2020 IEEE/ASME International Conference on
Advanced Intelligent Mechatronics (AIM), Boston, USA
(Virtual Conference), July 6-9, 2020

A Bidirectional Alignment Control Approach for Planar LED-based Free-Space Optical Communication Systems

Pratap Bhanu Solanki, Shaunak D. Bopardikar and Xiaobo Tan

Abstract— **Achieving and maintaining line-of-sight (LOS) is essential for free-space optical (FSO) communication systems due to the high directionality of the optical signals. We consider the problem of achieving LOS in a planar setting, where each of the two parties (called agents) makes a move only based on its own light intensity measurement. The problem is formulated as a discrete time dynamical system, where each agent seeks to maximize its own reward function that depends on the states of both agents, who make their moves in parallel. In particular, the reward function corresponds to the light intensity measurements made by each agent. While the two reward functions are non-conflicting (i.e., the optimization of one reward function helps the optimization of the other), the constraints of no information exchange between the agents, no access to states, and parallel (non-sequential) actions pose significant challenges. A novel iterative optimization algorithm meeting all of these constraints is proposed. We show that the proposed algorithm brings the system to a neighborhood of the LOS in a finite number of steps, when the reward functions are of the Gaussian form as supported by the experimental data. The effectiveness of the approach is evaluated in simulation by comparison with extremum-seeking control, where it shows an order of magnitude better performance in terms of the convergence speed.**

I. INTRODUCTION

The free-space optical (FSO) communication systems are promising for high-bandwidth wireless underwater communication. For the FSO communication systems, establishing and maintaining near line-of-sight (LOS) links is essential as the light signals are highly directional. It is desirable to achieve LOS, without relying on the communication between the agents (e.g., robots, which are often moving) as the quality of communication link itself depends on the LOS. This paper proposes, analyzes and evaluates one such scheme for two transmitter-receiver pairs with planar motion to achieve and maintain LOS from arbitrary allowable initial configurations.

Several approaches have been proposed to address the LOS requirement in optical communication systems. Multiple LEDs and/or multiple photo-diodes have been used to avoid the need of active pointing during optical-communication [1],[2]. Soysal *et al.* used quadrant photo-detector [3] with Kalman filtering to simultaneously estimate the azimuthal and elevation errors, which were further used as feedback for beam tracking. Jeon *et al.* [4] used position sensitive detector (PSD) and MEMS-based optical scanner for active bi-directional optical links. In our previous works, [5], [6],[7], we demonstrated one-sided tracking using a single photo-diode on the receiver, where only one agent uses an active alignment mechanism along with extended Kalman filtering to achieve the LOS. In all of these aforementioned works, active alignment is demonstrated in simulation or experimental setup, but they typically do not provide any theoretical guarantee on the stability or the convergence rate.

This paper considers the setup of two agents seeking to establish LOS for bidirectional communication in a planar setting, and formulate it as a two-agent optimization problem wherein each agent is assumed to have a local measurement of its reward function of the form, $y_i = g(x_i)h(x_{3-i})$, $i \in \{1, 2\}$, where x_1 and x_2 are the states of the two agents as relevant to LOS, and the functions g and h are of Gaussian type, as supported by experimental data. This makes the reward functions non-conflicting; both have the same maximizer (x_1^*, x_2^*), which corresponds to the LOS configuration. Based on the setup, the agents do not have access to their own state. In addition to the no communication constraint between the agents, they are assumed to act simultaneously at each time step, which eliminates candidate solutions based on sequential actions of the agents that can greedily optimize their instantaneous reward functions. The problem is to design a scheme such that the states of the agents reach an ϵ-neighborhood of the LOS configuration within a finite number of steps.

The problem in this paper can be considered as a special case of a multi-agent optimization problem in which each agent seeks to optimize their own cost (reward) function that depends on the state of the other agents. In multi-agent game-theoretic formulations, gradient play is a popular technique that converges to a Nash equilibrium for the game under mild technical assumptions [8], [9]. However, these techniques require that each agent has access to the gradient of its own cost function. Passivity-based tools have also been studied for multi-agent synchronization and extremum-seeking problems [10], [11]. However, these require information exchange between the agents or in certain cases, access to the gradient of the local cost function with respect to the state. A version of extremum-seeking control algorithm [12] is applicable to the problem structure, and is thus used as a benchmark approach for the comparison with the proposed method in this paper. The algorithm exhibits delay in convergence due to significant exploration by each agent.

The contributions of this paper are as follows. We propose a novel approach that only requires the information of the current and the immediately preceding reward function measurements by each agent. Under the assumption that

*This work was supported by the National Science Foundation(ECCS 1446793, IIS 1734272)

P.B. Solanki, S.D. Bopardikar and X. Tan are with the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI, 48824 USA. Email: prabhanu@msu.edu (P. B. S.), shaunak@egr.msu.edu (S.D.B.), xbtan@egr.msu.edu (X.T.).

the reward functions are a product of Gaussian functions, we prove that from any initial values of the states, and as per the proposed approach, both agents reach a specified ϵ-neighborhood of the optimal solution (corresponds to the LOS, and without the loss of generality can be considered as the origin) in a finite number of steps. The algorithm is evaluated in simulation, and a comparison with extremum-seeking control demonstrates *an order of magnitude advantage in terms of the speed of convergence.*

The rest of the paper is organized as follows. In Section II, the hardware description of the system is presented, followed by problem formulation in the state-space domain. In Section III, the proposed scheme is presented along with the main result of the paper and its mathematical proof. The simulation results along with a brief description of the implementation of the extremum-seeking approach are presented in Section IV. Finally, concluding remarks are provided in Section V.

II. System Setup and Modelling

A. System Setup

Consider two agents (e.g., robots) in a planar environment as illustrated in Fig. 1. The line joining the robots is the LOS. The distance between the robots is d. The optical axis of robot $\mathbf{R}_i$ makes angle θ_i with the LOS line, where $i \in \{1,2\}, \theta_i \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. In this work, we assume that the positions of the agents are fixed and they can only change their angle θ_i.

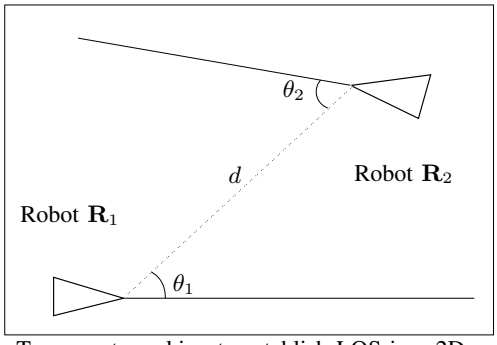

Fig. 1: Two agents seeking to establish LOS in a 2D scenario.

Fig. 2 shows the transceiver hardware setup for LED-based free-space optical communication, which has been developed in our prior work [6] and can be mounted on mobile robots. Each transceiver has two devices: a photo-diode with a lens and a light-emitting-diode (LED) with a lens. For the purpose of mathematical modelling, the agents are considered as points and the optical axes of the devices are assumed to be aligned for each robot. Moreover, the transceiver is mounted on a rotating platform, which enables the adjustment of transceiver orientation.

B. Mathematical Modeling

The light signal strength model adopted here largely follows [13] with minor adjustments to suit the experimental prototype considered in this paper. The light signal strength measured by the photo-diode of agent $\mathbf{R}_1$ (in V) and transmitted by agent $\mathbf{R}_2$'s LED is given by,

$$V_d = C_p \frac{e^{-cd}}{d^2} h(\theta_2) g(\theta_1), \tag{1}$$

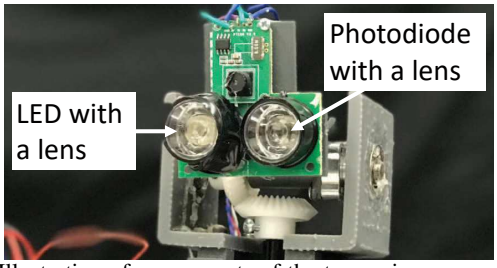

Fig. 2: Illustration of components of the transceiver on each agent.

where C_p is a constant of proportionality, c is the attenuation coefficient of the transmission medium, and $h(\theta_2)$ corresponds to the angular intensity distribution of $\mathbf{R}_2$'s LED (same as the optical signal radiant intensity [13]). Similarly, $g(\theta_1)$ is the angular sensitivity of the photo-diode of $\mathbf{R}_1$. An expression analogous to (1) holds for the light intensity measured by $\mathbf{R}_2$. It is to be noted that $h(\cdot)$ and $g(\cdot)$ are setup dependent and can be experimentally characterized. Fig. 3 shows the data collected on our setup for $g(\cdot)$ and the corresponding Gaussian fitting function. The collected data for the function $h(\cdot)$ has a similar Gaussian fitting function but with a smaller width.

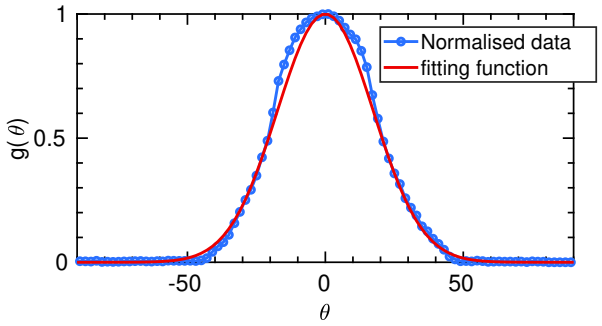

Fig. 3: Illustration of the Gaussian approximation of the fitting functions of photo-diode sensitivity curve.

The state variables x_1 and x_2 are the angles θ_1 and θ_2, respectively. Now, since the robot positions are fixed, the distance d is constant and can be merged into a new proportional constant C, resulting in the measurement functions

$$\mathbf{y}_k \triangleq \begin{bmatrix} y_{1,k} \\ y_{2,k} \end{bmatrix} = \begin{bmatrix} Ch(x_{2,k})g(x_{1,k}) \\ Ch(x_{1,k})g(x_{2,k}) \end{bmatrix} = \begin{bmatrix} Ce^{-(\frac{x_{1,k}^2}{a^2} + \frac{x_{2,k}^2}{b^2})} \\ Ce^{-(\frac{x_{1,k}^2}{b^2} + \frac{x_{2,k}^2}{a^2})} \end{bmatrix}, \tag{2}$$

where a and b represents the width of the Gaussian fitting functions $g(\cdot)$ and $h(\cdot)$ with $b < a$. To facilitate the state-space formulation, we consider the following dynamics for the states:

$$x_{i,k+1} = x_{i,k} + u_{i,k}, \tag{3}$$

where $u_{i,k}$ represents the control of the i-th agent with

$$u_{i,k} = U(y_{i,k}, ..., y_{i,0}, u_{i,k-1}, ..., u_{i,0}) \; : \; \|u_{i,k}\| \leq \delta, \tag{4}$$

and $U : \mathbb{R}^{2k+1} \longrightarrow \mathbb{R}$. In this case, δ is the size of the maximum allowable step size, assumed to be the same for both agent. Eq. (4) captures the constraint of no communication between agents since the control term of an agent i can

only depend on its own history of measurements and control inputs. Next, we state a generalized problem statement for the aforementioned class of two-agent systems.

Problem Statement: Design a control law of the form (4) for a system defined in (2) and (3) such that for any permissible initial state $\mathbf{x}_0$, $\exists$ a finite $K \in \mathbb{Z}^+ : \|\mathbf{x}_K\|_1 \leq \epsilon \in O(\delta)$. $\|\cdot\|_1$ denotes the ℓ_1-norm (*Manhattan* norm) of a vector and the term ϵ defines the size of the neighborhood around the LOS, and as that size tends to zero, the agents achieve perfect LOS.

III. Algorithms and Analysis

In this section, we propose a control law complying with the requirements in eq. (4), with an analysis to show that the states of the agents are within an ϵ-neighborhood of the origin (optimum) in a finite number of steps K.

A. Proposed Control Law and Main Result

Consider the following simple control law:

$$\begin{bmatrix} u_{1,k} \\ u_{2,k} \end{bmatrix} = \begin{bmatrix} \text{sgn}(y_{1,k} - y_{1,k-1})u_{1,k-1} \\ \text{sgn}(y_{2,k} - y_{2,k-1})u_{2,k-1} \end{bmatrix}, \tag{5}$$

$$\text{where, } \text{sgn}(p) = \begin{cases} +1, & \text{if } p \geq 0, \\ -1, & \text{otherwise,} \end{cases} \tag{6}$$

which essentially means that an agent takes an action in the direction of increasing light intensity measurement, and when it observes a reduction in its measurement, it switches to an action in the opposite direction. The initial direction $(u_{i,0})$ is chosen at random: $\mathbf{u}_0 = (u_{1,0}, u_{2,0}) \in \{[+\delta, +\delta]^T, [+\delta, -\delta]^T, [-\delta, +\delta]^T, [-\delta, -\delta]^T\}$, $\delta > 0$. Fig. 4 illustrates a sample trajectory under eq. (5) from an arbitrary initial condition S. Theorem 3.1 summarizes the guarantees associated with the control law.

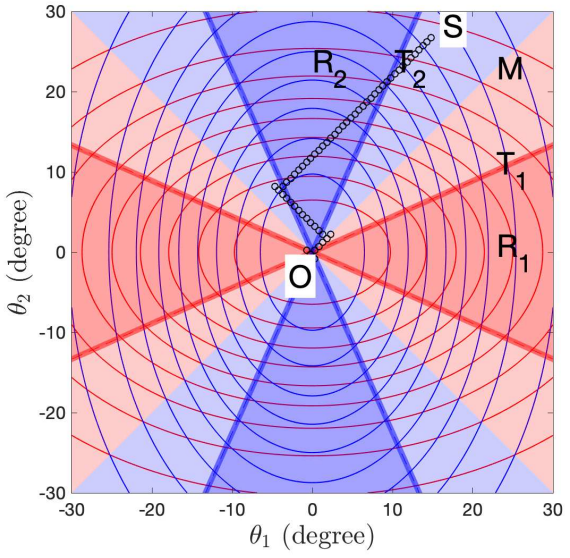

Fig. 4: Illustration of a sample trajectory along with level contours for measurement functions and regions R_1, R_2 and M defined in the proof of Theorem 3.1 in subsection III-B.

*Theorem 3.1 (**Main Result**):* Under the assumption that $a > b$, from any initial condition $\mathbf{x}_0$, the control law defined in eq. (5) satisfies the requirements in the problem statement for the system defined by eqs. (3) and (2) with

$$1)\ \epsilon = \frac{3\delta}{2}\left(1 + \frac{a^2}{b^2}\right), \text{ and} \tag{7}$$

$$2)\ K = \frac{\sqrt{2}\|\mathbf{x}_0\|_1}{\delta} + 4\left(\left\lceil \frac{\ln(\frac{2\|\mathbf{x}_0\|_1}{3\delta})}{\ln(\frac{a^2+b^2}{a^2-b^2})} \right\rceil + 1\right). \tag{8}$$

Remark 1: Suppose that the system is already inside the ϵ-neighborhood and leaves the neighborhood, then the theorem applies to the new initial condition and the state returns to the neighborhood in finite steps.

The next subsection is dedicated to the proof of Theorem 3.1.

B. Proof of Theorem 3.1

Consider the transformed measurement $\mathbf{z}$:

$$\mathbf{z}_k \triangleq \begin{bmatrix} z_{1,k} \\ z_{2,k} \end{bmatrix} = \begin{bmatrix} -\ln(y_{1,k}) \\ -\ln(y_{2,k}) \end{bmatrix} = \begin{bmatrix} \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} \\ \frac{x_1^2}{b^2} + \frac{x_2^2}{a^2} \end{bmatrix}.$$

The level contours of $\mathbf{z}$ and equivalently, of y_1 and y_2, are ellipses as illustrated in Fig. 4. Consequently, the control law defined in Eq (5) can be rewritten as:

$$\begin{bmatrix} u_{1,k} \\ u_{2,k} \end{bmatrix} = \begin{bmatrix} -\text{sgn}(z_{1,k} - z_{1,k-1})u_{1,k-1} \\ -\text{sgn}(z_{2,k} - z_{2,k-1})u_{2,k-1} \end{bmatrix}. \tag{9}$$

We begin with the notion of an improving control direction.

*Definition 3.1 (**Improving control direction**):* Consider the following conditional function

$$F_{imp}(\mathbf{x}_k, \mathbf{u}_k) = \begin{cases} 1, & \text{if } \nabla z_{1,k} \cdot \mathbf{u}_k < 0 \ \&\ \nabla z_{2,k} \cdot \mathbf{u}_k < 0 \\ 0, & \text{otherwise} \end{cases},$$

with $\nabla z_{1,k} = \left[2x_1/a^2 2x_2/b^2\right]$, $\nabla z_{2,k} = \left[2x_1/b^2 2x_2/a^2\right]$. For any point $\mathbf{x}_k$, a control term $\mathbf{u}_k$ is *improving* if $F_{imp}(\mathbf{x}_k, \mathbf{u}_k) = 1$. □

For the set of points $\mathbf{x}$ such that $\|\mathbf{x}\|_1 > \epsilon$, the allowable orientation space is divided into the following regions, based on the notion of improving control direction:

- $R_1 : \left|\frac{x_1}{a^2}\right| - \frac{\epsilon}{3a^2} > \left|\frac{x_2}{b^2}\right|$, $\mathbf{u} = [-\text{sgn}(x_1)\delta, \pm\delta]^T$ (10)
- $T_1 : \left|\frac{x_1}{a^2}\right| - \frac{\epsilon}{3a^2} \leq \left|\frac{x_2}{b^2}\right| \leq \left|\frac{x_1}{a^2}\right| + \frac{\epsilon}{3a^2}$. (11)
- $M : \left|\frac{x_1}{a^2}\right| + \frac{\epsilon}{3a^2} < \left|\frac{x_2}{b^2}\right|$, $\left|\frac{x_2}{a^2}\right| + \frac{\epsilon}{3a^2} < \left|\frac{x_1}{b^2}\right|$, $\mathbf{u} = [-\text{sgn}(x_1)\delta, -\text{sgn}(x_2)\delta]^T$ (12)
- $T_2 : \left|\frac{x_2}{a^2}\right| - \frac{\epsilon}{3a^2} \leq \left|\frac{x_1}{b^2}\right| \leq \left|\frac{x_2}{a^2}\right| + \frac{\epsilon}{3a^2}$. (13)
- $R_2 : \left|\frac{x_2}{a^2}\right| - \frac{\epsilon}{3a^2} > \left|\frac{x_1}{b^2}\right|$, $\mathbf{u} = [\pm\delta, -\text{sgn}(x_2)\delta]^T$ (14)

The regions T_1 and T_2 are *transition* regions, where an improvement direction becomes tangent to one of the level contours. Moreover, if a point is outside the ϵ neighborhood ($\|\mathbf{x}\|_1 > \epsilon$), then

$$\Rightarrow \frac{\|\mathbf{x}\|_1}{a^2} > \frac{\epsilon}{a^2} \Rightarrow \left|\frac{x_1}{a^2}\right| + \left|\frac{x_2}{b^2}\right| > \frac{\epsilon}{a^2} \text{ as } a > b, \tag{15}$$

$$\text{similarly, } \|\mathbf{x}\|_1 > \epsilon \Rightarrow \left|\frac{x_1}{b^2}\right| + \left|\frac{x_2}{a^2}\right| > \frac{\epsilon}{a^2}. \tag{16}$$

Fig. 5 illustrates the defined regions along with the above inequality bounds.

*Lemma 1 (**Signum and absolute value properties**):* For $p \in \mathbb{R}$ and $q \in \mathbb{R}$, the following properties hold true.

$$1)\ \text{sgn}(pq) = \text{sgn}(p)\,\text{sgn}(q) \tag{17}$$

$$2)\ |p| > |q| \Rightarrow \text{sgn}(p+q) = \text{sgn}(p) \tag{18}$$

$$3)\ |p| - |q| \leq |p - q| \leq |p| + |q| \tag{19}$$

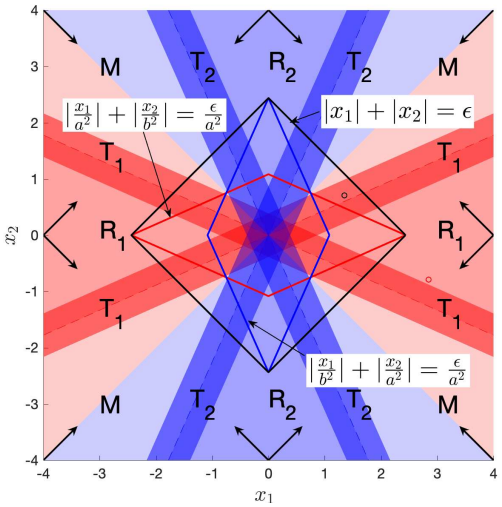

Fig. 5: Illustration of regions defined based on the notion of the improving direction.

4) $p = \text{sgn}(p)|p|, \ |p| = \text{sgn}(p)p$ (20)
5) $q > 0 \Rightarrow \text{sgn}(qp) = \text{sgn}(p)$ (21)

These properties can be easily verified and will be repeatedly used in the subsequent analysis.

*Lemma 2 (**Initialization**):* At any given initial point $\mathbf{x}_0$ with $\|\mathbf{x}_0\|_1 > \epsilon$, and an initial control term $\mathbf{u}_0$ with the control law defined in eq. (9),

$$F_{imp}(\mathbf{x}_0, \mathbf{u}_{k_{imp}}) = 1 \tag{22}$$

for some $k_{imp} \in \{1, 2, 3, 4\}$. □

The lemma essentially says that in the beginning, from any arbitrary point outside the ϵ neighborhood around origin, with any arbitrary initial direction, the system takes up to 4 steps to get into an improving control direction. The proof the lemma is provided in appendix A.

*Lemma 3 (**Subsequent steps**):* If $\mathbf{x}_k \in R$, $\mathbf{x}_{k+1} \in R$, $R \in \{R_1, M, R_2\}$ and $F_{imp}(\mathbf{x}_k, \mathbf{u}_k) = 1$, then,

1) $\mathbf{u}_{k+1} = \mathbf{u}_k$ and
2) $F_{imp}(\mathbf{x}_{k+1}, \mathbf{u}_{k+1}) = 1$. □

This result implies that once the control is in improving direction, it would continue to be so until the system state changes its region. The proof of the lemma can be derived using the properties of the regions.

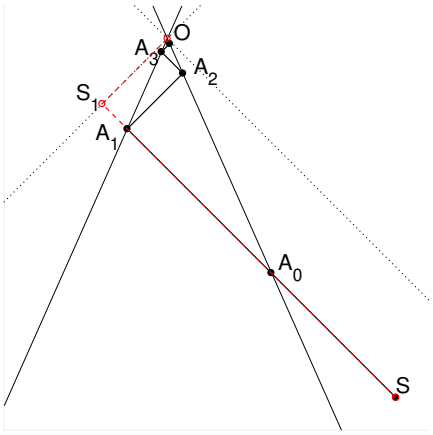

Fig. 6: Geometry of the evolved path.

Outline of the proof of Theorem 3.1. Now consider the path of sample trajectory of the system illustrated in Fig. 6, where the system starts at a general point S in M (we shall see that the initial point can be any other point in any other region). According to Lemma 2, it takes up to $k_{imp} \leq 4$ number of steps to come to an improving direction. Subsequently, by Lemma 3, the system continues to move in the same improving direction until it reaches T_2 at the point A_0. It is to be noted that the location of the point after initial correction dictates which region the trajectory aims to, namely:

- T_1: If $\left|x_{1,k_{imp}}\right| > \left|x_{2,k_{imp}}\right|$.
- T_2: If $\left|x_{1,k_{imp}}\right| < \left|x_{2,k_{imp}}\right|$.
- If $\left|x_{1,k_{imp}}\right| = \left|x_{2,k_{imp}}\right|$, then the trajectory points toward the origin.

In the current example, $\left|x_{1,k_{imp}}\right| < \left|x_{2,k_{imp}}\right|$, hence the trajectory aims toward T_2. Now, when the system crosses the point A_0, the system is already in improving direction and therefore it continues in the same direction inside T_2. Thereafter, it reaches R_2 with an already improved direction. Then inside R_2 the system continues in the same improving direction (Lemma 3). On the path inside R_2, it transits from the fourth quadrant to the third quadrant. So the control direction becomes $[\text{sgn}(x_1), -\text{sgn}(x_2)]^T$ which is also a valid improving control direction in R_2. Hence there is no change in direction of evolution. Then it continues till it reaches T_2 at point A_1. Here, it takes $k_{imp} = 4$ steps to correct itself to the new improving direction. From Fig. 6, this correction can be considered as *reflection*. Henceforth, the system follows the same pattern to reach A_2, A_3, ... and so on till it reaches the ϵ neighborhood of the origin. Since all inequalities hold for $\|\mathbf{x}\|_1 > \epsilon$, we now proceed to bound the number of steps.

Calculation of path length and number of steps: First, we will calculate the length of the path from S to the ϵ-neighborhood of the origin, say l_{total}, which satisfies

$$l_{total} = \text{len}(SA_1) + \sum_{j=1}^{n_{ref}} \text{len}(A_j A_{j+1}),$$

where $\text{len}(\cdot)$ represents the length and n_{ref} is the number of reflections before the system reaches the ϵ-neighborhood,

$$\text{len}\big(A_{n_{ref}} O\big) > \epsilon > \text{len}\big(A_{n_{ref}+1} O\big). \tag{23}$$

From Fig. 6, it follows that

$$\sum_{j=1}^{n_{ref}} \text{len}(A_j A_{j+1}) < \text{len}(A_1 S_1) + \text{len}(S_1 O).$$

So the total path length can be bounded as

$$l_{total} \leq \text{len}(SS_1) + \text{len}(S_1 O).$$

From Fig. 6, the path length of SS_1 and S_1O are the distances of the point S from the lines $x_1 = x_2$ and $x_1 = -x_2$ respectively. This results in

$$l_{total} \leq \frac{|x_{1,0} + x_{2,0}|}{\sqrt{2}} + \frac{|x_{1,0} - x_{2,0}|}{\sqrt{2}} \leq \sqrt{2}\|\mathbf{x}_0\|_1. \tag{24}$$

It is to be noted that in the beginning and at each reflection there are up to 4 steps for correction. Hence the total number

of steps would be

$$n_{total} \leq \frac{l_{total}}{\delta} + 4\left(n_{ref} + 1\right).$$

Using similar triangle geometry and eq. (24) it can be shown that

$$n_{total} \leq \frac{\sqrt{2}\|\mathbf{x}_0\|_1}{\delta} + 4\left(\left\lceil \frac{\ln(\frac{2\|\mathbf{x}_0\|_1}{3\delta})}{\ln(\frac{a^2+b^2}{a^2-b^2})} \right\rceil + 1\right) = K. \quad (25)$$

Hence, we have shown the existence of a finite upper bound depending on the initial state $\mathbf{x}_0$ and problem parameters. ■

IV. Simulation results

In this section, we report the results of a simulation study of the proposed approach, along with a comparison with the benchmark approach: *extremum-seeking (ES) control algorithm* [12]. We first discuss a two-variable discrete-time version of ES algorithm that is implemented for comparison. The block diagram in Fig. 7 illustrates the details of the implementation. Starting from **step 1**, the perturbation signals $A\sin(\omega_p k)$ and $A\cos(\omega_p k)$ are added to each of the current mean (state without perturbation) of the states: $\hat{x}_{1,k}$ and $\hat{x}_{2,k}$. The mean is initialized with the given initial condition $[x_{1,0}, x_{2,0}]^T$. This mean corresponds to the present configuration of the agents which is not known to them. In **step 2** the resulting system outputs, generated by measurement model applied on the perturbed states, are passed through high-pass filters (HPF). In **step 3** each of the filtered output is multiplied by the corresponding perturbation signals to generate the biases $\xi_{1,k}$ and $\xi_{2,k}$, which are then used to get control terms $u_{1,k}$ and $u_{1,k}$ to move the mean of the states, completing a feedback loop of the system.

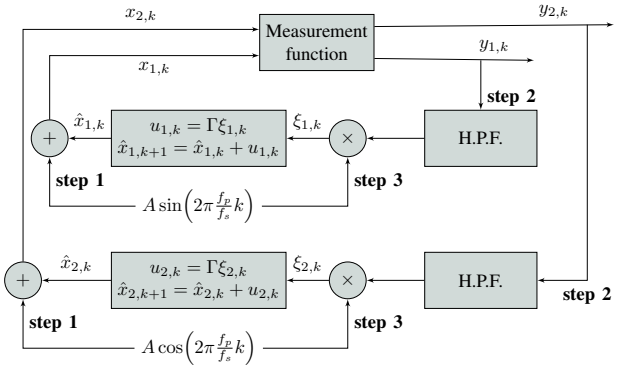

Fig. 7: Block diagram for extremum seeking control.

In our simulations, we considered relative motion between the agents R_1 and R_2. This relative motion emulates the unwanted disturbances or the movement of robots due to higher level tasks. Here, instead of the origin, the LOS orientation is at $\mathbf{x}^* = [x_1^*, x_2^*]^T$ such that

$$[\theta_1, \theta_2]^T = \mathbf{x} - \mathbf{x}^*,$$

where the new state $\mathbf{x}$ is defined in absolute sense (i.e., pointing angles of each agent with respect to the global x axis) and the LOS orientation circles the origin with a radius of r and f_M cycles per second, leading to

$$\begin{bmatrix} x_{1,k}^* \\ x_{2,k}^* \end{bmatrix} = \begin{bmatrix} r\cos(2\pi f_M k/f_s) \\ r\cos(2\pi f_M k/f_s) \end{bmatrix}.$$

Table I lists the parameters used in the simulation. For a fair comparison between the proposed algorithm and the extremum-seeking (ES) algorithm, we choose the same sampling frequency f_s for each of them and the value of the step-size δ of the proposed approach is chosen to be same as the value of perturbation amplitude A in the ES algorithm.

TABLE I: Parameters used in simulation.

Parameter	Value	Description
a	30°	Gaussian width for function g
b	20°	Gaussian width for function h
f_s	100 Hz	Sampling frequency
f_p	25 Hz	Perturbation frequency $(= \omega_p f_s)$ for ES
A	1°	Perturbation amplitude
Γ	500	Controller gain for ES
r	5°	Amplitude of circular motion for time varying optimum case
δ	1°	Step size for the algorithm

Fig. 8 shows the path of the trajectories of the states for a simulation run. The states start from the same initial condition and track the moving LOS direction. Fig. 9 illustrates the evolution of the states and the output for each of the robot corresponding to the simulation run. From Fig. 9, it can be observed that the states corresponding to the proposed approach converge to the trajectory of the optimum significantly faster than the states of ES approach. Furthermore, there is a small lag in the tracking for the ES approach. This delay in tracking for ES algorithm is attributed to the large amount of time spent in the collection of samples. The outputs of each agent using both approaches eventually converge close to the maximum value, which is 1 V. It is important to note that both of the aforementioned algorithms have steady state-oscillations, which are essential to track a moving optimum point.

Additionally, to test the repeatability and characterize the performance and limitations of the algorithms, we simulate the algorithms over a range of speeds of motion on a logarithmic scale; $f_M \in [0.001, 10]$ Hz. To illustrate the average tracking performance, we consider a metric called *mean steady-state error* e_m, which is defined as

$$e_m = \frac{1}{100}\sum_{k=n_f-99}^{n_f} \|\mathbf{x}_k - \mathbf{x}_k^*\|_2, \quad (26)$$

where n_f is the total number of iterations in one simulation run of an algorithm and $\|\cdot\|_2$ denotes the Euclidean norm of a vector. The number 100 is considered to effectively capture the average of steady-state points. Next, for each

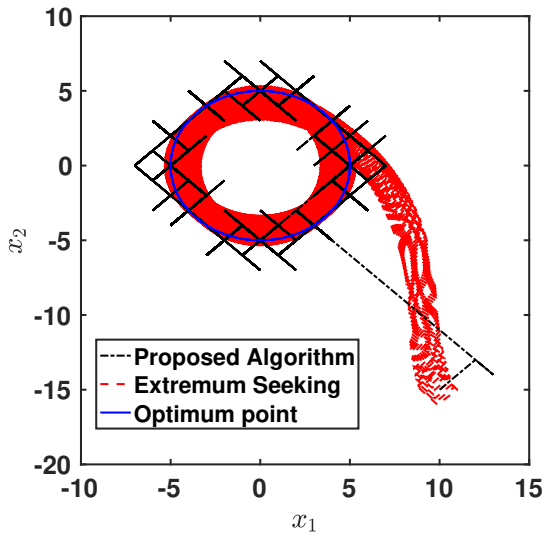

Fig. 8: Illustration of path of states tracking the moving LOS orientation that is circling the origin with frequency $f_M = 0.05$ Hz.

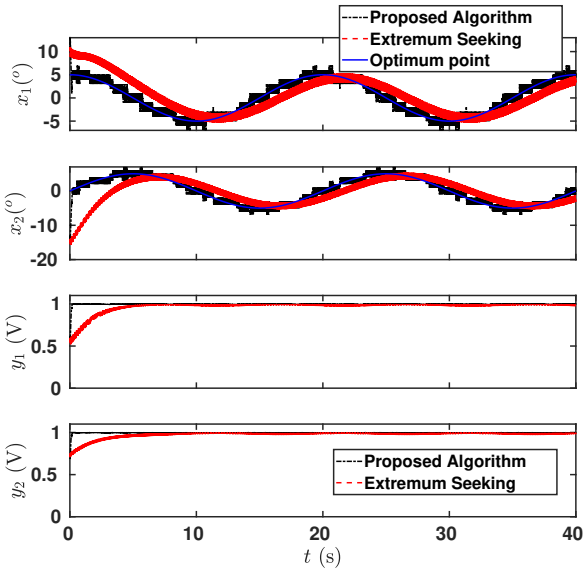

Fig. 9: Illustration of evolution of states and corresponding output for the simulation run corresponding to Fig. 8. One can observe that the ES approach trajectory has a small lag in tracking the optimum point.

of the considered frequency of motion, we generate 1000 initial points from the set ($\mathbf{x}_0 \in (-30^o, 30^o) \times (-30^o, 30^o)$) using the Latin Hypercube Sampling (LHS) technique in Matlab, and perform a simulation run for each of these initial conditions. The average (say r_{ave}) and the standard deviation of e_m from all of these runs yield the cumulative tracking performance at the particular frequency of motion. Fig. 10 shows the average tracking performance with the error bars over a range of speeds of motion. Fig. 10 also shows that the proposed algorithm is more effective in tracking higher-frequency moving optimum, maintaining a small error of atmost ϵ in magnitude up to 1 Hz. In comparison, the ES algorithm is only able to achieve the small tracking error for up to a frequency of 0.02 Hz. This *indicates that the proposed approach converges more than an order of magnitude faster than the ES algorithm.*

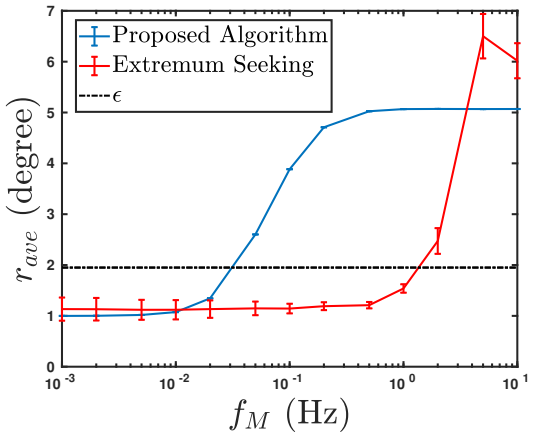

Fig. 10: Average mean steady-state errors in tracking a moving optimum over a range of f_M by each of the two algorithms. The error bars represent $\pm$standard deviation in the mean steady state error.

V. Conclusion and Future Work

In this work, we formulated a bidirectional optical alignment control problem as an optimization problem for a discrete-time dynamical system, where the origin is the optimal point for the reward function for each of the agents. For the case when the reward functions are a product of Gaussians, we proposed an output feedback control law following the constraint that the control command of an agent can depend only on the information accessible to that agent. Through rigorous analysis, we showed that for any initial condition, the proposed control law drives the system to an ϵ-neighborhood of the LOS in a finite number of steps. The efficacy of the proposed algorithm was further verified in simulation in a moving optimum scenario where the algorithm shows an order of magnitude faster convergence in comparison to the extremum-seeking control approach.

For future work, we plan to analyze a more general scenario wherein the measurement functions are not necessarily product of Gaussians. Moreover, we plan to verify the algorithm on an experimental setup involving two robots. Moving forward, we will extend the approach to a 3D scenario with moving robots, where each robot controls its own azimuthal and elevation angles to achieve the LOS [6].

Appendix

A. Proof of Lemma 2

Consider, at any iteration k,

$$\operatorname{sgn}(z_{1,k+1} - z_{1,k}) \overset{(3),(21)}{=} \operatorname{sgn}\left(\nabla z_{1,k} \cdot \mathbf{u}_k + \delta \frac{\epsilon}{3a^2}\right) \tag{27}$$

When $\left|\left|\frac{x_1}{a^2}\right| - \left|\frac{x_2}{b^2}\right|\right| > \frac{\epsilon}{3a^2}$, then from the properties of absolute values,

$$\left|\left|\frac{x_1 u_1}{a^2}\right| - \left|\frac{x_2 u_2}{b^2}\right|\right| > \frac{\delta\epsilon}{3a^2} \Rightarrow \left|\left|\frac{x_1 u_1}{a^2} + \frac{x_2 u_2}{b^2}\right|\right| > \frac{\delta\epsilon}{3a^2}$$

$$\Rightarrow \operatorname{sgn}\left(\nabla z_{1,k} \cdot \mathbf{u}_k + \frac{\delta\epsilon}{3a^2}\right) \overset{(18)}{=} \operatorname{sgn}(\nabla z_{1,k} \cdot \mathbf{u}_k) \tag{28}$$

Similarly, when $\left|\left|\frac{x_2}{a^2}\right| - \left|\frac{x_1}{b^2}\right|\right| > \frac{\epsilon}{3a^2}$, we get

$$\operatorname{sgn}(z_{2,k+1} - z_{2,k}) \overset{(18)}{=} \operatorname{sgn}(\nabla z_{2,k} \cdot \mathbf{u}_k). \tag{29}$$

For Region R_1:

$$u_{1,1} = u_{1,0} \operatorname{sgn}(z_{1,1} - z_{1,0}), \overset{(21,3)}{=} -\delta \operatorname{sgn}(x_{1,0})$$

using (10) it implies that $\mathbf{u}_1$ is improving direction. Following similar steps **for Region** R_2, $u_{2,1} = \operatorname{sgn}(u_{1,0}, u_{2,0}, x_{1,0})$, and using (14) $\mathbf{u}_1$ is improving direction, and hence $k_{imp} = 1$ for both of the regions. **For Region** M, using eqs. (9), (28), (27), we get

$$\mathbf{u}_1 = \begin{bmatrix} -\delta \operatorname{sgn}\left(\nabla z_{1,0} \cdot \mathbf{u}_0 u_{1,0}\right) \\ -\delta \operatorname{sgn}\left(\nabla z_{2,0} \cdot \mathbf{u}_0 u_{2,0}\right) \end{bmatrix} = \begin{bmatrix} -\delta \operatorname{sgn}(x_{2,0} u_{1,0} u_{2,0}) \\ -\delta \operatorname{sgn}(x_{1,0} u_{1,0} u_{2,0}) \end{bmatrix},$$

$$\mathbf{u}_2 = \begin{bmatrix} -\operatorname{sgn}(z_{1,2} - z_{1,1}) u_{1,1} \\ -\operatorname{sgn}(z_{2,2} - z_{2,1}) u_{2,1} \end{bmatrix} \overset{(3,28,15,16)}{=} \begin{bmatrix} -\delta \operatorname{sgn}(x_{1,0}) \\ -\delta \operatorname{sgn}(x_{2,0}) \end{bmatrix},$$

which essentially means that $\mathbf{u}_2$ is improving direction and $k_{imp} = 2$ for M. It is to be noted that the above expression of $\mathbf{u}_2$ is also an improving control direction for any point in any of the regions of interest.

Now consider the **transition region** T_1 where,

$$\left|\left|\frac{x_{1,0}}{a^2}\right| - \left|\frac{x_{2,0}}{b^2}\right|\right| < \frac{\epsilon}{3a^2} \text{ and } \left|\left|\frac{x_{2,0}}{a^2}\right| - \left|\frac{x_{1,0}}{b^2}\right|\right| > \frac{\epsilon}{3a^2} \tag{30}$$

From eq. (27), (29)

$$\begin{bmatrix} u_{1,1} \\ u_{2,1} \end{bmatrix} = -\begin{bmatrix} \operatorname{sgn}\left(\frac{x_{1,0}u_{1,0}}{a^2} + \frac{x_{2,0}u_{2,0}}{b^2} + \delta\frac{\epsilon}{3a^2}\right) \\ \operatorname{sgn}\left(\frac{x_{1,0}u_{1,0}}{b^2} + \frac{x_{2,0}u_{2,0}}{a^2} + \delta\frac{\epsilon}{3a^2}\right) \end{bmatrix}. \tag{31}$$

Here in this region, the evolution of next few states is captured in two separate cases depending on the initial condition. The two cases are as follows:

- **Case-I:** $\operatorname{sgn}(x_{1,0}u_{1,0}) = \operatorname{sgn}(x_{2,0}u_{2,0})$.
  Here eqs. (15), (16), (31) result in $\mathbf{u}_1 = [-\delta \operatorname{sgn}(x_{1,0}),\ -\delta \operatorname{sgn}(x_{2,0})]^T \Rightarrow k_{imp} = 1$.
- **Case-II:** $\operatorname{sgn}(x_{1,0}u_{1,0}) = -\operatorname{sgn}(x_{2,0}u_{2,0})$,
  Here, eq. (30) and eq. (31) result in

  $$\mathbf{u}_1 = -[u_{1,0},\ \delta \operatorname{sgn}(x_{2,0})]^T. \tag{32}$$

  Now consider $u_{1,2}$ given by,

  $$u_{1,2} = -\operatorname{sgn}\left(\frac{x_{1,1}u_{1,1}}{a^2} + \frac{x_{2,1}u_{2,1}}{b^2} + \frac{\epsilon}{3a^2}\right)u_{1,1}. \tag{33}$$

  There arise two further sub-cases:
  - **Case-II(a):** $\operatorname{sgn}(x_{1,0}) = \operatorname{sgn}(u_{1,0})$, using eqs. (30), (33) we get $u_{1,2} - \delta \operatorname{sgn}(x_{1,0})$. (34)
  - **Case-II(b):** $\operatorname{sgn}(x_{1,0}) = -\operatorname{sgn}(u_{1,0})$, using eqs. (15), (33) we get $u_{1,2} - \delta \operatorname{sgn}(x_{1,0})$. (35)

  Hence, $u_{1,2} = -\delta \operatorname{sgn}(x_{1,0})$ in both cases (a) and (b). Now considering the component $u_{2,2}$,

  $$u_{2,2} = -\operatorname{sgn}\left(\frac{x_{1,1}u_{1,1}}{b^2} + \frac{x_{2,1}u_{2,1}}{a^2} + \frac{\delta\epsilon}{3a^2}\right)u_{2,1} \tag{36}$$

  Similar to the cases II(a) and (b), two sub-cases arise for the second component of $\mathbf{u}_2$:
  - **Case II(c):** $\operatorname{sgn}(x_{2,0}) = \operatorname{sgn}(u_{2,0})$, using eqs. (16) and (36), we get $u_{2,2} = -\delta \operatorname{sgn}(x_{2,0})$. Here, combining eqs. (34) and (35), $\mathbf{u}_2$ is an improving direction.
  - **Case II(d):** $\operatorname{sgn}(x_{2,0}) = -\operatorname{sgn}(u_{2,0})$, using eqs. (30) and (36), we get $u_{2,2} = \delta \operatorname{sgn}(x_{2,0})$ In this case $\mathbf{u}_2$ is not an improving direction. However, if we consider $\mathbf{x}_2$, with $u_{2,0} : \delta \operatorname{sgn}(u_{2,0}) = -\delta \operatorname{sgn}(x_{2,0})$ along with eq. (32), we obtain

    $$\mathbf{x}_2 = [x_{1,0} + u_{1,0} + u_{1,1},\ x_{2,0} + u_{2,0} + u_{2,1}]^T = \mathbf{x}_0$$

    Hence the system comes back to its initial state $\mathbf{x}_0$ with control direction:

    $$\mathbf{u}_2 = [-\delta \operatorname{sgn}(x_{1,0}),\ \delta \operatorname{sgn}(x_{2,0})]^T$$

    where it can be verified that $\operatorname{sgn}(x_{1,2}u_{1,2}) = \operatorname{sgn}(x_{2,2}u_{2,2})$ and $\operatorname{sgn}(u_{2,2}) = \operatorname{sgn}(x_{2,2})$. These are the exact conditions for case II (c). Consequently, by our previous arguments, $\mathbf{u}_4$ becomes improving direction with respect to $\mathbf{x}_0$ ($k_{imp} = 4$).

Thus, in all cases of initial control term $\mathbf{u}_0$ for a point $\mathbf{x}_0 \in T_1$, eq. (22) holds true with $k_{imp} \leq 4$. By symmetry, **transition region** T_2 can be analyzed using arguments similar to those for T_1. Hence, we have shown that from any initial point $\mathbf{x}_0$ in the defined regions, the control direction becomes improving with respect to the initial point after a finite number of steps $k_{imp} \leq 4$. ■

REFERENCES

[1] D. Anguita, D. Brizzolara, and G. Parodi, "Building an underwater wireless sensor network based on optical ommunication: Research challenges and current results," in *Sensor Technologies and Applications, 2009. SENSORCOMM'09. Third International Conference on*. IEEE, 2009, pp. 476–479.

[2] I. Rust and H. Asada, "A dual-use visible light approach to integrated communication and localization of underwater robots with application to non-destructive nuclear reactor inspection," *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pp. 2445–2450, May 2012.

[3] G. Soysal and M. Efe, "Kalman filter aided cooperative optical beam tracking," *Radioengineering*, pp. 242–248, June 2010.

[4] S. Jeon and H. Toshiyoshi, "MEMS tracking mirror system for a bidirectional free-space optical link," *Applied Optics*, vol. 56, no. 24, p. 6720, 2017.

[5] P. B. Solanki, M. Al-Rubaiai, and X. Tan, "Extended Kalman filter-based active alignment control for LED optical communication," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 4, pp. 1501–1511, Aug 2018.

[6] P. B. Solanki and X. Tan, "Extended Kalman filter-based 3D active-alignment control for LED communication," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, May 2018, pp. 1–8.

[7] P. Solanki, S. Bopardikar, and X. Tan, "Active alignment control-based led communication for underwater robots," 2020, manuscript submitted to International Conference on Intelligent Robots and Systems.

[8] N. Li and J. R. Marden, "Designing games for distributed optimization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 230–242, April 2013.

[9] J. S. Shamma and G. Arslan, "Dynamic fictitious play, dynamic gradient play, and distributed convergence to nash equilibria," *IEEE Transactions on Automatic Control*, vol. 50, no. 3, pp. 312–327, March 2005.

[10] N. Chopra and M. W. Spong, *Passivity-Based Control of Multi-Agent Systems*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 107–134. [Online]. Available: https://doi.org/10.1007/978-3-540-37347-6˙6

[11] E. Biyik and M. Arcak, "Gradient climbing in formation via extremum seeking and passivity-based coordination rules," in *2007 46th IEEE Conference on Decision and Control*, Dec 2007, pp. 3133–3138.

[12] M. Krstic and K. B. Ariyur, *Multiparameter Extremum Seeking*. John Wiley Sons, Ltd, 2004, ch. 2, pp. 21–45. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/0471669784.ch2

[13] M. Doniec, M. Angermann, and D. Rus, "An end-to-end signal strength model for underwater optical communications," *IEEE Journal of Oceanic Engineering*, vol. 38, no. 4, pp. 743–757, Oct 2013.