of the spatial gradients of the unknown field at the current sensor location (Assumption 2) and incorporate the learning scheme (9) with the modified guidance
$$\begin{cases} \dot{\xi}_s(t) = -\upsilon \mathrm{sign}\left(\varepsilon(t)\varepsilon_\xi(t)\right) - \beta_\xi\left(\xi_s(t) - \xi_d\right) \\ \dot{\zeta}_s(t) = -\upsilon \mathrm{sign}\left(\varepsilon(t)\varepsilon_\zeta(t)\right) - \beta_\zeta\left(\zeta_s(t) - \zeta_d\right) \end{cases} \tag{14}$$
where $\beta_\xi, \beta_\zeta > 0$ are user-defined velocity gains.

2) do not assume the availability of the spatial gradients of the unknown field at the current sensor location and instead design the guidance law $u(t)$ in (13) such that
$$\lim_{t\to\infty} \theta_s(t) = \theta_d, \tag{15}$$
and at the same time ensure the PE condition (12) is satisfied.

The first option does not assume any agent motion dynamics and implements the learning (9) with the modified guidance (14). The second option selects the controller $u(t)$ from the family of controllers that generate candidate paths within Ω satisfying (12) and also satisfy the regulation (15).

When the final constraint on the limit of the accumulated amount $z(t)$ is included, the following optimal control problem formulation
$$\min \int_0^{t_f} 1 d\tau \quad \text{s.t.} \quad \begin{cases} z(t_f) \le z_{threshold} \\ \theta_s(t_f) = \theta_d \end{cases}$$
produces open-loop policies [1], [2]. When the concurrent estimation of the unknown spatial field is added in order to produce the appropriate guidance, the problem becomes computationally intractable.

In order to arrive in closed-loop policies with adaptive learning, then the above two path-dependent adaptive guidance schemes must be modified.

IV. Intermittent adaptive estimation of spatial fields via mobile sensing agents

In general, an adaptive-based guidance may not be performed as a single stage guidance; rather one has a *planning stage* and a *learning stage*. In the planning stage, the agent is not moving and uses the most recent estimate of the spatial field to move to the next way point. In the learning stage, where the agent is moving, it uses the measurements to implement the adaptive learning scheme.

We denote by τ_{plan} the duration of the planning stage and by τ_{travel} the duration of the learning stage (and also travel stage). The entire maneuver has a duration
$$\tau_{cycle} = \tau_{plan} + \tau_{travel}. \tag{16}$$
During the planning stage, the sensor is immobile and thus the agent speed is set to zero $\dot{\xi}_s = 0$, $\dot{\zeta}_s = 0$. During this time interval, the agent is using the current estimates of the spatial field to plan the guidance during the travel stage. The entire time interval is decomposed into N subintervals, each of duration τ_{cycle} with
$$t_i = (i-1)\cdot\tau_{cycle}, \quad i = 1, \ldots, N.$$
Alternatively, if a time interval is not specified, one considers cycles with duration τ_{cycle}. *Planning occurs* in times $t \in [t_i, t_i + \tau_{plan})$ and *travel occurs* during $t \in [t_i + \tau_{plan}, t_i + \tau_{cycle})$. The planning stage for $t \in [t_i, t_i + \tau_{plan})$ uses the state estimate from the previous cycle at $\widehat{x}(t_i, \widehat{\theta})$. In other words, while the state estimate $\widehat{x}(t,\theta)$ of the time-invariant function $x(\theta)$ is time varying, the planning state uses the recent estimate frozen at the end of the previous cycle, i.e., use an arrested adaptation estimate.

The activities in each stage are now summarized.

Planning stage: For $t \in [t_i, t_i + \tau_{plan})$, use the arrested estimate $\widehat{x}(t_i, \theta)$ which is time-invariant. Implement the level-set based guidance presented in [10], [11]. While the sensor is not moving, it continues to collect spatial field information and use it in the adaptation (9) modified to
$$\dot{\widehat{\alpha}}(t) = -\gamma\varepsilon(t)\Phi(\theta_s(t_i)), \quad t \in [t_i, t_i + \tau_{plan}). \tag{17}$$
Learning stage: For $t \in [t_i + \tau_{plan}, t_i + \tau_{cycle})$, the mobile agent uses the prescribed guidance developed at the planning stage and implements the adaptation
$$\dot{\widehat{\alpha}}(t) = -\gamma\varepsilon(t)\Phi(\theta_s(t)), \quad t \in [t_i + \tau_{plan}, t_i + \tau_{cycle}). \tag{18}$$
The above are tabulated in Algorithm 1.

Remark 3: Please note that the adaptation of the vector α is always active even when the sensor is immobile during the planning stage. This means that the adaptive estimation of the spatial field is always updated. However, there are two differences with a continuously adapted spatial field. The first one is that the arrested adaptive estimate $\widehat{x}(t_i, \theta) = \Phi^T(\theta)\widehat{\alpha}(t_i)$, which is a time-invariant function, is used during the planning stage and the second one is that during the planning stage the adaptation uses (17) and not (18).

V. Numerical examples

The unknown spatial field is selected as
$$x(\theta) = \sum_{i=1}^{3}\sum_{j=1}^{2} \alpha_{ij} g_i(\xi) h_j(\zeta) = \sum_{i=1}^{6} a_i \phi_i(\theta),$$
where
$$a_1 = \alpha_{11} = 2, \quad a_2 = \alpha_{12} = 1, \quad a_3 = \alpha_{21} = 1.75,$$
$$a_4 = \alpha_{22} = 2, \quad a_5 = \alpha_{31} = 1.75, \quad a_6 = \alpha_{32} = 1.25,$$

Algorithm 1 Arrested adaptation-based evacuation guidance in $[t_i, t_i + \tau_{cycle}[$

1: **initialize:** Select the planning stage τ_{plan} and travel stage τ_{travel} durations. Using the initial estimates of the parameters $\widehat{\alpha}(0)$, set up the initial estimate $\widehat{x}(0;\theta)$ of the unknown spatial field. Using the initial agent location $(\xi_s(0), \zeta_s(0))$ and the desired location (escape exit) (ξ_d, ζ_d), determine the first path planning for the interval $t \in [\tau_{plan}, \tau_{plan} + \tau_{travel})$. Obtain sensor measurements in both $t \in [0, \tau_{plan})$ and $t \in [\tau_{plan}, \tau_{plan} + \tau_{travel})$ and use them to implement the adaptive learning (9).
2: **iterate:** $i = 2$
3: **loop**
4: Define next cycle $[t_i, t_i + \tau_{cycle}) = [(i-1)\tau_{cycle}, i\tau_{cycle})$ with $t_i = (i-1)\tau_{cycle}$. For each $t \in [t_i, t_i + \tau_{cycle})$ continue to obtain sensor measurements and implement the adaptation (17) regardless of the agent motion.
5: In the i^{th} *planning stage* of duration τ_{plan} with $t \in [t_i, t_i + \tau_{plan})$, use the most recent arrested estimate of the spatial field $\widehat{x}(t_i, \theta)$ to plan the path for $t \in [t_i + \tau_{plan}, t_i + \tau_{cycle})$ using the level-set based guidance in [10], [11]. Continue the adaptation (18).
6: In the i^{th} *travel stage* of duration τ_{travel} with $t \in [t_i + \tau_{plan}, t_i + \tau_{cycle})$, implement the level-set based path planning developed at the most recent planning stage.
7: At the end of the i^{th} *cycle* $t_{i+1} = t_i + \tau_{cycle}$, update the adaptive estimate of the spatial field using $\widehat{\alpha}(t_{i+1})$
8: **if** $\sqrt{(\xi_s(t_{i+1}) - \xi_d)^2 + (\zeta_s(t_{i+1}) - \zeta_d)^2} > 0$ **then**
9: $i \leftarrow i + 1$
10: **goto** 2
11: **else**
12: terminate-reached safety exit
13: **end if**
14: **end loop**

defined over the spatial domain $\Omega = [0, L_\xi] \times [0, L_\zeta] = [0, 60] \times [0, 30]$m. The regressor functions are given by

$$g_i(\xi) = 50e^{-(\xi - \mu_{\xi,i})^2/(2\sigma_\xi^2)}, \quad \mu_{\xi,i} = \frac{L_\xi(2i+1)}{10}, \quad \sigma_\xi = \frac{L_\xi}{12},$$
$$h_j(\zeta) = 40e^{-(\zeta - \mu_{\zeta,j})^2/(2\sigma_\zeta^2)}, \quad \mu_{\zeta,j} = \frac{L_\zeta j}{3}, \quad \sigma_\zeta = \frac{L_\zeta}{7}.$$

The adaptive guidance (14) was implemented with $\upsilon = 7$m/s, $\beta_\xi = 0.3, \beta_\zeta = 0.8$ and the scheme (17), (18) had $\gamma = 0.1$. For initial guess of the adaptive estimates, we used $\widehat{a}_i(0) = 0.1$ and the initial sensor location was $\xi_s(0) = L_\xi/5$, $\zeta_s(0) = L_\zeta/6$ with the desired safety exit at $(\xi_d, \zeta_d) = (L_\xi, \frac{5L_\zeta}{6})$.

A baseline learning scheme uses a straight path with

$$\dot{\xi}_s = \upsilon\cos(\theta), \ \ \dot{\zeta}_s = \upsilon\sin(\theta), \ \ \theta = \tan^{-1}\left(\frac{\zeta_d - \zeta_s(0)}{\xi_d - \xi_s(0)}\right).$$

It is easily seen that a straight path from $(\xi_s(0), \zeta_s(0))$ to (ξ_d, ζ_d) is 52m long and requires 7.42s to complete. Based on the selected expression for $x(\theta)$, in such a trajectory, the total accumulated amount is $z(7.42) = 18,763.3$. The adaptive guidance scheme requires $t = 20.64$s to reach safety and the total distance travelled is 60.9m. The accumulated amount at the exit in this case is $z(20.4) = 68\%$ of the straight path case. Table I examines the effects of the travel duration τ_{travel} with a fixed $\tau_{cycle} = 6$s on the accumulated amount and the time to safety. The nonlinear dependence of the escape time on the variation of τ_{travel} is due to the fact that adaptation continues even during the τ_{plan} stage. In this particular case, the best value for $\tau_{travel} = 4$s with $\tau_{plan} = 2$s as it yields the smallest possible accumulated amount, which is 68% of the straight path case. Figure 2 depicts the distance from safety for both cases. It is observed that with the fixed path guidance, an escapee reaches safety in minimum time as expected whereas with the adaptive guidance reaches after 20s. However, it reaches safety with

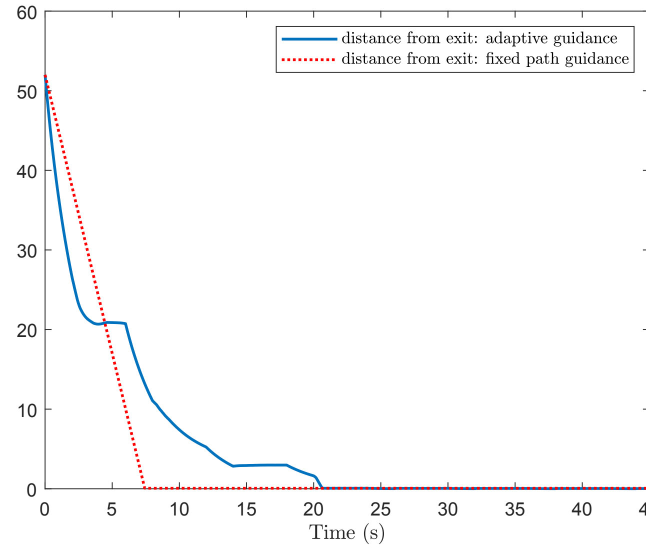

Fig. 2: Distance of agent from safety exit; straight-path trajectory is completed in $t = 7.4286$s whereas the adaptive guidance requires $t = 20.64$s.

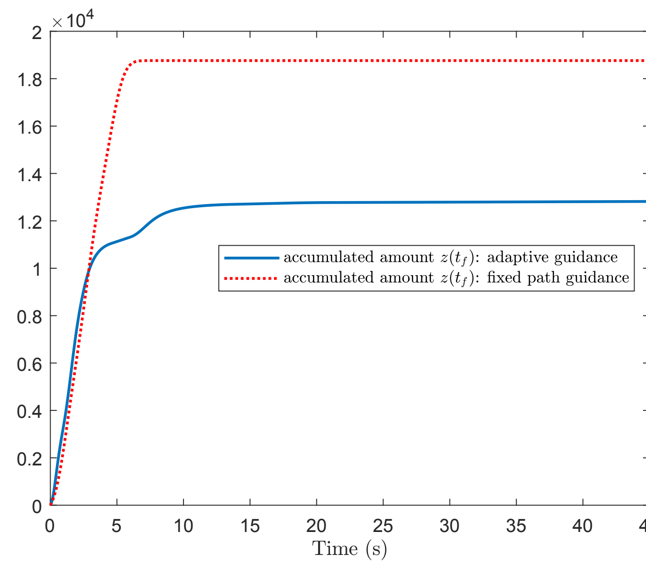

Fig. 3: Accumulated amount using a straight-path trajectory and the proposed adaptive one; the accumulated amount at exit is $z(7.42) = 1.8763 \times 10^4$ for the straight-path trajectory, whereas the adaptive case has $z(20.64) = 1.7963 \times 10^4$.

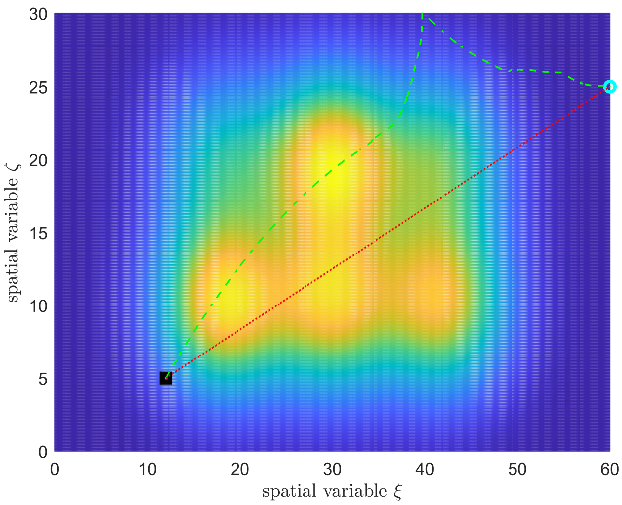

Fig. 4: Fixed path (···) and proposed adaptive path (−−) in indoor environment.

stage τ_{travel} (s)	escape time (s)	accumulated amount (%)
1	11.74	0.7157
2	16.30	0.7195
3	15.46	0.7224
4	20.64	0.6832
5	44.70	0.9570
straight path	7.42	1

TABLE I: Effects of the travel duration τ_{travel} in a fixed cycle $\tau_{cycle} = \tau_{plan} + \tau_{travel} = 6$s on the time to safety, and the accumulated amount as a percentage of the amount accumulated using the straight path.

a much smaller amount equal to 68% of the amount for the fixed path case. Figure 3 depicts the accumulated amount up to time t and thus for the fixed path guidance it reports an amount of $z(7.4286) = 1.8763 \times 10^4$ concentration units. The corresponding amount for the adaptive guidance is roughly 68% of that quantity reported at $t = 20.64$s. A top-view of the two trajectories over the spatial distribution of the spatial field are depicted in Figure 4.

VI. Conclusions

A scheme combining field-dependent guidance of an evacuee in indoor contaminated environments and adaptive learning of a spatial field with a mobile sensor was presented. The earlier level-set based guidance for human evacuation over harmful environments described by spatial functions assumed that the spatial field was constant in time and known. To utilize such a field-dependent guidance, the time-varying estimate of the spatial field was frozen in time at a particular time instance within the planning cycle via an arrested adaptive scheme in order to utilize the level-set based planning. This ensured that the accumulated amount of the substance over the duration of the evacuation was below a prescribed threshold. Numerical results depicting aspects of the field-dependent guidance were included.

An immediate extension involves the inclusion of multiple safety exits at the indoor environment boundary along with a decision to switch to a different safety exit mid-flight by continuously assessing the projected accumulated amounts.

References

[1] M. A. Demetriou and E. Bakolas, "Navigating over 3d environments while minimizing cumulative exposure to hazardous fields," *Automatica*, vol. 115, p. 108859, 2020.

[2] ——, "Optimal navigation over spatiotemporally varying hazardous fields," in *Proc. of the IEEE Conf. on Decision and Control*, 2021, pp. 2607–2612.

[3] E. F. F. Zermelo, "Über das navigationsproblem bei ruhender oder veränderlicher windverteilung," *Zamm-zeitschrift Fur Angewandte Mathematik Und Mechanik*, vol. 11, pp. 114–124, 1931.

[4] A. V. Savkin and H. Huang, "Optimal aircraft planar navigation in static threat environments," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 53, no. 5, pp. 2413–2426, 2017.

[5] D. Marthaler and A. L. Bertozzi, "Tracking environmental level sets with autonomous vehicles," in *Recent Deveopments in Cooperative Control and Optimization*, ser. Cooperative Systems. Kluwer Academic Publishing, 2013, vol. 3, pp. 317–332.

[6] T. Lolla, M. P. Ueckermann, K. Yiğit, P. J. Haley, and P. F. J. Lermusiaux, "Path planning in time dependent flow fields using level set methods," in *Proc. of the IEEE International Conf. on Robotics and Automation*, 2012, pp. 166–173.

[7] A. S. Matveev and M. S. Nikolaev, "Hybrid control for tracking environmental level sets by nonholonomic robots in maze-like environments," *Nonlinear Analysis: Hybrid Systems*, vol. 39, 2021.

[8] C. A. Parkinson, *Models for Human Navigation and Optimal Path Planning Using Level Set Methods and Hamilton-Jacobi Equations*. ProQuest LLC, Ann Arbor, MI, 2020, thesis (Ph.D.)–University of California, Los Angeles.

[9] K. Dantu and G. S. Sukhatme, "Detecting and tracking level sets of scalar fields using a robotic sensor network," in *Proc. of the IEEE International Conf. on Robotics and Automation*, 10-14 April 2007.

[10] M. A. Demetriou and M. Kontopyrgos, "A level-set based approach for the path planning of human evacuations in contaminated indoor environments," in *Proc. of the European Control Conference*, Naples, Italy, June 25-28 2019, pp. 1914–1919.

[11] ——, "A path planning algorithm for human evacuations with an environment dependent motion," in *Proc. of the IEEE Conf. on Decision and Control*, December 11-13 2019, pp. 274–279.

[12] X. Lan and M. Schwager, "Rapidly exploring random cycles: Persistent estimation of spatiotemporal fields with multiple sensing robots," *IEEE Trans. on Robotics*, vol. 32(5), pp. 1230–1244, Oct 2016.

[13] ——, "Learning a dynamical system model for a spatiotemporal field using a mobile sensing robot," in *Proc. of the American Control Conference*, 24-26 May 2017, pp. 170–175.

[14] H. Bai, "Motion-dependent estimation of a spatial vector field with multiple vehicles," in *Proc. of the IEEE Conf. on Decision and Control*, 2018, pp. 1379–1384.

[15] M. A. Demetriou, "Adaptive techniques for the online estimation of spatial fields with mobile sensors," *IEEE Transactions on Automatic Control*, vol. 68, no. 9, pp. 5669–5676, 2023.

[16] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Englewood Cliffs, NJ: Prentice Hall, 1995.

[17] G. Roussos, D. V. Dimarogonas, and K. J. Kyriakopoulos, "3D navigation and collision avoidance for nonholonomic aircraft-like vehicles," *Internat. J. Adapt. Control Signal Process.*, vol. 24, pp. 900–920, 2010.

[18] C. Vrohidis, P. Vlantis, C. P. Bechlioulis, and K. J. Kyriakopoulos, "Prescribed time scale robot navigation," *IEEE Robotics and Automation Letters*, vol. 3(2), pp. 1191–1198, 2018.

[19] W. Kaplan, *Advanced Calculus*, 3rd ed. Addison-Wesley Publishing Company, Reading, MA, 1984.

[20] A. C. Dings, C. M. Gatsonis, and M. A. Demetriou, "A psychological approach for the path planning of human evacuations in contaminated indoor environments," *IFAC-PapersOnLine*, vol. 53(2), 2020.

[21] P. Kachroo, S. Al-nasur, S. Wadoo, and A. Shende, *Pedestrian Dynamics: Feedback Control of Crowd Evacuation*. Berlin-Heidelberg: Springer-Verlag, 2008.

[22] E. Cristiani, B. Piccoli, and A. Tosin, *Multiscale modeling of pedestrian dynamics*. Springer, Cham, 2014, vol. 12.