Stochastic Source Seeking for Mobile Robots in Obstacle Environments Via the SPSA Method

Eduardo Ramírez-Llanos and Sonia Martínez

***Abstract*—This paper considers a class of stochastic source-seeking problems to drive a mobile robot to the minimizer of a source signal. Our approach is first analyzed in an obstacle-free scenario, where measurements of the signal at the robot location and information of a contact sensor are required. We extend our results to environments with obstacles under mild assumptions on the step size. Our approach builds on the simultaneous perturbation stochastic approximation idea to obtain information of the signal field. We prove the practical convergence of the algorithms to a ball whose size depends on the step size that contains the location of the source. The novelty relies in that we consider nondifferentiable convex functions, a fixed step size, and the environment may contain obstacles. Our proof methods employ nonsmooth Lyapunov function theory, tools from convex analysis, and stochastic difference inclusions. Finally, we illustrate the applicability of the proposed algorithms in a two-dimensional scenarios.**

***Index Terms*—Algorithm design and analysis, intelligent robots, nonlinear dynamical systems, optimization, stochastic systems.**

I. Introduction

Stochastic source-seeking algorithms are used in mobile robotics to find a source of a radiation-like signal in GPS-denied environments. Applications range from biology, in understanding bacterial foraging, to security, for search and rescue operations and chemical detection. In a typical setting, the robot samples the signal emitted by the source by exploring the environment through a stochastic motion. The samples are used to steer and climb the gradient of the signal field, where this field might represent the spatial distribution of magnetic force, a thermal signal, or a chemical concentration.

Our approach is inspired by the simultaneous perturbation stochastic approximation (SPSA) algorithm. The SPSA method is a well-known algorithm usually applied to estimate the gradient of a cost function from measurements. It was first proposed in [1] and since then it has been successfully applied in many optimization problems, such as parameter estimation, simulation optimization, resource allocation, and robotics.

The SPSA method uses a monotonically decreasing step size to solve an unconstrained optimization problem. In mobile robots, a decreasing step size is not an option since it is impossible to navigate with infinite precision, which is implied by a step size converging to zero. We propose a modified version of the SPSA algorithm that uses a small, but constant, step size for environments that may contain obstacles.

Manuscript received August 24, 2017; revised January 6, 2018; accepted June 30, 2018. Date of publication July 10, 2018; date of current version March 27, 2019. This work was supported in part by the AFOSR award 11RSL548, in part by the NSF-CMMI award 1434819, and in part by the Colciencias award 784-2017. Recommended by Associate Editor U. V. Shanbhag. *(Corresponding author: Eduardo Ramírez-Llanos.)*

The authors are with the Department of Mechanical and Aerospace Engineering, University of California, La Jolla, CA 92093 USA (e-mail: ejllanos@gmail.com; soniamd@ucsd.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2018.2854646

A. Literature Review

There are many approaches to stochastic source seeking for mobile robots in position-denied environments. An example is given in [2], where an extremum seeking approach is employed with nonholonomic vehicles, or the application of the SPSA algorithm to mobile robots in [3] and [4]. We follow the approach provided in [3], where Azuma *et al.* designed a controller to drive a robot to the source by applying the SPSA algorithm without the use of the position information. The algorithm they proposed uses standard assumptions on SPSA, such as the thrice differentiability of the cost function and a monotonically decreasing step size. Those assumptions fit very well in many applications where direct measurements of the gradient are not available. An alternative is to use a small constant step size, which has been successfully applied in diverse areas, such as combustion control [5], mobile robots [4], and tracking and adaptive control [6].

In [5], a variation of the SPSA algorithm is proposed that decreases the oscillation against the constraints. The proposed algorithm is applied to an automotive combustion engine problem. Although [5] uses a constant step size, no theoretical guarantees are given for fixed step sizes. A model-free algorithm is proposed in [4], based on stochastic approximation to find a source in environments with obstacles and using a constant step size. A decreasing step size is not desirable because the robot might get trapped in a location where the magnitude of the gradient is small. The convergence of the algorithm in [4] is shown through an experiment in a real-world scenario; however, no theoretical guarantees are provided.

In [6], an algorithm inspired by SPSA is proposed for unconstrained optimization. The algorithm uses a constant step size to minimize a cost function for three different tracking problems: a random walk, an optimization of a unmanned aerial vehicle (UAV's) flight, and a load balancing. A drawback of their algorithm is that the cost function is assumed to be once differentiable and it solves an unconstrained optimization problem. In [7], He *et al.* studied the convergence of the SPSA method when the cost function is nondiferentiable. However, the analysis also employs a decaying step size.

Following a different line to the gradient-free algorithms, Taylor and LaValle [8] proposed an approach to guide a robot through an unknown obstacle environment using sensed information from a single intensity source. The algorithm is similar in spirit to the well-known family of bug algorithms [9]. Bug algorithms are classic reactive motion-planning algorithms for point robots that have a limited knowledge of their environment. With respect to previous algorithms in the literature, the approach provided in [8] requires less sensing information. A drawback of their approach is the assumption on the availability of the gradient generated by the intensity source, which in general may not be true.

B. Statement of Contributions

We propose a stochastic source-seeking algorithm to drive a robot to an unknown source signal by only using measurements of the signal field. Our algorithm builds on the SPSA algorithm. The novelty of our approach is that we consider nondifferentiable convex functions, fixed step size, and the environment may have obstacles. We prove practical convergence to a ball whose size depends on the step size that contains the location of the source. For the proof, we use the Lyapunov theory together with tools from convex analysis and stochastic difference inclusions. Our proof does not rely on the stochastic approximation theory as is usually the case for algorithms in the literature based on SPSA. Finally, we show the applicability of the proposed algorithm in a two-dimensional (2-D) scenario for the source-seeking problem.

II. Preliminaries

This section presents notation, notions of convex analysis, and stochastic stability theory that are used in the following.

A. Notation

We denote by $\mathbb{Z}_{\geq 0}$ the set of nonnegative integers, $\mathbb{Z}_{>0}$ the set of positive integers, $\mathbb{R}^n_{>0}$ the positive orthant of $\mathbb{R}^n$, for some $n \in \mathbb{Z}_{>0}$, and I_n the identity matrix of size $n \times n$. For $x = [x_1, \ldots, x_n] \in \mathbb{R}^n$ with nonzero entries, we define $x^{-1} \triangleq [x_1^{-1}, \ldots, x_n^{-1}]^\top$. The two-norm of a vector is denoted by $\|.\|$. A function f is $o(h)$, and we write $f(x) = o(h(x))$ as $x \to x_0$, if $\lim_{x \to x_0} \frac{f(x)}{h(x)} = 0$. A function f is $O(h)$, and we write $f(x) = O(h(x))$ as $x \to x_0$, if there exists $\delta, M \in \mathbb{R}_{>0}$ such that $\|f(x)\| \leq M\|h(x)\|$ for $\|x - x_0\| \leq \delta$. For a closed set $S \subset \mathbb{R}^n$ and $x \in \mathbb{R}^n$, $|x|_S = \inf_{y \in S} |x - y|$ is the Euclidean distance to S. A function $\phi : \mathbb{R}^n \to \mathbb{R}$ is *upper semicontinuous* if $\limsup_{i \to +\infty} \phi(x_i) \leq \phi(x)$ whenever $\lim_{i \to +\infty} x_i = x$. Given sets S and T, a *set-valued map*, denoted by $h : S \rightrightarrows T$, associates an element of S with a subset of T. The symbol $\mathbb{I}_S(x)$ denotes the indicator function of $\mathbb{I}_S$. A set-valued map $M : \mathbb{R}^p \rightrightarrows \mathbb{R}^n$ is *outer semicontinuous* if, for each sequence $(x_i, y_i) \to (x, y)$ as $i \to +\infty$, in $\in \mathbb{R}^p \times \mathbb{R}^n$, and satisfying $y_i \in M(x_i)$ for all $i \in \mathbb{Z}_{\geq 0}$, it holds that $y \in M(x)$. A mapping M is *locally bounded* if, for each bounded set $K \subset \mathbb{R}^p$, $M(K) \triangleq \cup_{x \in K} M(x)$ is bounded.

B. Convex Analysis Notions

The notions we introduce here follow [10] and [11]. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a closed, proper, and convex function. The *subgradient* of f is the set-valued map $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ defined by the subgradient set $\partial f(x) = \{\xi \in \mathbb{R}^n \mid f(x') \geq f(x) + \xi^\top (x' - x)\}$. We refer to $df(x)$ as the *semiderivative* function, which is the support function of the nonempty, compact, and convex set $\partial f(x)$, that is, $df(x)(w) = \sup\{\xi^\top w \mid \xi \in \partial f(x)\}$. The first-order expansion of f for any point x is given by

$$f(x + w) = f(x) + df(x)(w) + o(\|w\|). \tag{1}$$

We say that f satisfies the *superquadratic growth condition* if there exists $\rho > 0$ such that

$$f(y) \geq f(x) + df(x)(y - x) + \frac{\rho}{2}\|y - x\|^2 \tag{2}$$

for $x, y \in \mathbb{R}^n$. In particular, a strongly convex function satisfies the superquadratic growth condition. When f is differentiable, the superquadratic growth condition is equivalent to assuming that $\rho I_n \leq \nabla^2 f(x)$, for $x \in \mathbb{R}^n$.

C. Stability for Stochastic Difference Inclusions

The notions we introduce here follow [12]. Consider a discrete-time, stochastic difference inclusion

$$x^+ \in \mathcal{H}_\alpha(x, v^+), \quad v \sim \mu \tag{3}$$

where x^+ is the state after an instantaneous change, $\mathcal{H}_\alpha : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is a set-valued map for some $n, m \in \mathbb{Z}_{>0}$ parameterized by $\alpha \in \mathbb{R}_{>0}$, which assigns nonempty set values, and $x \in \mathbb{R}^n$ is the state. The notation v^+ and v refers to sequences of random input variables as explained next. Consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω denotes the set of all possible outcomes, $\mathcal{F}$ is the σ-field associated with Ω, and $\mathbb{P}$ is the probability function that assigns a probability to events in $\mathcal{F}$. In particular, we assume $\mathbf{B}(\mathbb{R}^m) \subseteq \mathcal{F}$, where $\mathbf{B}(\mathbb{R}^m)$ is the Borel field. In (3), we use v^+ and v as a place holder for a sequence of independent, identically distributed (i.i.d.) random variables $\mathbf{v} \triangleq \{\mathbf{v}_k\}_{k=0}^\infty$; that is, $\mathbb{P}(\mathbf{v}_k \in F) = \mathbb{P}(\{w \in \Omega \mid \mathbf{v}_k(w) \in F\})$ is well defined and independent of k for each $F \in \mathbf{B}(\mathbb{R}^m)$. We use $\mathcal{F}_k$ to denote the collection of sets $\{w \in \Omega \mid (\mathbf{v}_0(w), \ldots, \mathbf{v}_k(w)) \in F\}$, where $F \in \mathbf{B}((\mathbb{R}^m)^{k+1})$, which are the sub-$\sigma$-fields of $\mathcal{F}$ that form the minimal filtration of the sequence $\mathbf{v}$. Due to the i.i.d. property, each random variable has the same probability measure $\mu : \mathbf{B}(\mathbb{R}^m) \to [0, 1]$ defined as $\mu(F) \triangleq \mathbb{P}(\mathbf{v}_k \in F)$ and, for almost all $w \in \Omega$

$$E[f(\mathbf{v}_0, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1})|\mathcal{F}_k](w) = \int_{\mathbb{R}^m} f(\mathbf{v}_0(w), \ldots, \mathbf{v}_k(w), v)\mu(dv)$$

for each $k \in \mathbb{Z}_{\geq 0}$ and each measurable $f : (\mathbb{R}^m)^{k+2} \to \mathbb{R}$.

The sequence of random variables $\mathbf{x} \triangleq \{\mathbf{x}_k\}_{k \geq 0}$, where $\mathbf{x}_k : \operatorname{dom} \mathbf{x}_k \subset \Omega \to \mathbb{R}^n$, $k \in \mathbb{Z}_{\geq 0}$ with $\mathbf{x}_0 = x$ for all $w \in \Omega$ and $\operatorname{dom} \mathbf{x}_{k+1} \subset \operatorname{dom} \mathbf{x}_k$, is called a random process starting at $x \in \mathbb{R}^n$. We say that $\mathbf{x}$ is *adapted to the natural filtration* of $\mathbf{v}$ if $\mathbf{x}_{k+1}$ is $\mathcal{F}_k$ measurable for each $k \in \mathbb{Z}_{\geq 0}$, i.e., $\mathbf{x}_{k+1}^{-1}(F) \in \mathcal{F}_k$ for each $F \in \mathbf{B}(\mathbb{R}^m)$.

Let $\mathbf{x}$ be a random process starting from $x \in \mathbb{R}^n$, which is adapted to the natural filtration of $\mathbf{v}$. Let $J_{\mathbf{x}} : \Omega \to \mathbb{Z}_{\geq 0} \cup \{\infty\}$ be a random variable that denotes the number of elements in the sequence $\mathbf{x}$. Then, $\mathbf{x}$ is a *random solution* of (3) starting at $x \in \mathbb{R}^n$, denoted as $\mathbf{x} \in \mathcal{S}(x)$, if $\mathbf{x}_0 = x$, $\mathbf{x}_{k+1}(w) \in \mathcal{H}_\alpha(\mathbf{x}_k(w), \mathbf{v}_{k+1}(w))$ for all $w \in \operatorname{dom} \mathbf{x}_{k+1} \triangleq \{w \in \Omega \mid k + 1 \leq J_{\mathbf{x}}\}$ and $k \in \mathbb{Z}_{\geq 0}$. We impose the following regularity condition on $\mathcal{H}$.

Assumption 1: $\mathcal{H}$ is locally bounded and $v \to \operatorname{graph}(\mathcal{H}_\alpha(\cdot, v)) \triangleq \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mid y \in \mathcal{H}_\alpha(x, v)\}$ is measurable with closed values.

Definition 1: (Mean-square practically exponentially stable (MSP-ES) equilibrium): We say that the equilibrium point of (3) is MSP-ES if there exists $\alpha^* \in (0, 1)$, positive real numbers β, $\lambda < \frac{1}{\alpha^*}$, γ, and η such that for all $\alpha \in (0, \alpha^*]$, we have

$$E[\|x_k\|^2] \leq \beta(1 - \alpha\lambda)^k \|x_0\|^2 + \gamma\alpha^\eta \qquad \forall k \in \mathbb{Z}_{\geq 0}.$$

Proposition 1 (see [13]): Consider the system (3) under Assumption 1. If there exists an upper semicontinuous function $V : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, positive constants c_1, c_2, λ, K, $\alpha^* \in (0, 1)$, and $\eta > 1$ such that for all $\alpha \in (0, \alpha^*)$

$$c_1\|x\|^2 \leq V(x) \leq c_2\|x\|^2$$
$$\int_{\mathbb{R}^m} \max_{h \in \mathcal{H}_\alpha} V(h)\mu(dv) \leq (1 - \alpha\lambda)V(x) + \alpha^\eta K \tag{4}$$

then, the equilibrium point is MSP-ES for (3). •

III. Problem Statement

This section describes the stochastic source-seeking problem for GPS-denied environments. The problem has been studied for obstacle-free environments in, e.g., [2], [3], and [4]. In particular, we follow the approach provided in [3], except that we consider boundaries and obstacles in the environment. Suppose that a sufficiently small robot moves in $\mathbb{R}^n$ and its motion is described in the world coordinate frame by

$$\left(\dot{p}^\top, \dot{\theta}^\top, \dot{\phi}^\top\right)^\top = G(p(t), \theta(t), \phi(t))u(t) \tag{5}$$

where $G : \mathbb{R}^n \times \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^{(n+n_1+n_2)\times m}$ is a function describing the robot dynamics, $p(t) \in \mathbb{R}^n$ and $\theta(t) \in \mathbb{R}^{n_1}$ are the translational and orientational positions in the world of coordinate frame, and $\phi(t) \in \mathbb{R}^{n_2}$ and $u(t) \in \mathbb{R}^m$ are the internal posture and the control input, respectively.

Let $\mathcal{E}$ be the *environment* where the robot moves, which is assumed to be convex and compact. A *tower* broadcasts a signal, which is modeled by an intensity function f over $\mathbb{R}^n$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be the signal mapping, in which $f(p)$ yields the intensity at $p \in \mathbb{R}^n$, generated from a tower at p^*. The location of the tower p^* can be or not in $\mathcal{E}$. The environment $\mathcal{E}$ and the signal mapping f are unknown to the robot. The robot aims to solve the following optimization problem without knowledge of its absolute position (p, θ)

$$\min_{p\in\mathcal{E}} f(p) \tag{6}$$

by only using (noisy) measurements of $f(p(t))$. We consider two scenarios for $\mathcal{E}$: when $\mathcal{E}$ does not have obstacles and when it does. For the obstacle-free scenario, we assume $\mathcal{E}$ is a convex compact set. In both scenarios, the problem is to design an algorithm with guaranteed practical convergence to a small ball containing p^*. When $p^* \notin \mathcal{E}$, the robot should converge in practical way to a ball containing the closest point from $\mathcal{E}$ to p^*. We cannot use standard algorithms based on the explicit form of f or its gradient because the expression for f is not available. The robot only has measurements given by sensors, and the measurements of f may be noisy but we neglect this noise.

The robot is equipped with two sensors. First, a contact sensor $l_\mathcal{E}$ that indicates whether the robot is touching the environment boundary or any obstacle inside the environment, i.e., $l_\mathcal{E}(p) = 1$ if $p \in \partial\mathcal{E}$ and $l_\mathcal{E}(p) = 0$ otherwise. Second, the robot is equipped with an intensity sensor l_I. It indicates the strength of the signal from position p, i.e., $l_I(p) = f(p)$. Since the robot does not have position information in the coordinate frame, it is necessary to adapt (5) to a body fixed frame. The position of the robot in the body frame at time t is given by

$$\begin{pmatrix} z(t) \\ \psi(t) \\ \varphi(t) \end{pmatrix} = \begin{pmatrix} R(-\theta(\tau))(p(t) - p(\tau)) \\ \theta(t) - \theta(\tau) \\ \phi(t) \end{pmatrix}$$

where t expresses a future time after τ, $(z(t), \psi(t), \varphi(t)) \in \mathbb{R}^n \times \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ are the new coordinates, and $R(-\theta(\tau))$ is the rotation matrix of an angle $-\theta(\tau)$.

IV. Proposed Algorithm for the Obstacle-Free Scenario

In this section, we assume that there are no obstacles in $\mathcal{E}$, where $\mathcal{E}$ is described by a convex compact set. To find p^*, we propose the following algorithm, which is similar in spirit to the stochastic approximation algorithm for fixed step size:

$$p_{k+1} = \Pi_\mathcal{E}[p_k - \alpha g(p_k, \delta(p_k, R_k v_k), R_k, v_k)] \tag{7}$$

where $k \in \mathbb{Z}_{\geq 0}$ and we assume that R_k is a sequence of rotation matrices set *a priori*; see below a justification of terms for the algorithm. To simplify the notation for aid in analysis, we write the above algorithm as a discrete-time dynamical system as follows $p^+ = \Pi_\mathcal{E}[p - \alpha g(p, \delta(p, Rv), R, v)]$, where $p \in \mathbb{R}^n$ is the current state, $p^+ \in \mathbb{R}^n$ is the state at the next time step, $\Pi_\mathcal{E}$ is the projection on a convex compact set $\mathcal{E}$ (i.e., $\Pi_\mathcal{E}[p] = \mathrm{argmin}_{z\in\mathcal{E}} \|z - p\|_2$), and $g : \mathbb{R}^n \times \mathbb{R}^2 \times \mathrm{SO}(n) \times \mathbb{R}^n \to \mathbb{R}^n$ is given as follows:

$$g(p, \delta(p, Rv), R, v) = \begin{cases} R\frac{f(p+\delta_1 Rv)-f(p-\delta_2 Rv)}{\delta_1+\delta_2} v^{-1}, & \text{if } \delta_1 + \delta_2 > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

Here, $f : \mathbb{R}^n \to \mathbb{R}$ is the function to be minimized, $\alpha \in \mathbb{R}_{>0}$ is the step size, and $R \in \mathrm{SO}(n)$ is the uncertain time-varying rotation matrix, which by definition is an orthogonal matrix. In (8), $\delta = (\delta_1, \delta_2)$ is a mapping $\delta_i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, for $i \in \{1, 2\}$, which is given by

$$\delta_1(p, Rv) = \begin{cases} \bar{\delta}_1, & \text{if } p + \bar{\delta}_1 Rv \in \mathcal{E} \\ \mathrm{dist}_{+Rv}(p, \partial\mathcal{E}), & \text{otherwise} \end{cases}$$

$$\delta_2(p, Rv) = \begin{cases} \bar{\delta}_2, & \text{if } p - \bar{\delta}_2 Rv \in \mathcal{E} \\ \mathrm{dist}_{-Rv}(p, \partial\mathcal{E}), & \text{otherwise} \end{cases}$$

where $\bar{\delta}_1, \bar{\delta}_2 \in \mathbb{R}_{\geq 0}$ are given constants satisfying $\bar{\delta}_1 + \bar{\delta}_2 > 0$, $\mathrm{dist}_{\pm Rv}(p, \partial\mathcal{E})$ is the distance between the point p and the set $\partial\mathcal{E}$ along the direction $\pm Rv$, and the random variable $\{\mathbf{v}_k\}_{k\in\mathbb{Z}_{\geq 0}}$ takes values in $\{-1, 1\}^n$. We assume that there is a routine that gives the distance from p to the position where the robot found the obstacle. This can be designed using information of the acceleration and the contact sensor $l_\mathcal{E}$.

To summarize, the algorithm in (7) is composed of two steps: first, the exploration step, which is given when the robot computes δ. In this step, the robot moves to two positions in the directions $+Rv$ or $-Rv$. Note here that the time-varying rotation matrix R helps the robot to explore any point in a ball centered at the robot position. For example, it is enough to set it up equal to the identity matrix or cover repeatedly a set of fixed matrices. This ball can have either a radius given by δ_1 or δ_2. Second, the gradient approximation step, which consists of sampling the intensity at each location of the exploration step, i.e., sampling at $\pm Rv$. With that information, the robot computes g and moves in its direction, where g can be seen as an approximation to the gradient, see Lemma 1. To handle the implementation of the projection operator in (7), we use (22). We first analyze (7) and then as a corollary explain how the results hold for (22). We make the following assumption on the sequence of random variables $\mathbf{v}$.

Assumption 2: (On the characteristics of the random input): The sequence of random variables $\{\mathbf{v}_k\}_{k\in\mathbb{Z}_{\geq 0}}$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbf{v}_k : \Omega \to \{-1, 1\}^n$, is i.i.d. with $E[\mathbf{v}_k] = 0$ for each $k \in \mathbb{Z}_{\geq 0}$.

Remark 1: For the easiness of presentation, we neglect the presence of noise in the observations of f. However, from the analysis in Section V, practical convergence in expected value to the tower can still be achieved under appropriate statistical properties on the noise. •

V. Convergence for the Obstacle-Free Scenario

In this section, we derive the convergence results for the algorithm in (7). In particular, we show practical convergence in probability to a ball with fixed radius depending on α and $\bar{\delta}_1 + \bar{\delta}_2$ under different assumptions. We are able to characterize the size of this ball under the assumption of strong convexity of the cost function as shown in Theorem 1. When we do not have enough information on the cost function, like differentiability, we prove practical convergence in probability to a ball that can be made arbitrarily small by tuning α and $\bar{\delta}$ as shown in Theorem 2. We begin by providing two supporting lemmas.

Lemma 1: (SPSA approximation to the gradient): Let Assumption 2, on the characteristics of the random input, hold.

Assume that f is convex, finite, and twice differentiable. Then, if $\delta_1 + \delta_2 > 0$, we have

$$g_i(p, \delta, R, v) = \frac{\partial f(p)}{\partial p_i} + b_i + c_i \quad (9)$$

where $b_i = \sum_{l,j,q,j\neq l} R_{il} R_{qj} \frac{v_j}{v_l} \frac{\partial f(p)}{\partial p_q}$, for $i \in \{1,\ldots,n\}$, $c = \frac{Rv^{-1}}{2(\delta_1+\delta_2)} v^\top R^\top (\delta_1^2 \nabla^2 f(p_1) - \delta_2^2 \nabla^2 f(p_2)) Rv$, and $p_j = p + \delta_j' Rv$ for some $\delta_j' \in [0,1]$ and $j \in \{1,2\}$. Otherwise, if $\delta_1 + \delta_2 = 0$, we have $g_i(p,\delta,R,v) = 0$ for $i \in \{1,\ldots,n\}$.

Proof: For the case when $\delta_1 + \delta_2 = 0$, by definition, it follows that $g_i(p,\delta,R,v) = 0$. Otherwise, when $\delta_1 + \delta_2 > 0$, by using a second-order Taylor expansion around p, there exists $\delta_1' \in [0,1]$ and $p_1 = p + \delta_1' Rv$ such that

$$f(p + \delta_1 Rv) = f(p) + \delta_1 v^\top R^\top \nabla_p f(p) + \frac{1}{2}\delta_1^2 v^\top R^\top \nabla^2 f(p_1) Rv. \quad (10)$$

Similarly, there is $\delta_2' \in [0,1]$ and $p_2 = p - \delta_2' Rv$ such that

$$f(p - \delta_2 Rv) = f(p) - \delta_2 v^\top R^\top \nabla_p f(p) + \frac{1}{2}\delta_2^2 v^\top R^\top \nabla^2 f(p_2) Rv. \quad (11)$$

Subtracting (11) from (10) and dividing the result by $\delta_1 + \delta_2$

$$\frac{f(p+\delta_1 Rv) - f(p-\delta_2 Rv)}{\delta_1+\delta_2} = v^\top R^\top \nabla_p f(p) + \frac{1}{2(\delta_1+\delta_2)} v^\top R^\top (\delta_1^2 \nabla^2 f(p_1) - \delta_2^2 \nabla^2 f(p_2)) Rv.$$

Multiplying the last equation by Rv^{-1} we have

$$R\frac{f(p+\delta_1 Rv) - f(p-\delta_2 Rv)}{\delta_1+\delta_2} v^{-1} = v^\top R^\top \nabla_p f(p) R v^{-1} + \frac{Rv^{-1}}{2(\delta_1+\delta_2)} v^\top R^\top (\delta_1^2 \nabla^2 f(p_1) - \delta_2^2 \nabla^2 f(p_2)) Rv. \quad (12)$$

We analyze next the ith component of the first term of the right-hand side (RHS) of the last equation as follows:

$$\left(v^\top R^\top \nabla_p f(p) R v^{-1}\right)_i = \sum_{l=1}^n \frac{R_{il}}{v_l} \sum_{j=1}^n v_j \sum_{q=1}^n R_{qj} \frac{\partial f(p)}{\partial p_q}$$
$$= \sum_{l=1}^n \frac{R_{il}}{v_l} \sum_{j\neq l}^n v_j \sum_{q=1}^n R_{qj} \frac{\partial f(p)}{\partial p_q} + \sum_{l=1}^n R_{il} \frac{v_l}{v_l} \sum_{q=1}^n R_{ql} \frac{\partial f(p)}{\partial p_q}$$
$$= \frac{\partial f(p)}{\partial p_i} + b_i \quad (13)$$

where we have used the fact that R is an orthogonal matrix. Replacing (13) in (12), (9) follows. ∎

Lemma 2 (Optimality bounds): Assume f is convex, finite, and satisfies the superquadratic growth condition in (2). Then, for all $\xi \in \partial f(p)$ and $p, p^* \in \mathbb{R}^n$, it holds that

$$(p - p^*)^\top \xi \geq \frac{\rho}{2}\|p^* - p\|^2 \quad (14)$$

and

$$\|\xi\| \geq \frac{\rho}{2}\|p^* - p\|. \quad (15)$$

Proof: We prove first inequality (14). By the assumption on the superquadratic growth condition (2) it holds that $f(p^*) \geq f(p) + (p^* - p)^\top \xi + \frac{\rho}{2}\|p^* - p\|^2$, for all $p, p^* \in \mathbb{R}^n$, and $\xi \in \partial f(p)$. Subtracting $f(p)$ from both sides, we have $f(p^*) - f(p) \geq (p^* - p)^\top \xi + \frac{\rho}{2}\|p^* - p\|^2$. By noticing that $f(p^*) - f(p) \leq 0$, we have $0 \geq (p^* - p)^\top \xi + \frac{\rho}{2}\|p^* - p\|^2$. Then, (14) follows. Next, we prove (15). By noting that the RHS of (14) is bigger or equal than zero, it follows $|(p - p^*)^\top \xi| \geq \frac{\rho}{2}\|p^* - p\|^2$. By using the Cauchy–Schwarz inequality, it follows that $\|p - p^*\|\|\xi\| \geq \frac{\rho}{2}\|p^* - p\|^2$, which implies (15). ∎

The next theorem shows the algorithm convergence when f is twice differentiable.

Theorem 1: (Convergence when f is twice differentiable): Let Assumption 2, on the characteristics of the random input, hold. Assume that f is convex, finite, twice differentiable, $\rho I_n \leq \nabla^2 f(p) \leq \Gamma I_n$, and $\|\nabla_p f(p)\| \leq M$. Then, for any initial state p_0, the solution p^* of the system (7) is MSP-ES with ultimate bound $\mathcal{O} = \mathcal{E} \backslash Z$, where

$$Z = \left\{p \in \mathcal{E} \big| \|p - p^*\|^2 \geq \frac{\alpha}{\rho}(M^2(n^2+2) + \frac{1}{4}(\bar{\delta}_1 + \bar{\delta}_2)^2 \Gamma^2 n^3)\right\}. \quad (16)$$

Proof: Without loss of generality assume $\delta_1(p, Rv) + \delta_2(p, Rv) > 0$. This is the case because, at any time $k > 0$ for which $\delta_1(p, Rv) + \delta_2(p, Rv) = 0$, with probability one, the dynamics in (7) will generate a feasible direction in finite time in $\mathcal{E}$ satisfying $\delta_1(p, Rv) + \delta_2(p, Rv) > 0$.

Without loss of generality assume $p^* \in \mathcal{E}$ (the projection of p^* on $\mathcal{E}$ is in $\mathcal{E}$ and unique.) By the nonexpansive property of the projection operation, the dynamics in (7), and the fact that $p^* \in \mathcal{E}$, we have

$$\|p^+ - p^*\|^2 = \|\Pi_{\mathcal{E}}[p - \alpha g(p, \delta(p, Rv), R, v)] - p^*\|^2$$
$$\leq \|p - \alpha g(p, \delta(p, Rv), R, v) - p^*\|^2$$
$$= \|p - \alpha(\nabla_p f(p) + b + c) - p^*\|^2$$
$$= \|p - p^*\|^2 - 2\alpha(\nabla_p f(p) + b + c)^\top (p - p^*) + \alpha^2\|\nabla_p f(p) + b + c\|^2$$

where $b_i = \sum_{l,j,q,j\neq l} R_{il} R_{qj} \frac{v_j}{v_l} \frac{\partial f(p)}{\partial p_q}$, for $i \in \{1,\ldots,n\}$, $c = \frac{Rv^{-1}}{2(\delta_1+\delta_2)} v^\top R^\top (\delta_1^2 \nabla^2 f(p^1) - \delta_2^2 \nabla^2 f(p^2)) Rv$, $p^j = p + \delta_j' Rv$ for some $\delta_j' \in [0,1]$ and $j \in \{1,2\}$ (see Lemma 1 to learn how to get b and c).

Let $V : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, $V = \|p - p^*\|^2$, and define $\Delta V = \|p^+ - p^*\|^2 - \|p - p^*\|^2$. We have

$$\Delta V \leq -2\alpha(\nabla_p f(p) + b + c)^\top (p - p^*) + \alpha^2 \|\nabla_p f(p) + b + c\|^2.$$

By using (14), we have that $-(p - p^*)^\top \nabla_p f(p) \leq -\frac{\rho}{2}\|p - p^*\|^2$. It follows that

$$\Delta V \leq -\alpha\rho\|p - p^*\|^2 - 2\alpha(b + c)^\top (p - p^*) + \alpha^2\|\nabla_p f(p) + b + c\|^2.$$

By taking expectation operator $E[V(p^+)|\mathcal{F}_k]$, since $\mathbf{v}_k$ is i.i.d. with $E[\mathbf{v}_k] = 0$ for each $k \in \mathbb{Z}_{\geq 0}$, and by noticing that $E[\mathbf{v}_k^{-1}] = E[\mathbf{v}_k]$, it implies that $E[b] = 0$. Next, we show that $E[c_i] = 0$ for $i \in \{1,\ldots,n\}$. We rewrite $c = m(v^\top H v) Rv$, where $m = \frac{1}{2(\delta_1+\delta_2)}$, $H \triangleq R^\top (\delta_1^2 \nabla^2 f(p_1) - \delta_2^2 \nabla^2 f(p_2)) R$, and $H = (h_{ij})$, and we use the fact that $v = v^{-1}$. Then

$$E[c_i] = mE\left[\sum_{l=1}^n R_{il} v_l \sum_{s=1}^n v_s \sum_{j=1}^n h_{sj} v_j\right]$$
$$= m(q_i + z_i)$$

where $q_i = E[R_{ii} v_i \sum_{s=1}^{n} v_s \sum_{j=1}^{n} h_{sj} v_j]$ and $z_i = E[\sum_{l \neq i} R_{il} v_l \sum_{s=1}^{n} v_s \sum_{j=1}^{n} h_{sj} v_j]$. Expanding q_i

$$
\begin{aligned}
q_i &= E\left[R_{ii} v_i \left(v_i \sum_{j=1}^{n} h_{ij} v_j + \sum_{s \neq i} v_s \sum_{j=1}^{n} h_{sj} v_j\right)\right] \\
&= R_{ii} E\left[h_{ii} v_i^3 + v_i^2 \sum_{j \neq i} h_{ij} v_j + v_i \sum_{s \neq i} v_s^2 h_{ss}\right. \\
&\quad \left. + v_i \sum_{s \neq i} v_s \sum_{j \neq s} h_{sj} v_j \right] \\
&= 0
\end{aligned}
$$

where we have used the assumption that $\mathbf{v}_k$ is i.i.d. with $E[\mathbf{v}_k] = 0$ and the fact that $v_i^3 = v_i$ for $i \in \{1, \ldots, n\}$. Analogous to the last procedure, we expand z_i

$$
\begin{aligned}
z_i &= E\left[\sum_{l \neq i} R_{il} v_l^2 \sum_{j=1}^{n} h_{lj} v_j + \sum_{l \neq i} R_{il} v_l \sum_{s \neq l} v_s \sum_{j=1}^{n} h_{sj} v_j\right] \\
&= E\left[\sum_{l \neq i} R_{il} v_l^3 h_{ll} + \sum_{l \neq i} R_{il} v_l^2 \sum_{j \neq l} h_{lj} v_j + \sum_{l \neq i} R_{il} v_l \sum_{s \neq l} v_s^2 h_{ss}\right. \\
&\quad \left. + \sum_{l \neq i} R_{il} v_l \sum_{s \neq l} v_s \sum_{j \neq s} h_{sj} v_j \right] \\
&= 0.
\end{aligned}
$$

Thus, $E[c] = 0$. Therefore

$$
\begin{aligned}
E[\Delta V | \mathcal{F}_k] &\leq -\alpha\rho \|p - p^*\|^2 + \alpha^2 (\|\nabla_p f(p)\|^2 \\
&\quad + E[\|b\|^2 + \|c\|^2 | \mathcal{F}_k]).
\end{aligned} \tag{17}
$$

Note that $E[\|c\|^2] \leq \frac{1}{4} \Gamma^2 n^3 (\delta_1 + \delta_2)^2$ and from (13), we have

$$
\begin{aligned}
E[\|b\|^2 | \mathcal{F}_k] &= E[\|v^\top R^\top \nabla_p f(p) R v^{-1} - \nabla_p f(p)\|^2 | \mathcal{F}_k] \\
&\leq E[\|v^\top R^\top\|^2 \|\nabla_p f(p)\|^2 \|R v^{-1}\|^2 + \|\nabla_p f(p)\|^2 | \mathcal{F}_k] \\
&\leq M^2 (n^2 + 1)
\end{aligned}
$$

where we have used $\|\nabla_p f\| \leq M$. Using above upper bounds and replacing them in (17), one has

$$
E[\Delta V | \mathcal{F}_k] \leq -\alpha\rho V(p) + \frac{\alpha^2}{4} (\delta_1 + \delta_2)^2 \Gamma^2 n^3 + \alpha^2 M^2 (n^2 + 2).
$$

It follows that $E[\Delta V | \mathcal{F}_k] \leq -\alpha\rho V(p) + \alpha^2 J$, where $J = \frac{1}{4}(\delta_1 + \delta_2)^2 \Gamma^2 n^3 + M^2 (n^2 + 2)$. Reorganizing these terms, we have $E[V(p^+) | \mathcal{F}_k] \leq (1 - \alpha\rho) V(p) + \alpha^2 J$. Therefore, by Proposition 1, the equilibrium point is MSE-ES. Notice that the max inside the integral in (4) simplifies to a point because we do not have a differential inclusion. The set $\mathcal{O}$ given in (16) follows by noticing that $E[\Delta V | \mathcal{F}_k] \leq 0$ if $\|p - p^*\|^2 \geq \mathcal{O}$ and by noticing that $\delta_1 + \delta_2 \leq \bar{\delta}_1 + \bar{\delta}_2$. ■

Remark 2: In Theorem 1, the knowledge of Γ and M are not required to prove convergence. Given that $\mathcal{E}$ is assumed to be compact, the existence of Γ is guaranteed. Since f is assumed locally Lipschitz, then there always exist a finite M such that $\|\nabla_p f(p)\| \leq M$. We use those values to characterize the size of the ball where the trajectories converge to in expectation. •

If f is nondifferentiable, we are not able to characterize the size of the ball as in Theorem 1. However the next result shows practical convergence in probability to p^* and that this ball can be made arbitrarily small by reducing α and $\bar{\delta}_1 + \bar{\delta}_2$ without the assumption on the superquadratic growth condition on f.

Theorem 2: (Convergence when f is nonsmooth): Let Assumption 2, on the characteristics of the random input, hold. Assume that f is convex and finite with a unique minimizer p^*. Then, for any initial state p_0, the solution p^* of the system (7) is MSP-ES.

Proof: The proof employs the same Lyapunov function as that of Theorem 1, and the nonexpansive property of the projection operator to bound its difference. However, in order to find upper bounds, we can not resort to the differentiability properties of f and we do not use the assumption on its superquadratic growth condition. Since f is assumed to be convex and locally Lipschitz, then the set-valued map ∂f is locally bounded, upper semicontinuous, and takes nonempty, compact, and convex values [14]. Using the last fact, the sup in (1) can be replaced by a max, and then

$$
f(p + \delta_1 R v) = f(p) + \delta_1 v^\top R^\top \bar{\xi} + o(\delta_1 \|Rv\|) \tag{18}
$$

where $\bar{\xi} = \operatorname{argmax}_{\xi \in \partial f_s(p)} \{\xi^\top R v\}$. Similarly

$$
f(p - \delta_2 R v) = f(p) - \delta_2 v^\top R^\top \underline{\xi} + o(\delta_2 \|Rv\|) \tag{19}
$$

where $\underline{\xi} = \operatorname{argmin}_{\xi \in \partial f(p)} \{\xi^\top R v\}$. Subtracting (19) from (18) and dividing the result by $\delta_1 + \delta_2$, we have

$$
\begin{aligned}
\frac{f(p + \delta_1 R v) - f(p - \delta_2 R v)}{\delta_1 + \delta_2} &= \frac{1}{\delta_1 + \delta_2} \Big(v^\top R^\top (\delta_1 \bar{\xi} + \delta_2 \underline{\xi}) \\
&\quad + o(\delta_1 \|v\|) - o(\delta_2 \|v\|) \Big)
\end{aligned}
$$

where we have used the assumption that R is an orthogonal matrix, then $o(\delta_i \|Rv\|) = o(\delta_i \|v\|)$ for $i \in \{1, 2\}$. Multiplying the last equation by $R v^{-1}$, we have

$$
\begin{aligned}
\frac{f(p + \delta_1 R v) - f(p - \delta_2 R v)}{\delta_1 + \delta_2} R v^{-1} &= v^\top R^\top (\delta_1 \bar{\xi} + \delta_2 \underline{\xi}) \frac{R v^{-1}}{\delta_1 + \delta_2} \\
&\quad + \frac{R v^{-1}}{\delta_1 + \delta_2} (o(\delta_1 \|v\|) - o(\delta_2 \|v\|)).
\end{aligned} \tag{20}
$$

We analyze the ith component of the first term of the RHS of the last equation to obtain

$$
\begin{aligned}
\left(v^\top R^\top (\delta_1 \bar{\xi} + \delta_2 \underline{\xi}) \frac{R v^{-1}}{\delta_1 + \delta_2} \right)_i &= \sum_{l=1}^{n} \frac{R_{il}}{v_l} \sum_{j=1}^{n} v_j \sum_{q=1}^{n} R_{qj} \frac{\delta_1 \bar{\xi}_q + \delta_2 \underline{\xi}_q}{\delta_1 + \delta_2} \\
&= \frac{\delta_1 \bar{\xi}_i + \delta_2 \underline{\xi}_i}{\delta_1 + \delta_2} + b_i
\end{aligned} \tag{21}
$$

where $b_i = \sum_{l,j,q, j \neq l} R_{il} R_{qj} \frac{v_j}{v_l} \frac{\delta_1 \bar{\xi}_q + \delta_2 \underline{\xi}_q}{\delta_1 + \delta_2}$. Replacing (21) in (20), it follows that

$$
\frac{f(p + \delta_1 R v) - f(p - \delta_2 R v)}{\delta_1 + \delta_2} R v^{-1} = \frac{\delta_1 \bar{\xi} + \delta_2 \underline{\xi}}{\delta_1 + \delta_2} + b + c
$$

where $c = \frac{R v^{-1}}{\delta_1 + \delta_2} (o(\delta_1 \|v\|) - o(\delta_2 \|v\|))$. Then, one has

$$
g(p, \delta(p, Rv), R, v) = \frac{\delta_1 \bar{\xi} + \delta_2 \underline{\xi}}{\delta_1 + \delta_2} + b + c.
$$

With probability one, we can assume that the dynamics in (7) generates a feasible direction in finite time such that $\delta_1(p, Rv) + \delta_2(p, Rv) > 0$. We can also assume that $p^* \in \mathcal{E}$.

In addition, resorting to the nonexpansive property of the projection operation

$$\begin{aligned}\|p^+ - p^*\|^2 &= \|\Pi_{\mathcal{E}}[p - \alpha g(p, \delta(p, Rv), R, v)] - p^*\|^2\\ &\le \|p - \alpha g(p, \delta(p, Rv), R, v) - p^*\|^2\\ &\le \|p - \alpha\Big(\frac{\delta_1\bar{\xi} + \delta_2\underline{\xi}}{\delta_1 + \delta_2} + b + c\Big) - p^*\|^2.\end{aligned}$$

It follows

$$\|p^+ - p^*\|^2 \le \|p - p^*\|^2 + \alpha^2\|\frac{\delta_1\bar{\xi} + \delta_2\underline{\xi}}{\delta_1 + \delta_2} + b + c\|^2 - 2\alpha\left(\frac{\delta_1\bar{\xi} + \delta_2\underline{\xi}}{\delta_1 + \delta_2} + b + c\right)^\top (p - p^*).$$

Let $V : \mathbb{R}^n \to \mathbb{R}_{\ge 0}, V = \|p - p^*\|^2$, and define $\Delta V = \|p^+ - p^*\|^2 - \|p - p^*\|^2$. Then, we have

$$\Delta V \le -2\alpha\left(\frac{\delta_1\bar{\xi} + \delta_2\underline{\xi}}{\delta_1 + \delta_2} + b + c\right)^\top (p - p^*) + \alpha^2\|\frac{\delta_1\bar{\xi} + \delta_2\underline{\xi}}{\delta_1 + \delta_2} + b + c\|^2.$$

Let $f_s : \mathbb{R}^n \to \mathbb{R}$ be a convex function satisfying the superquadratic growth condition for some $\rho \in \mathbb{R}_{>0}$ such that $f_s(p^*) = f(p^*), \xi_s^\top(p - p^*) \le \xi^\top(p - p^*)$, $\xi \in \partial f(p)$, and $\xi_s \in \partial f_s(p)$ for all $p \in \mathcal{E}$. Notice that f_s always can be found since p^* is assumed unique and $\mathcal{E}$ is a compact set. Using the last fact, there exists $\rho > 0$ such that $-\xi^\top(p - p^*) \le -\frac{\rho}{2}\|p - p^*\|^2$. Thus

$$\Delta V \le -\alpha\rho\frac{\delta_1 + \delta_2}{\delta_1 + \delta_2}\|p - p^*\|^2 - 2\alpha(b + c)^\top(p - p^*) + \alpha^2\left\|\frac{\delta_1\bar{\xi} + \delta_2\underline{\xi}}{\delta_1 + \delta_2} + b + c\right\|^2.$$

By noticing that $E[b] = E[c] = 0$, one has

$$E[\Delta V|\mathcal{F}_k] \le -\alpha\rho\|p - p^*\|^2 + \alpha^2 E[\|\frac{\delta_1\bar{\xi} + \delta_2\underline{\xi}}{\delta_1 + \delta_2} + b + c\|^2|\mathcal{F}_k].$$

From here, the proof follows similar steps as the proof of Theorem 1, where we use ξ instead of $\nabla_p f$, and consider $O(1)$ terms instead of the upper bound of the Hessian. We omit the steps for conciseness. ■

When the robot moves in a GPS-denied environment, the implementation of the projection operator $\Pi_{\mathcal{E}}$ in (7) is challenging. However, if $p^* \in \mathcal{E}$ and we use

$$p^+ = p - \alpha(p, \delta(p, Rv), R, v)g(p, \delta(p, Rv), R, v) \tag{22}$$

where $\alpha : \mathbb{R}^n \times \mathbb{R}^2 \times \mathrm{SO}(n) \times \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\alpha(p, \delta(p, Rv), R, v) = \begin{cases}\bar{\alpha}, & \text{if } p - \bar{\alpha}g(p, \delta(p, Rv), R, v) \in \mathcal{E}\\ \mathrm{dist}_{+g}(p, \partial\mathcal{E}), & \text{otherwise}\end{cases}$$

and all other variables defined as in (8) for given $\bar{\alpha} \in \mathbb{R}_{>0}$. Then, the results of Theorem 2 hold as is shown in the following corollary.

Corollary 1: Let Assumption 2, on the characteristics of the random input, hold. Assume that $f : \mathbb{R}^n \to \mathbb{R}$ is convex and finite with a unique minimizer $p^* \in \mathcal{E}$. Then, for any initial state $p_0 \in \mathcal{E}$, the solution p^* of the system (22) is MSP-ES.

Proof: The proof follows by noticing that, whenever the robot detects a boundary, it stops until the next iteration of the dynamics in (22). Then, three considerations have to be taken into account as in Theorem 2. First, notice that we can assume $\delta_1(p, Rv) + \delta_2(p, Rv) > 0$ similarly as before. Second, the robot stops at any time $k > 0$ while executing (22) only when it finds a boundary. Then, the positive term in the RHS of the upper bound of $E[\Delta V|\mathcal{F}_k]$ given in the proof of Theorem 2 parameterized by α is less or equal, with probability one, to the same positive term parameterized by $\bar{\alpha}$. The last fact is explained by noticing that $\alpha \le \bar{\alpha}$. Third, when the robot is in $\partial\mathcal{E}$, it moves to a feasible direction in a finite time, with probability one. ■

VI. Proposed Algorithm and Convergence Analysis for Environments With Obstacles

In this section, we propose an algorithm to find the tower when there are obstacles in the environment and the intensity function f is convex. We introduce some notation used in this section. Let $\mathcal{O}$ be a disjoint set of obstacles. Each $O \in \mathcal{O}$ is assumed to be closed with a connected piecewise-analytic boundary that is finite in length. Furthermore, the obstacles in $\mathcal{O}$ are pairwise disjoint. There may be a countably infinite number of obstacles, but at most a finite number are contained in any fixed disk. The obstacle set $\mathcal{O}$ may contain an outer obstacle O_{outer} that is unbounded; all other obstacles are bounded. Here, we consider a point robot that moves in $\mathbb{R}^2$ according to a kinematic differential drive model. We use the following motion primitives that are used in Algorithm 1.

U_{fwd}: The robot executes the dynamics in (22), stopping only if it contacts an obstacle.

U_{fol}: The robot travels around the obstacle boundary clockwise, maintaining contact to the right at all times, with a fixed step size $\bar{\alpha}$, and executing $g(p, \delta, Rv)$ at every time step. The robot stops implementing U_{fol} only when $g(p, \delta, Rv)$ points to a feasible direction (i.e., a direction that gives a point in $\mathcal{E} \setminus \mathcal{O}$).

Remark 3: The proposed motion primitives are inspired by those in [8]. However, notice that U_{fol} is different from the U_{fol} proposed in [8] as the latter is defined to stop at a local minimum. Our motion primitive is defined to stop when the estimated gradient points toward a feasible direction. •

Due to the fixed step size, it could be that the robot does not detect a point to leave an obstacle when traveling around it. If this happens, the robot will be trapped in the obstacle for all time. This case can be prevented under the following conditions. Assume that we know the Lipschitz constant L, i.e., $|f(p) - f(p')| \le L\|p - p'\|$, then we consider a constant B such that $B \triangleq L$. We define $C \in \mathbb{R}_{>0}$ to indicate an upper bound of $E[\|b + c\|]$ for all $p \in \mathcal{E} \setminus \mathcal{O}$.

Assumption 3: (On the step size given B): We assume that $\bar{\alpha}$ is small enough such that $\bar{\alpha}\max\{B, C\}$ is strictly less than the minimum distance between two separated obstacles. In addition, we assume that the robot is able to circumnavigate any obstacle $O \in \mathcal{O}$ using the constant step size $\bar{\alpha}$. Finally, we assume $\bar{\delta}_1$ and $\bar{\delta}_2$ less or equal than $\bar{\alpha}$.

The following lemma is used in the main result of this section, given in Theorem 3.

Lemma 3 (see [8]): For every obstacle boundary ∂O and every possible tower location $\mathbb{R}^2 \setminus O$, there exists at least one intensity local minimum $p \in \partial O$ for which the topological disk centered at the tower $(0, 0)$ with radius $\|p\|$ is disjoint from the interior of O.

Theorem 3: (Convergence with obstacles in the environment): Let Assumption 2, on the characteristics of the random input, and Assumption 3, on the step-size given B and C, hold. Assume that f is convex and finite with a unique minimizer p^*. Then, for any initial

Algorithm 1: Algorithm for convex intensity functions.

1: $I_L = f(p)$
2: **while** not hitting an obstacle **do**
3: execute U_{fwd}
4: **end while**
5: $I_H = f(p)$
6: **while** $f(p) - \bar{\alpha}B \geq I_H$ **do**
7: execute U_{fol}
8: **end while**
9: go to Step 1

$p_0 \in \mathcal{E} \setminus \mathcal{O}$, Algorithm 1 causes the robot to reach a ball of radius $\bar{\alpha}$ containing the tower with probability one.

Proof: Since f is assumed to be convex, when the robot moves using the estimated gradient in Step 3, the average distance to the tower is decreased, as shown in Theorem 2. This implies that, on average, f is nonincreasing as the robot moves. In fact, f is decreasing because on average, the gradient points to the greatest decrease rate of the function f.

After the execution of Steps 2–4 for the first time, either the ball of size depending on $\bar{\alpha}$ that contains the tower is located or the robot hits the boundary of an obstacle. If the former happens, then the algorithm terminates successfully. Thus, let us assume the latter. By Step 5, I_H stores the intensity at the boundary point where the robot hits the obstacle. At this point, the robot follows around the obstacle.

It might seem that an infinite loop is possible by failure to satisfy the condition of Step 6 or by the motion being blocked by an obstacle boundary. However, this is not the case because of the following three facts: first, let denote by P_m the set of points $p \in \partial O$ where there is a local minimum. By Lemma 3, P_m is not empty and second, there exists $p_m \in P_m$ for which the gradient $\nabla_p f(p_m)$ (or an element of the generalized gradient, when f is nondifferentiable) points to a feasible direction (i.e., a direction that gives a point in $\mathcal{E} \setminus \mathcal{O}$). Then, in expectation $g(p_m, \delta, Rv)$ points to a feasible direction for such points. In addition, third, by the definition of B and because the robot follows the obstacle boundary while the condition in Step 6 holds, the robot keeps detecting all balls of radius $\bar{\alpha}B$ containing all the local minima of ∂O. In particular, the robot will reach a ball of radius $\bar{\alpha}B$ containing a minimum point p_m as in Lemma 3. Because p_m is a local minimum it must be that $f(p_m) \leq I_H$. Now, the robot reaches p' such that $\|p_m - p'\| \leq \bar{\alpha}$. By the Lipschitz condition on f, it holds that $f(p') < f(p_m) + \bar{\alpha}B$, which implies that $f(p') < I_H + \bar{\alpha}B$. Thus, the while loop will be exited at this point p'. Then, I_L is reassigned to $I_L = f(p')$ by the re-execution of Step 3, and the robot leaves the obstacle by reapplying Steps 2–4.

When the robot leaves the obstacle through this point, it is guaranteed that it will not contact a different obstacle in the next iteration since it is assumed that all obstacles are separated by a distance bigger than $\bar{\alpha}\max\{B, C\}$. That is, a deadlock by bouncing between obstacles is not possible. By Step 6, the leaving point from the obstacle boundary ∂O is closer to the goal than the hitting point. Thus, even if an obstacle boundary ∂O is contacted a finite number of times due to nonconvexities, the robot will eventually leave the obstacle behind after a finite number of hits. In an environment with a finite number of obstacles, the robot will reach a ball that contains the minimizer of the intensity function. ■

Remark 4: There are results available in the literature on tight upper bounds on the length of bug paths when the gradient information is known, which characterizes the running time to be of the order of this length. For the unknown gradient case, an upper bound on the expected length of bug paths should be very similar. Indeed, this seems to be the case in view of the simulation results that we have obtained; see Fig. 2. In the known gradient case, a tight upper bound on the length of bug paths for e.g., the i-bug algorithm in [8] (which is similar to ours) is given by $L + \sum_{O \in \mathcal{O}} n_O c_O$, where L is the length of the straight line joint from the initial position to the maximizer, n_O is the number of unblocked local maxima along O (a local maximum at a point $p \in \partial O$ is called unblocked if the robot can freely move toward the tower from p), and c_O is its perimeter. The feasible direction search in our algorithm, affected by v and R_k and the obstacle, perturbs the bug path but the resulting behavior is on average similar to the known gradient case. Then, we conjecture that an upper bound on the expected length of the path multiplied by a factor upper bounding the time it takes the robot to find a feasible direction should be a good indicative of the order of the running time of the algorithm. However, this is just a conjecture, and a thorough analysis falls out of the scope of this paper. •

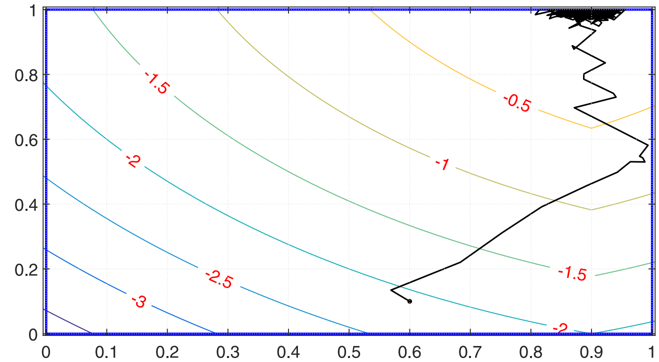

Fig. 1. Evolution of the mobile robot for $f = (p_1 - 0.9)^2 + |p_1 - 0.9| + (p_2 - 1)^2 + |p_2 - 1|$ with a box constraint $p \in [0, 1]^2$. The level sets of f are shown in colors and the trajectory of the robot is shown in black.

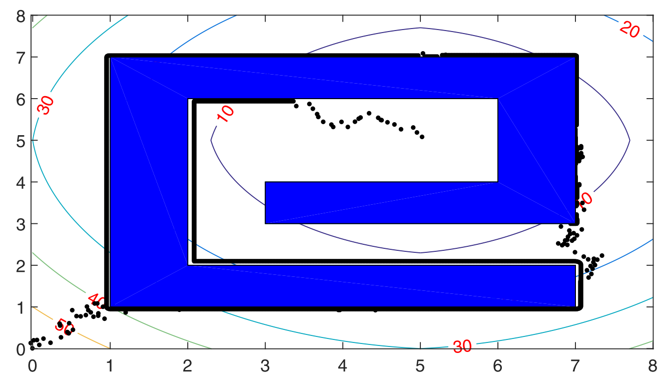

Fig. 2. Evolution of the robot for $f = (p_1 - 5)^2 + |p_1 - 5| + (p_2 - 5)^2 + |p_2 - 5|$ with a box constraint $p \in [0, 8]^2$. The level sets of f are shown in colors. The trajectory of the robot is shown in red and the obstacles are shown in blue.

VII. Simulations

Here, we show the response of (7) and Algorithm 1 to a particular source-seeking problem, in Figs. 1 and 2, respectively. In both simulations, we use a point robot that moves in $\mathbb{R}^2$ according to a kinematic differential drive model. Fig. 1 illustrates the evolution of the mobile robot to a source $f = (p_1 - 0.9)^2 + |p_1 - 0.9| + (p_2 - 1)^2 + |p_2 - 1|$ with a box constraint $p \in [0, 1]^2$. Notice that the function f is nondifferentiable and strongly convex, then it satisfies the conditions on Theorem 2. The tower is located at $p^* = [0.9, 1]^\top$. This simulation uses $\alpha = \bar{\delta}_1 = \bar{\delta}_2 = 0.02$ and $R_k = I_n$ for all $k \geq 0$, and $v_k \in \{-1, 1\}^2$ takes values drawn from the Bernoulli distribution, for which each outcome has equal probability. We have introduced additional white noise in both the measurements of the intensity signal and in the state, with standard deviations of 0.001 and 0.01, respectively. The robot starts at $p_0 = [.6, .1]^\top$ and it converges to a ball containing the optimizer p^*, which in turn can be made arbitrarily small by decreasing the parameters α, $\bar{\delta}_1$, and $\bar{\delta}_2$.

Fig. 2 illustrates the evolution of the mobile robot to a source $f = (p_1 - 5)^2 + |p_1 - 5| + (p_2 - 5)^2 + |p_2 - 5|$ with a box constraint $p \in [0, 8]^2$. Notice that the function f is nondifferentiable and strongly convex, then it satisfies the conditions on Theorem 2. The tower is located at $p^* = [5, 5]^\top$. This simulation uses $\alpha = 0.1$, $\bar{\delta}_1 = \bar{\delta}_2 = 0.01$, and $R_k = I_n$ for all $k \geq 0$, and $v_k \in \{-1, 1\}^2$ takes values drawn from the Bernoulli distribution, for which each outcome has equal probability. The robot starts at $p_0 = [0, 0]^\top$. Additive white noises in both measurements, in the intensity and the state, are introduced. The standard deviations are 0.01 and 0.05, respectively. The robot converges to a ball containing the optimizer p^*, which in turn can be made arbitrarily small by decreasing the parameters α, $\bar{\delta}_1$, and $\bar{\delta}_2$.

VIII. Conclusion

Building on the SPSA method, we have introduced a novel algorithm that allows a mobile robot to find the minimizer of an emitting signal. The novelty of our approach is that we consider nondifferentiable convex functions and fixed step size for compact convex environments. Inspired by the well-known family of bug algorithms, we have extended our approach to include obstacles in the environment. In particular, we prove convergence to a ball around the optimizer of the emitting signal whose size depends on the step size. The proof relies on the Lyapunov theory together with tools from convex analysis and stochastic difference inclusions. Finally, we show the applicability of the proposed algorithm in a 2-D scenario for the source-seeking problem in scenarios where observation noise is present. Overall, our proposed modification captures the mean behavior of the bug algorithms, and its applicability seems to work in scenarios where the conditions for convergence do not hold. A thorough analysis of the complexity of the algorithm is left for future work.

References

[1] J. C. Spall, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," *IEEE Trans. Autom. Control*, vol. 37, no. 3, pp. 332–341, Mar. 1992.

[2] S. Liu and M. Krstic, "Stochastic source seeking for nonholonomic unicycle," *Automatica*, vol. 46, no. 9, pp. 1443–1453, 2010.

[3] S. Azuma, M. S. Sakar, and G. J. Pappas, "Stochastic source seeking by mobile robots," *IEEE Trans. Autom. Control*, vol. 57, no. 9, pp. 2308–2321, Sep. 2012.

[4] N. Atanasov, J. L. Ny, N. Michael, and G. J. Pappas, "Stochastic source seeking in complex environments,"in *Proc. IEEE Int. Conf. Robot. Automat.*, Saint Paul, MN, USA, May 2012, pp. 3013–3018.

[5] D. Popovic, "Topics in extremum seeking," Ph.D. dissertation, Dept. elect. comput., Univ. California at Santa Barbara, Santa Barbara, CA, USA, 2004.

[6] O. Granichin and N. Amelina, "Simultaneous perturbation stochastic approximation for tracking under unknown but bounded disturbances," *IEEE Trans. Autom. Control*, vol. 60, no. 6, pp. 1653–1658, Jun. 2015.

[7] Y. He, M. C. Fu, and S. I. Marcus, "Convergence of simultaneous perturbation stochastic approximation for nondifferentiable optimization," *IEEE Trans. Autom. Control*, vol. 48, no. 8, pp. 1459–1463, Aug. 2003.

[8] K. Taylor and S. M. LaValle, "Intensity-based navigation with global guarantees," *Auton. Robots*, vol. 36, no. 4, pp. 349–364, 2014.

[9] H. Choset, K. M. Lynch, S. Hutchinson, G. Kantor, L. E. Kavraki, and S. Thrun, *Principles of Robot Motion: Theory, Algorithms Implementations*. Cambridge, MA, USA: MIT Press, 2005.

[10] R. T. Rockafellar and R. J.-B Wets, *Variational Analysis*. New York, NY, USA: Springer, 1998.

[11] R. T. Rockafellar, "Second-order convex analysis," *J. Nonlinear Convex Anal.*, vol. 1, pp. 1–16, 1999.

[12] A. R. Teel, J. P. Hespanha, and A. Subbaraman, "Equivalent characterizations of input-to-state stability for stochastic discrete-time systems," *IEEE Trans. Autom. Control*, vol. 59, no. 2, pp. 516–522, Feb. 2014.

[13] J. I. Poveda, A. R. Teel, and D. Nesic, "Flexible Nash seeking using stochastic difference inclusions, "in *Proc. Amer. Control Conf.*, Chicago, IL, USA, Jul. 2015, pp. 2236–2241.

[14] J. Cortés, "Discontinuous dynamical systems—A tutorial on solutions, nonsmooth analysis, and stability," *IEEE Control Syst. Mag.*, vol. 28, no. 3, pp. 36–73, May 2008.