

# Nonholonomic Source Seeking in Switching Random Fields

Shun-ichi Azuma, Mahmut Selman Sakar, and George J. Pappas

**Abstract**—We consider the problem of designing controllers for nonholonomic mobile robots converging to the source (minimum) of a field. In addition to the mobility constraints posed by the nonholonomic dynamics, we assume that the field is completely unknown to the robot and the robot has no knowledge of its own position. Furthermore, the field is randomly switching. In this paper, we combine ideas from stochastic approximations and nonholonomic control, in order to address this challenging problem. In particular, we develop a rotation-invariant and forward-sided version of the simultaneous-perturbation stochastic algorithm, which is much more suitable for sensor-free navigation. Based on this algorithm, we design source seeking controllers for nonholonomic robots and prove convergence to the unknown source with probability 1. The proposed controllers are demonstrated by numerical simulations.

## I. INTRODUCTION

Source seeking is a mixed problem of *search* and *navigation* as shown in Fig. 1: when a mobile robot is placed in an environment where an *unknown* signal field is introduced, find a controller steering the robot to the source (the unknown maximizer) *without* the position information. The field could express the spacial distribution of magnetic force, heat, or chemical concentration. The robot is navigated by only using the measurements of the signal at the positions.

Unlike the existing results, e.g., [1]–[5], we are interested in the *stochastic* source seeking, which involves a randomly switching field, and we want to solve it for a *general class* of mobile robots. This is motivated by two facts. First, though switching fields appear in many applications, they have never been handled so far. An example with a switching field is the base station placement for wireless communication, which is to find the best location in terms of the terminal density. In this case, the signal field corresponds to the radio field made by a number of terminal units, which randomly switches depending on their usage. Next, to our best knowledge, there is no solution dealing with various types of robots in a unified way. In fact, the existing results have focused on specific robots, e.g., integrators in [3], nonholonomic unicycles in [2], [4], [5], and underwater vehicles [1].

This paper addresses the stochastic source seeking problem for mobile robots in the general form  $\dot{x} = G(x)u$ . Our approach is to generate a stochastic trajectory converging to the unknown source and to steer the robot along the trajectory. Here, the idea of a stochastic optimization

This work was supported by the ONR-MURI award N00014-08-1-0696 and The Kyoto University Foundation.

S. Azuma is with Graduate School of Informatics, Kyoto University, Uji, Kyoto 611-0011, Japan [sazuma@i.kyoto-u.ac.jp](mailto:sazuma@i.kyoto-u.ac.jp). M.S. Sakar and G.J. Pappas are with Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA [{sakarmah, pappasj}@seas.upenn.edu](mailto)

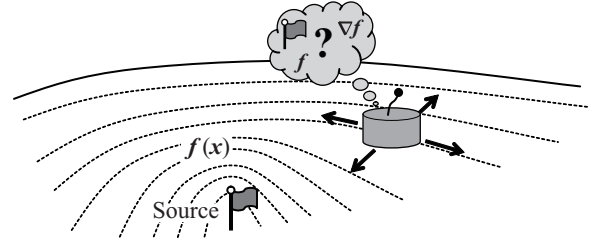


Fig. 1. Source seeking problem.

technique, called the *simultaneous-perturbation stochastic approximation (SPSA)* [6], is utilized to obtain the trajectory.

The contributions of this paper are as follows. First, we extend the SPSA algorithm to an appropriate form for the source seeking by mobile robots. Since the original algorithm generates a trajectory based on the world coordinate frame, it is impossible to drive the robot along the trajectory without a position sensor. In contrast, the new algorithm provides a trajectory in a time-varying coordinate frame, which fits the sensor-free navigation. Based on this, source seeking controllers, which drive the robot to the source with probability 1, can be characterized by the combination of point-to-point controllers. Second, simple source seeking controllers are presented for the unicycle by exploiting a special structure. This shows that the stochastic source seeking is achieved by very simple movements.

**Notation:** Let  $\mathbf{R}$ ,  $\mathbf{R}_+$ , and  $\mathbf{N}$  be the real number field, the set of positive real numbers, the set of nonnegative integers, respectively. We denote by  $0$  and  $I$  the zero matrix and the identity matrix of appropriate dimensions. For  $\theta \in \mathbf{R}$ ,  $R_2(\theta)$  is the two-dimensional rotation matrix. We use  $\|x\|$  to express the Euclidian norm of the vector  $x$ . Finally, for the scalar-valued function  $f : \mathbf{R}^n \rightarrow \mathbf{R}$ , the gradient is denoted by  $\nabla f(x)$ , i.e.,  $\nabla f(x) := [\partial f(x)/\partial x_1 \ \cdots \ \partial f(x)/\partial x_n]^\top \in \mathbf{R}^n$  where  $x_i$  is the  $i$ th element of the vector  $x \in \mathbf{R}^n$ .

## II. STOCHASTIC SOURCE SEEKING PROBLEM

### A. Problem Formulation

Consider the feedback system in Fig. 2.

The robot  $P$  is given by

$$P : \begin{bmatrix} \dot{x}(t) \\ \dot{\theta}(t) \\ \dot{\phi}(t) \end{bmatrix} = G(x(t), \theta(t), \phi(t)) u(t) \quad (1)$$

where  $x(t) \in \mathbf{R}^{n_1}$  and  $\theta(t) \in \mathbf{R}^{n_2}$  are the translational and orientational positions in the world coordinate frame,  $\phi(t) \in \mathbf{R}^{n_3}$  is the internal posture relative to the absolute position  $(x(t), \theta(t))$ ,  $u(t) \in \mathbf{R}^m$  is the control input, and  $G : \mathbf{R}^{n_1} \times \mathbf{R}^{n_2} \times \mathbf{R}^{n_3} \rightarrow \mathbf{R}^{(n_1+n_2+n_3) \times m}$  is a nonlinear function describing the dynamics. We assume that  $P$  is in a

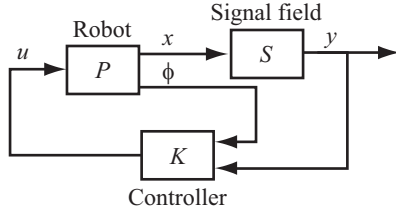


Fig. 2. Control system for source seeking.

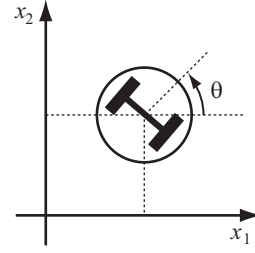


Fig. 3. Unicycle.

two- or three-dimensional space, i.e.,  $n_1 \in \{2, 3\}$  and  $n_2 \in \{1, 2, 3\}$ . An example of  $P$  is the nonholonomic unicycle in Fig. 3, which is described by

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{\theta}(t) \end{bmatrix} = \begin{bmatrix} \cos \theta(t) & 0 \\ \sin \theta(t) & 0 \\ 0 & 1 \end{bmatrix} u(t) \quad (2)$$

where  $x_1(t), x_2(t), \theta(t) \in \mathbf{R}$ . Note that the state variable  $\phi$  is not required for the unicycle but will be used for more complicated robots such as a four-wheeled vehicle.

The signal field  $S$  is a stochastic transducer from the information on  $x$  to a scalar signal, which is given by

$$S: y(t) = f_{\sigma(t)}(x(t)) \quad (3)$$

where  $y(t) \in \mathbf{R}$  expresses the signal strength (but assume that a smaller value expresses a stronger signal level) and  $f_{\sigma} : \mathbf{R}^{n_1} \rightarrow \mathbf{R}$  are thrice differentiable strictly convex functions. Furthermore,  $\sigma(t) \in \{1, 2, \dots, N\}$  is the piecewise constant random signal, given as  $\sigma(t) = \sigma_i$  on the time interval  $[i\eta, (i+1)\eta)$  where  $i \in \mathbf{N}$ ,  $\eta \in \mathbf{R}_+$ , and  $\sigma_i$  is the i.i.d. random variable from a probability distribution  $q : \{1, 2, \dots, N\} \rightarrow [0, 1]$ . An example of  $S$  is shown in Section IV-C. We denote by  $E[f_{\sigma}(x)|x]$  the conditional expected value for  $\sigma$ , and we call the translational position  $\operatorname{argmin}_{x \in \mathbf{R}^{n_1}} E[f_{\sigma}(x)|x]$  the *source*. The minimum of  $E[f_{\sigma}(x)|x]$  expresses the strongest signal level in the expectation sense. Due to the convexity of  $f_{\sigma}$ ,  $E[f_{\sigma}(x)|x]$  is a strictly convex function of  $x$ .

The controller  $K$  is a (causal) dynamical system, whose inputs are  $\phi$  and  $y$ . So, in  $K$ , the information on the internal posture  $\phi$  is available by internal sensors of  $P$  (e.g., potentiometers), but that on the absolute position  $(x, \theta)$  is not.

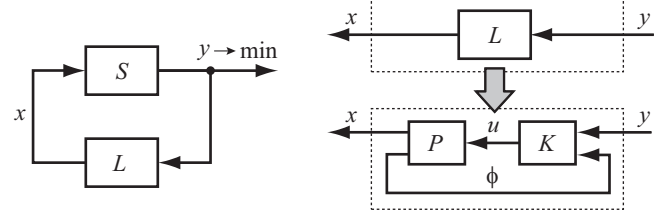
Then we consider the following problem.

**Problem 1 (Stochastic Source Seeking):** For the feedback system in Fig. 2, suppose that  $P$  and  $S$  are given, but assume that  $S$  (i.e.,  $f_{\sigma}$  and  $q$ ) is *unknown*. Then find a controller  $K$  seeking the *source*, that is, satisfying

$$\lim_{t \rightarrow \infty} y(t) = \min_{x \in \mathbf{R}^{n_1}} E[f_{\sigma}(x)|x] \quad \text{w.p.1.} \quad (4)$$

Note that we have *no* information on  $S$  except for a few assumptions (e.g., the convexity of  $f_{\sigma}$ ). This implies that two challenging issues are involved in the problem. First, even if we focus on only the static optimization problem

$$\min_{x \in \mathbf{R}^{n_1}} E[f_{\sigma}(x)|x], \quad (5)$$



(a) Step 1: Feedback exploration.

(b) Step 2: Controller extraction.

Fig. 4. Proposed solution framework to Problem 1.

typical methods, based on the explicit form of  $E[f_{\sigma}(x)|x]$  or its gradient, cannot be employed. Second, in the feedback system in Fig. 2, it is impossible to estimate the absolute position of  $P$  through  $y$ . In fact,  $S$  and  $\theta$  are completely unknown to  $K$ . Thus  $K$  has to generate the control input with no position information.

So we often use the body fixed coordinate frame. The frame at time  $\tau$  is denoted by  $\Sigma(\tau)$ , i.e.,

$$\Sigma(\tau) : \begin{bmatrix} z(t) \\ \psi(t) \\ \varphi(t) \end{bmatrix} = \begin{bmatrix} R_{n_1}(-\theta(\tau))(x(t) - x(\tau)) \\ \theta(t) - \theta(\tau) \\ \phi(t) \end{bmatrix} \quad (6)$$

where  $t$  expresses a future time after time  $\tau$ .

### B. Solution Framework for Problem 1

The solution framework proposed here is outlined in Fig. 4. As easily imagined, Problem 1 consists of two issues: the *exploration* of the solution to the static problem (5), and the *control* of the robot. So we decompose Problem 1 into the following problems.

**Problem 2 (Feedback Exploration):** Consider the feedback system in Fig. 4 (a), composed of the signal field  $S$  and a dynamical system  $L$ . For a given  $S$  (but unknown), find a set  $\mathbf{L}$  of systems  $L$  satisfying (4) under any initial state. ■

**Problem 3 (Controller Extraction):** For the robot  $P$ , suppose that a solution  $\mathbf{L}$  to Problem 2 is given. Then find a controller  $K$  such that the feedback connection of  $P$  and  $K$  in Fig. 4 (b) is equivalent to a system  $L^* \in \mathbf{L}$  in the sense of the  $y$ - $x$  relation. ■

Problem 2 is to find systems  $L$  which asymptotically solve the problem (5) with the measurements of  $y$ . Problem 3 is to decompose a system  $L$  into the robot  $P$  and a controller  $K$ , where the system in Fig. 4 (a) is regarded as that in Fig. 2. It is clear that the resulting  $K$  is a solution to Problem 1.

### III. FEEDBACK EXPLORATION BY SIMULTANEOUS- PERTURBATION STOCHASTIC APPROXIMATION

To solve Problem 2, we employ the idea of a stochastic approximation technique, called the *simultaneous-perturbation stochastic approximation (SPSA)* [6]. Since, as stated in Section I, the original SPSA algorithm is not applicable to our source seeking problem, we extend the SPSA algorithm to a more appropriate version.

#### A. Generalized Simultaneous-Perturbation Stochastic Approximation for Source Seeking

A general form of the stochastic approximation algorithm is given by

$$x_{k+1} = x_k - a_k d_k(x_k, c_k \xi_k, \varepsilon_k) \quad (7)$$

where  $x_k \in \mathbf{R}^n$  is the state,  $a_k, c_k \in \mathbf{R}_+$  are the gains,  $\xi_k \in \mathbf{R}^n$  is the random variable with the probability distribution  $p_k : \mathbf{R}^n \rightarrow [0, 1]$ ,  $\varepsilon_k \in \mathbf{R}^l$  is the random noise, and  $d_k : \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^l \rightarrow \mathbf{R}^n$  is the search direction. The algorithm corresponds to the steepest decent if  $d_k(x_k, c_k \xi_k, \varepsilon_k) = \nabla f(x_k)$  for an objective function  $f : \mathbf{R}^n \rightarrow \mathbf{R}$ .

For algorithm (7), we propose the search direction

$$d_k(x_k, c_k \xi_k, \varepsilon_k) := T_k \begin{bmatrix} \frac{(f(x_k + c_{k1} T_k \xi_k) + \varepsilon_{k+}) - (f(x_k - c_{k2} T_k \xi_k) + \varepsilon_{k-})}{(c_{k1} + c_{k2}) \xi_{k1}} \\ \vdots \\ \frac{(f(x_k + c_{k1} T_k \xi_k) + \varepsilon_{k+}) - (f(x_k - c_{k2} T_k \xi_k) + \varepsilon_{k-})}{(c_{k1} + c_{k2}) \xi_{kn}} \end{bmatrix} \quad (8)$$

where  $f : \mathbf{R}^n \rightarrow \mathbf{R}$  is a function,  $\xi_{ki} \in \mathbf{R}$  is the  $i$ th element of the random vector  $\xi_k$ , and  $\varepsilon_{k+}, \varepsilon_{k-} \in \mathbf{R}$  are the noise terms on  $\varepsilon_k$  (i.e.,  $\varepsilon_k := \varepsilon_{k+} - \varepsilon_{k-}$ ),  $c_{k1}, c_{k2} \in \{0\} \cup \mathbf{R}_+$  are the gains, and  $T_k \in \mathbf{R}^{n \times n}$  are matrices. For  $c_k := \max\{c_{k1}, c_{k2}\}$ , we assume

$$c_k > 0. \quad (9)$$

Here,  $f$ ,  $\varepsilon_{k+}$ , and  $\varepsilon_{k-}$  are the problem parameters, that is, depending upon the problem to be solved, while  $a_k$ ,  $\xi_k$  (i.e.,  $p_k(\xi_k)$ ),  $c_{k1}$ ,  $c_{k2}$ , and  $T_k$  are the design parameters of the algorithm. We call the algorithm by (7) and (8) the *generalized simultaneous-perturbation stochastic approximation algorithm* or simply the *G-SPSA algorithm*.

Under several conditions, the G-SPSA algorithm solves the static optimization problem

$$\min_{x \in \mathbf{R}^n} f(x). \quad (10)$$

This fact is formalized in Propositions 1 and 2.

**Proposition 1:** Consider the search direction  $d_k(x_k, c_k \xi_k, \varepsilon_k)$  in (8). If

- the conditions on the problem parameters:
  - (A1)  $f$  is thrice differentiable,
  - (A2)  $E[\varepsilon_k | (x_0, x_1, \dots, x_k), \xi_k] = 0$  w.p.1 for all  $k \in \mathbf{N}$ ,
- the conditions on the design parameters:
  - (B1) (a) the probability distributions  $p_k$  ( $k = 0, 1, \dots$ ) are symmetric about zero (i.e.,  $E[\xi_k] =$

- 0), and  $|\xi_{ki}|$  and  $|\xi_{ki}^{-1}|$  are bounded w.p.1 for all  $(k, i) \in \mathbf{N} \times \{1, 2, \dots, n\}$ ,
- (b)  $\xi_{ki}$  and  $\xi_{kj}$  are independent for all  $(k, i, j) \in \mathbf{N} \times \{1, 2, \dots, n\}^2$  such that  $i \neq j$ ,

hold, then

$$E[d_k(x_k, c_k \xi_k, \varepsilon_k) | x_k] = T_k T_k^\top \nabla f(x_k) + O(c_k^2) \quad (\text{as } c_k \rightarrow 0), \quad (11)$$

where the left hand side expresses the expected value with respect to  $\xi_k$  and  $\varepsilon_k$ .

*Proof:* See Appendix. ■

Proposition 1 implies that, under several assumptions, the expected value of  $d_k(x_k, c_k \xi_k, \varepsilon_k)$  is nearly equal to the product of the time-varying matrix  $T_k T_k^\top$  and the gradient of  $f(x_k)$ . So if

$$T_1 T_1^\top = T_2 T_2^\top = \dots = T_\infty T_\infty^\top = T \quad (12)$$

holds for some nonsingular matrix  $T$ , the algorithm given by (7) and (8) is an approximation of the so-called fixed-point iteration for finding a root of  $T \nabla f(x) = 0$ . Furthermore, since the roots of  $T \nabla f(x) = 0$  are identical to those of  $\nabla f(x) = 0$ , it is expected that the algorithm finds a root of  $\nabla f(x) = 0$ . This conjecture is justified as follows.

**Proposition 2:** For the SPSA algorithm given by (7) and (8), assume that there exists a root  $x^* \in \mathbf{R}^n$  of the equation  $\nabla f(x) = 0$ . If

- the conditions on the problem parameters: (A1), (A2), and
  - (A3)  $E[f(x_k \pm c_k \xi_k)^2]$  is bounded for all  $k \in \mathbf{N}$ ,
  - (A4)  $E[\varepsilon_{k+}^2]$  and  $E[\varepsilon_{k-}^2]$  are bounded for all  $k \in \mathbf{N}$ ,
- the conditions on the design parameters: (B1) and
  - (B2)  $\lim_{k \rightarrow \infty} a_k = 0$ ,  $\sum_{k=0}^{\infty} a_k = \infty$ ,  $\lim_{k \rightarrow \infty} c_k = 0$ , and  $\sum_{k=0}^{\infty} a_k^2 / c_k^2 < \infty$ ,
  - (B3) the random vectors  $\xi_k$  ( $k = 0, 1, \dots$ ) are mutually independent, and  $x_k$  and  $\xi_k$  are mutually independent for all  $k \in \mathbf{N}$ ,
  - (B4)  $E[\xi_{ki}^{-2}]$  is bounded for all  $(k, i) \in \mathbf{N} \times \{0, 1, \dots, n\}$ ,
  - (B5)  $T_k$  ( $k = 0, 1, \dots$ ) satisfy (12) for some nonsingular matrix  $T$ ,
- the coupled conditions on the problem and design parameters:
  - (C1)  $x^*$  is an asymptotically stable equilibrium of  $\dot{x}(t) = -T \nabla f(x(t))$ , where the stability is in the Lyapunov sense and  $T$  is given in (B5),
  - (C2) there exists a compact set  $\mathbf{S} \subseteq \mathbf{R}^n$  such that the following conditions hold for every  $x_0 \in \mathbf{S}$ : (a)  $\dot{x}(t) = -T \nabla f(x(t))$  with  $x(0) = x_0$  results in  $x(\infty) = x^*$ ; (b)  $x_k \in \mathbf{S}$  infinitely often for almost all sample points of  $\xi_k$  ( $k = 0, 1, \dots$ ),
  - (C3)  $\sup_{k \in \mathbf{N}} \|x_k\| < \infty$  w.p.1

hold, then

$$\lim_{k \rightarrow \infty} x_k = x^* \quad \text{w.p.1.} \quad (13)$$

*Proof:* In a similar way to [6], this is proven by the Robbins-Monro Algorithm (see, e.g., [7]). ■

Therefore, under some conditions, a local minimum of the problem (10) is given by the SPSA algorithm.

The above conditions are fairly standard in stochastic approximation [6], [7]. Conditions (A1)–(A5) are concerned with the function  $f$  and the noise term  $(\varepsilon_{k+}, \varepsilon_{k-})$ . (A1) implies that  $f$  is smooth enough and (A2) resembles the common martingale difference noise assumption. (A3) and (A4) prescribes the second order moments of the function  $f$  with the random perturbation by  $\xi_k$  and of the noise term. Next, (B1)–(B5) are imposed for the parameters designed by users, and then they will be a guideline. A typical parameter choice of  $a_k, c_{k1}, c_{k2}, p_k(\xi_k)$ , and  $T_k$  will be

$$a_k := \frac{a}{(k+1)^\alpha}, \quad c_{k1} := \frac{c}{(k+1)^\gamma}, \quad c_{k2} := -c_{k1}, \quad (14)$$

$$p_k(\xi_k) := \prod_{i=1}^n B(\xi_{ki}), \quad T_k = I \quad (15)$$

where  $a, \alpha, c, \gamma \in \mathbf{R}_+$  are arbitrarily given so that  $\alpha \leq 1$  and  $\alpha - \gamma > 0.5$ , and  $B$  is the probability distribution for Bernoulli trial with outcome  $\pm 1$  and equal probabilities. (C1)–(C3) are technical conditions to guarantee the convergence. (C1) is common for decent-type algorithms. (C2) and (C3) are challenging to check, but it is known that they are not restrictive conditions in practice, as addressed in [6], [7].

Similar to the original SPSA algorithm [6], the G-SPSA algorithm has the following two features. First, the algorithm solves the problem (10) when neither  $f$  nor  $\nabla f$  is known, and instead, only noisy measurements of  $f$  are available. Next, the number of measurements to determine the search direction is only *two* and is independent of the dimension  $n$  of  $x$ . This merit can be understood by the fact that the number of measurements for the difference approximation of  $\nabla f = [\partial f(x)/\partial x_1 \cdots \partial f(x)/\partial x_n]^\top$  grows with  $n$ .

The proposed algorithm is equivalent to the original SPSA algorithm in [6] if

$$c_{k1} \equiv c_{k2}, \quad T_k \equiv I \quad (16)$$

in (8). Basically, the proposed algorithm is an extension by the coordinate transformation, while the following difference should be stressed:

- the search direction (8) is based on unequal two-sided perturbations  $c_{k1}T_k\xi_k \neq c_{k2}T_k\xi_k$ , while the original version is of equal perturbations  $c_{k1}T_k\xi_k = c_{k2}T_k\xi_k$ .
- the direction (8) is a time-varying function of  $(x_k, c_k\xi_k, \varepsilon_k)$  (by the time-dependent matrices  $T_k$ ) unlike the time-invariant original one.

These play a key role in solving the source seeking problem.

### B. SPSA Based Solution to Problem 2

Now, we present a solution to Problem 2 based on the SPSA algorithm given by (7) and (8). For the signal field  $S$ , let us introduce the random variable

$$\varepsilon(t) := f_{\sigma(t)}(x(t)) - E[f_{\sigma(t)}(x(t)) | x(t)], \quad (17)$$

whose conditional expected value is zero, i.e.,

$$E[\varepsilon(t) | x(t)] = 0 \quad \forall (t, x(t)) \in [0, \infty) \times \mathbf{R}^{n_1}. \quad (18)$$

By (3) and (17),  $S$  is expressed as

$$y(t) = E[f_{\sigma(t)}(x(t)) | x(t)] + \varepsilon(t). \quad (19)$$

Then by regarding  $x(t)$ ,  $E[f_{\sigma(t)}(x(t)) | x(t)]$ , and  $\varepsilon(t)$  as  $x_k$ ,  $f(x)$ , and  $\varepsilon_k$  in the SPSA algorithm, we obtain a solution to Problem 2.

**Theorem 1:** For the feedback system in Fig. 4 (a), suppose that  $S$  is given. Let  $x^* \in \mathbf{R}^n$  denote the root of the equation  $\nabla f(x) = 0$  ( $x^*$  is uniquely determined by the definition of  $f_\sigma$ ). Let also  $\mathbf{L}$  be the set of stochastic systems  $L$  such that

(i) SPSA condition ( $k = 0, 1, \dots, \infty$ )

$$\begin{cases} x(t_{k1}) = x(t_{k0}) + c_{k1}T_k\xi_k, \\ x(t_{k2}) = x(t_{k0}) - c_{k2}T_k\xi_k, \\ x(t_{(k+1)0}) = x(t_{k0}) - a_kT_k\hat{d}(y(t_{k1}), y(t_{k2}), (c_{k1}+c_{k2})\xi_k), \end{cases} \quad (20)$$

(ii) Boundness condition ( $k = 0, 1, \dots, \infty$ )

$$\begin{aligned} & \|x(t) - x(t_{k0})\| \\ & \leq r \max\{\|c_kT_k\xi_k\|, \|a_kT_k\hat{d}(y(t_{k1}), y(t_{k2}), (c_{k1}+c_{k2})\xi_k)\|\} \\ & \quad \forall t \in [t_{k0}, t_{(k+1)0}] \end{aligned} \quad (21)$$

hold for some

- monotonically nondecreasing time sequence  $(t_{00}, t_{01}, t_{02}, t_{10}, t_{11}, t_{12}, \dots)$  such that (a)  $t_{ij} \rightarrow \infty$  as  $i \rightarrow \infty$  and (b)  $t_{i(j+1)} - t_{ij} > \eta$  and  $t_{(i+1)0} - t_{i2} > \eta$  for all  $(i, j) \in \{0, 1, \dots\} \times \{0, 1\}$ ,
- parameters  $a_k, c_{k1}, c_{k2}, p_k, T_k$  ( $k = 1, 2, \dots$ ) satisfying (B1)–(B5) (and (9)),
- number  $r \in [1, \infty)$ ,

where

$$\hat{d}(y(t_{k1}), y(t_{k2}), (c_{k1}+c_{k2})\xi_k) := \begin{bmatrix} \frac{y(t_{k1})-y(t_{k2})}{(c_{k1}+c_{k2})\xi_{k1}} \\ \vdots \\ \frac{y(t_{k1})-y(t_{k2})}{(c_{k1}+c_{k2})\xi_{kn}} \end{bmatrix}. \quad (22)$$

Then if (C2) and (C3) hold, the set  $\mathbf{L}$  is a solution to Problem 2. ■

*Proof:* Consider a system  $L \in \mathbf{L}$ . Equation (19) and the first two equations of (20) give  $y(t_{k1}) = f(x(t_{k0}) + c_{k1}T_k\xi_k) + \varepsilon(t_{k1})$  and  $y(t_{k2}) = f(x(t_{k0}) - c_{k2}T_k\xi_k) + \varepsilon(t_{k2})$ . This implies that the third equation of (20) is the same as the G-SPSA algorithm given by (7) and (8). Furthermore, by the definitions of  $\sigma$ ,  $f_\sigma$ , and (17) and by the conditions (C3) and (b) imposed for the time sequence, it is clear that (A1)–(A4) and (C1) hold for  $f(x) := E[f_\sigma(x) | x]$ ,  $\varepsilon_{k+} := \varepsilon(t_{k1})$  and  $\varepsilon_{k-} := \varepsilon(t_{k2})$ . This and Proposition 2 prove

$$\lim_{k \rightarrow \infty} x(t_{k0}) = x^* \quad \text{w.p.1.} \quad (23)$$

On the other hand,  $\lim_{k \rightarrow \infty} \hat{d}(y(t_{k1}), y(t_{k2}), (c_{k1} + c_{k2})\xi_k) = 0$  holds w.p.1 subject to (23). Thus, under (B1), (B3), and (23), the right hand side of (21) converges to zero

as  $k \rightarrow \infty$ . This implies the convergence of the inter-sample behavior, i.e.,  $\lim_{t \rightarrow \infty} x(t) = \lim_{k \rightarrow \infty} x(t_{k0})$  w.p.1. This and (23) imply (4). Hence,  $\mathbf{L}$  is a solution to Problem 2. ■

#### IV. STOCHASTIC SOURCE SEEKING CONTROLLERS

This section gives a solution to Problem 3.

##### A. A General Solution to Problem 3

We can transform the condition (20) into

$$\begin{cases} z(t_{k1}) = c_{k1} R_{n_1}(-\theta(t_{k0})) T_k \xi_k, \\ z(t_{k2}) = -c_{k2} R_{n_1}(-\theta(t_{k0})) T_k \xi_k, \\ z(t_{(k+1)0}) = -a_k R_{n_1}(-\theta(t_{k0})) T_k \\ \quad \times \hat{d}(y(t_{k1}), y(t_{k2}), (c_{k1} + c_{k2}) \xi_k) \end{cases} \quad (24)$$

in the body fixed frame  $\Sigma(t_{k0})$ . Here, let us choose  $c_{k1} := 0$  and  $T_k := R_{n_1}(\theta(t_{k0}))$  ( $k = 1, 2, \dots$ ) by noting that they do not violate (9) and (B6) and that  $T_k T_k^\top \equiv I$ . We then have  $z(t_{k1}) = 0$ ,  $z(t_{k2}) = -c_{k2} \xi_k$ , and  $z(t_{(k+1)0}) = -a_k \hat{d}(y(t_{k1}), y(t_{k2}), c_{k2} \xi_k)$ . This is a condition excluding the position information of  $\theta(t_{k0})$  in the world coordinate frame and thus is suitable in the sensor-free situation.

In addition, for  $t_d \in \mathbf{R}_+$ ,  $\Delta z \in \mathbf{R}^{n_1}$ ,  $\Delta \psi \in \mathbf{R}^{n_2}$ , and  $r \in [1, \infty)$ , we denote by

$$v \left( t, t_d, 0 \rightarrow \begin{bmatrix} \Delta z \\ \Delta \psi \end{bmatrix}, r \right)$$

an input trajectory on the time interval  $[t_c, t_c + t_d]$  such that, in the body fixed frame at time  $t_c$ ,  $P$  is steered from  $[z(t_c) \ \psi(t_c)]^\top = 0$  (the current position) to  $[z(t_c + t_d) \ \psi(t_c + t_d)]^\top = [\Delta z \ \Delta \psi]^\top$  keeping  $\|z(t)\| \leq r \|\Delta z\|$  where  $t_c$  is the time when  $v$  is applied to  $P$ . This expresses a finite point-to-point controller for  $P$ .

Then our solution to Problem 3 is presented as follows.

**Theorem 2:** For the robot  $P$  and the set  $\mathbf{L}$  in Theorem 1, suppose that the following tuning parameters are given:

- monotonically nondecreasing time sequence  $(t_{00}, t_{02}, t_{10}, t_{12}, \dots)$  such that (a)  $t_{ij} \rightarrow \infty$  as  $i \rightarrow \infty$  and (b)  $t_{i(j+1)} - t_{ij} > \eta$  and  $t_{(i+1)0} - t_{i2} > \eta$  for all  $(i, j) \in \{0, 1, \dots\} \times \{0, 1\}$ ,
- $a_k, c_{k2}, p_k$  ( $k = 1, 2, \dots$ ) satisfying (B1)–(B5) subject to  $c_{k1} \equiv 0$ ,
- $\Delta \psi_{ki} \in \mathbf{R}^{n_2}$  ( $k = 1, 2, \dots, \infty, i = 1, 2$ ),
- $r \in [1, \infty)$ .

Then the controller  $K$  such that

$$u(t) = \begin{cases} v \left( t, t_{k2} - t_{k0}, 0 \rightarrow \begin{bmatrix} -c_{k2} \xi_k \\ \Delta \psi_{k1} \end{bmatrix}, r \right) & \text{if } t \in [t_{k0}, t_{k2}], \\ v \left( t, t_{(k+1)0} - t_{k2}, \right. \\ \quad \left. 0 \rightarrow \begin{bmatrix} c_{k2} \xi_k - a_k \hat{d}(y(t_{k1}), y(t_{k2}), c_{k2} \xi_k) \\ \Delta \psi_{k2} \end{bmatrix}, r \right) & \text{if } t \in [t_{k2}, t_{(k+1)0}] \end{cases} \quad (25)$$

for every  $k = 0, 1, \dots, \infty$  is a solution to Problem 3.

*Proof:* For the feedback system in Fig. 4 (b), suppose that  $K$  is given by (25). Then it is obvious that the two

conditions in Theorem 1 hold for  $c_{k1} := 0$  and  $T_k := R_{n_1}(\theta(t_{k0}))$ . ■

From Theorems 1 and 2 and the relation  $T_k T_k^\top \equiv I$ , our answer to the stochastic source seeking problem is presented.

**Theorem 3:** For the feedback system in Fig. 2, suppose that  $P$  and  $S$  are given. If (C2) and (C3) hold for the G-SPSA algorithm with the parameters  $a_k, c_{k2}, p_k$  specified in Theorem 2,  $c_{k1} := 0$ , and  $T_k := R_{n_1}(\theta(t_{k0}))$ , the controller  $K$  in Theorem 2 is a solution to Problem 1. ■

We comment on how to obtain the point-to-point control inputs  $v$  in (25). A unified method to obtain them is to utilize the so-called Lie bracket motion (see, e.g., [8]). By applying some periodic input, the robot  $P$  is steered into a direction in the translational space  $\mathbf{R}^{n_1}$ . The moving direction is determined by the input period, the initial state, and the accessibility distribution. Under some controllability assumptions, we can find a period and an amplitude of the periodic input which drives  $P$  to a desired position in  $\mathbf{R}^{n_1}$ .

##### B. Source Seeking Controllers for Unicycle

Now, we focus on the unicycle in Fig. 3 and show that there exist very simple source seeking controllers.

When  $\xi_k \in \{-\sigma, \sigma\}^n$  for  $\sigma \in \mathbf{R}_+$ , the vector  $T_k \xi_k$  is linearly dependent on  $d_k(x_k, c_k \xi_k, \varepsilon_k)$  in (8). Using this fact, we obtain the following result.

**Theorem 4:** For the feedback system in Fig. 2, suppose that  $P$  is the unicycle (2) and  $S$  is given. Let  $K$  be the discrete-time stochastic controller

$$\begin{cases} y_{pre}[\tau+1] = y[\tau], \\ u[\tau] = \begin{cases} \begin{bmatrix} 0 \\ \frac{1}{h} \delta[\tau] \end{bmatrix} & \text{if } \tau=0, 3, 6, \dots, \\ \begin{bmatrix} -\frac{\lambda}{h} c_k \\ 0 \end{bmatrix} & \text{if } \tau=1, 4, 7, \dots, \\ \begin{bmatrix} \frac{\lambda}{h} \left( c_k - a_k \frac{y_{pre}[\tau] - y[\tau]}{c_k} \right) \\ 0 \end{bmatrix} & \text{if } \tau=2, 5, 8, \dots \end{cases} \end{cases} \quad (26)$$

where  $\tau \in \mathbf{N}$  is the discrete time for the sampling period  $h \in (\eta, \infty)$  ( $\eta$  is the switching time period for  $f_\sigma$ ),  $y_{pre}[\tau] \in \mathbf{R}$  is the state to save  $y(t_{k1})$ ,  $\delta[\tau]$  is an i.i.d random number drawn from the uniform distribution on  $\{\pi/4, 3\pi/4, -3\pi/4, -\pi/4\}$ ,  $\lambda \in \mathbf{R}_+$  is given as  $\lambda := \sqrt{2}$ , and  $a_k$  and  $c_k$  ( $k = 0, 1, \dots$ ) are arbitrarily given sequences satisfying (B2). If (C2) and (C3) hold for the G-SPSA algorithm with  $a_k, c_{k1} := 0, c_{k2} := c_k$ , the Bernoulli-type distribution  $p_k$  in (15), and  $T_k := R_{n_1}(\theta(t_{k0}))$ ,  $K$  is a solution to Problem 1. ■

The controller  $K$  steers the robot  $P$  as shown in Fig. 5. Three steps on a line segment are repeated: the random turn, the forward movement, and the forward or backward movement. The first and second steps correspond to the first input in (25). The final step does to the second input, and its moving direction is determined by whether the slope  $y(t_{k1}) - y(t_{k2})$  is positive or not.

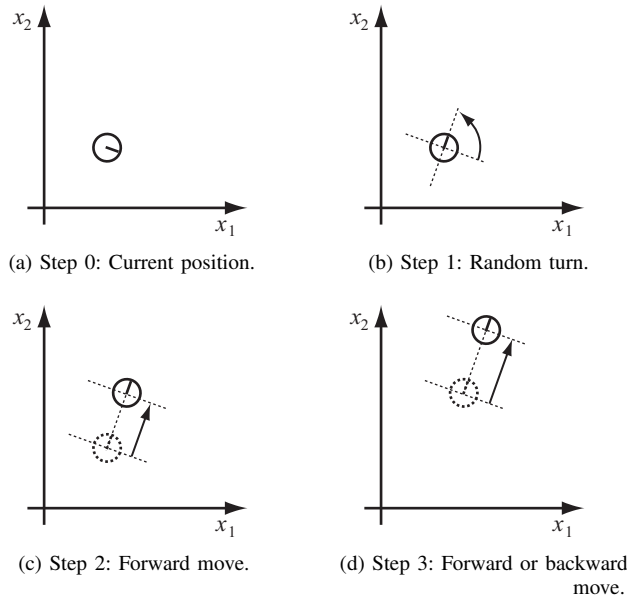


Fig. 5. Robot motion by proposed controller in Theorem 4 (by repeating these steps, the robot is guided to the source).

### C. Example

Consider Problem 2, where the robot  $P$  is the unicycle (2) and the signal field  $S$  is given by the functions

$$f_1(x) = \left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 110 \\ 80 \end{bmatrix} \right)^\top \begin{bmatrix} 0.01 & -0.005 \\ -0.005 & 0.01 \end{bmatrix} \left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 110 \\ 80 \end{bmatrix} \right),$$

$$f_2(x) = \left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 90 \\ 90 \end{bmatrix} \right)^\top \begin{bmatrix} 0.01 & 0.001 \\ 0.001 & 0.003 \end{bmatrix} \left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 90 \\ 90 \end{bmatrix} \right),$$

$$f_3(x) = \left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 109 \\ 110 \end{bmatrix} \right)^\top \begin{bmatrix} 0.02 & 0.003 \\ 0.003 & 0.01 \end{bmatrix} \left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 109 \\ 110 \end{bmatrix} \right)$$

and the probability distribution  $q(1) = 0.15$ ,  $q(2) = 0.15$ , and  $q(3) = 0.7$ . Its source is  $\operatorname{argmin}_{x \in \mathbf{R}^2} E[f_\sigma(x)|x] \simeq [109.13 \ 103.54]^\top$ . The controller  $K$  is given by Theorem 4 for  $h := 1$  and the parameters  $a_k$  and  $c_k$  are given by (14) with  $a := 15$ ,  $\alpha := 0.55$ ,  $c := 10$ , and  $\gamma := 0.03$ .

Fig. 6 illustrates the contour plot of  $E[f_\sigma(x)|x]$  and the moving trajectory of  $P$  from the initial state  $(x(0), \theta(0)) := ([182 \ 41]^\top, 10)$ , where the isosceles triangles express  $(x(t_{ki}), \theta(t_{ki}))$  ( $k = 0, 1, \dots, \infty, i = 0, 1, 2$ ). We see that the robot  $P$  is guided to the source by the simple controller.

**Acknowledgments:** The first author would like to thank Prof. J.C. Spall, Johns Hopkins University, for his valuable information on one-sided SPSA algorithms.

### APPENDIX: PROOF OF PROPOSITION 1

**Notation:** Let  $e_i$  be the  $i$ th standard basis of  $\mathbf{R}^n$ . For the thrice differentiable function  $f : \mathbf{R}^n \rightarrow \mathbf{R}$ , let  $f^{(1)}(x) := (\nabla f(x))^\top$  and  $f^{(j)}(x) := [\partial f^{(j-1)}(x)/\partial x_1 \ \dots \ \partial f^{(j-1)}(x)/\partial x_n] \in \mathbf{R}^{1 \times n^j}$  ( $j = 2, 3$ ). Using this,  $f(x+y)$  is expressed by Taylor's theorem as  $f(x+y) = f(x) + f^{(1)}(x)y + (1/2)f^{(2)}(x)(y \otimes y) + (1/6)f^{(3)}(\bar{x})(y \otimes y \otimes y)$  where  $\otimes$  expresses the Kronecker product and  $\bar{x}$  is a vector on the line segment between  $x$  and  $x+y$ .

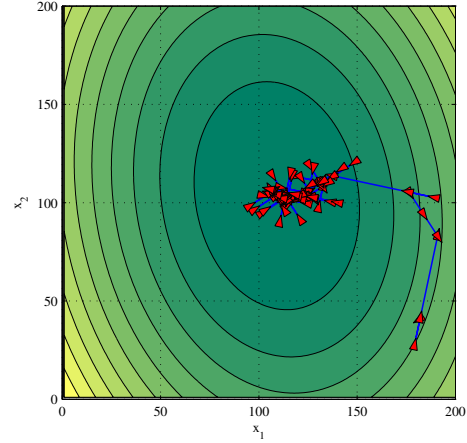


Fig. 6. Moving trajectory of unicycle by G-SPSA based controller.

**Proof:** Noting (A1), we apply Taylor's theorem to the terms  $f(x_k + c_{k1}T_k\xi_k)$  and  $f(x_k - c_{k2}T_k\xi_k)$  in (8). Then

$$\frac{f(x_k + c_{k1}T_k\xi_k) - f(x_k - c_{k2}T_k\xi_k)}{(c_{k1} + c_{k2})\xi_{ki}}$$

$$= f^{(1)}(x_k)T_k \frac{\xi_k}{\xi_{ki}} + \frac{(c_{k1} - c_{k2})}{2} f^{(2)}(x_k)(T_k \otimes T_k) \frac{\xi_k \otimes \xi_k}{\xi_{ki}}$$

$$+ \frac{1}{6} \left( \frac{c_{k1}^3}{c_{k1} + c_{k2}} f^{(3)}(\bar{x}_{k1}) + \frac{c_{k2}^3}{c_{k1} + c_{k2}} f^{(3)}(\bar{x}_{k2}) \right)$$

$$\times (T_k \otimes T_k \otimes T_k) \frac{\xi_k \otimes \xi_k \otimes \xi_k}{\xi_{ki}} \quad (27)$$

where  $\bar{x}_{k1}$  and  $\bar{x}_{k2}$  are vectors on the line segment between  $x_k + c_{k1}T_k\xi_k$  and  $x_k - c_{k2}T_k\xi_k$ . Here, from (B1) (b) and the fact that (B1) (a) implies  $E[\xi_{ki}^{-1}] = 0$ , the relations  $E[\xi_k/\xi_{ki}] = e_i$  and  $E[(\xi_k \otimes \xi_k)/\xi_{ki}] = 0$  hold, which gives

$$E \left[ \frac{f(x_k + c_{k1}T_k\xi_k) - f(x_k - c_{k2}T_k\xi_k)}{(c_{k1} + c_{k2})\xi_{ki}} \middle| x_k \right]$$

$$= f^{(1)}(x_k)T_k e_i + O(c_k^2) = e_i^\top T_k^\top \nabla f(x_k) + O(c_k^2).$$

Therefore, by applying this and (A2) to the expected value of the right hand side of (8), we get (11).

### REFERENCES

- [1] E. Burian *et al.*, "Gradient search with autonomous underwater vehicles using scalar measurements," *Proc. of IEEE Symp. on Auto. Underwater Veh. Tech.*, pp. 86–98, 1996.
- [2] C.G. Mayhew *et al.*, "Robust hybrid source-seeking algorithms based on directional derivatives and their approximations," *Proc. of IEEE Conf. on Decis. Contr.*, pp. 1735–1740, 2008.
- [3] A.R. Mesquita *et al.*, "Optimotaxis: A stochastic multi-agent optimization procedure with point measurements," *Hybrid Systems: Computation and Control*, Springer-Verlag, Berlin, pp. 358–371, 2008.
- [4] C. Zhang *et al.*, "Source seeking with non-holonomic unicycle without position measurement and with tuning of forward velocity," *Syst. Contr. Lett.*, 56-3, pp. 245–252, 2007.
- [5] J. Cochran and M. Krstic, "Nonholonomic source seeking with tuning of angular velocity," *IEEE Trans. Automat. Contr.*, 54-4, pp. 717–731, 2009.
- [6] J.C. Spall, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," *IEEE Trans. Automat. Contr.*, 37-3, pp. 332–341, 1992.
- [7] H.J. Kushner and D.S. Clark, *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, Springer-Verlag, 1978.
- [8] H. Nijmeijer and A.J. van der Schaft, *Nonlinear Dynamical Control Systems*, Springer-Verlag, 1990.