Multivariable Extremum Seeking Controllers for Multi-Beam Steering using Reconfigurable Metasurfaces

Abdullah Bin Masood*†, Vasos Vassiliou*†, Andreas Pitsillides*‡, Christos Liaskos§, and Marios Lestas¶

*Department of Computer Science, University of Cyprus, Nicosia, Cyprus

†CYENS-Centre of Excellence, Nicosia, Cyprus

‡Department of Electrical and Electronic Engineering Science, University of Johannesburg, South Africa (Visiting Professor)

§Department of Computer Science and Engineering, University of Ioannina, Ioannina, Greece

¶Department of Electrical Engineering, Frederick University, Nicosia, Cyprus

(masood.abdullah, vassiliou.v, andreas.pitsillides)@ucy.ac.cy, cliaskos@uoi.gr, eng.lm@frederick.ac.cy

***Abstract*—Intelligent Reflecting Metasurfaces constitute a revolutionary technology that can alleviate the blockage problem in mm-Wave communications. In this paper, we consider a metasurface coding that can realize beam splitting of an incident beam in multiple directions and we address the problem of feedback-based adaptive reconfiguration of the metasurface controller states so that the reflected beams are guided towards two intended receivers. As the receiver locations are unknown and to counter system uncertainties, we employ feedback to maximize the received power at both receivers, aiming for standalone operation of the metasurface. Changes in both the elevation and azimuth angles of the transmitted beams are considered, thus necessitating a design beyond standard extremum-seeking controllers, which have proved to be ineffective. Our approach involves online parameter identification techniques for gradient estimation, coupled with the method of steepest ascent. The effectiveness of the method in guiding the beams towards the reference values is shown through simulations.**

I. INTRODUCTION

Intelligent Reflective Surfaces (IRS)[1] is by now a well-established technology expected to be part of 6G wireless networks, due to their ability to programmatically alter the wireless channel, which up to now was a given unpredictable entity [1]. Metasurfaces constitute the most promising technology for the realization of IRSs due to their ability to exert precise control over the impinging electromagnetic waves, relative to competing technologies, such as phased arrays, due to the subwavelength sizing of the unit cells [2]. In the last few years, there has been a surge of interest in the design and analysis of metasurfaces, especially in the communications domain. Excellent recent reviews include [3]–[5]. As correctly pointed out in [6], although programmable metasurfaces have been a hot topic, the development of their real-time implementations is rarely touched. This is relevant to beam management and beamforming algorithms, which usually employ single-bit feedback and refined scanning of the parameter space in search of the optimal configuration based on the chosen cost metric. However, the need for finer training to achieve more accurate beam steering has been identified in previous literature [7] paving the way for the use of multi-bit feedback. This has been done, specifically for the metasurface paradigm in [8], [9], however, therein the thrust has been different, not considering the iterative calculation of the feedback signal. The concept of extremum-seeking control has been employed in [10] for a single user being served by the metasurface. The approach has been shown to be successful. However, its multivariable extensions for the simultaneous control of both the azimuth and elevation angles were observed to be problematic. Classical extremum seeking strategies [11], [12] are known to implement on average, a gradient estimation algorithm, with the obtained gradient estimate subsequently used to direct the parameters towards values that converge to the maximum of the considered cost function. Taking advantage of this observation, we consider the direct calculation of the gradient estimate, which is then utilized in the steepest ascent algorithm to maximize the cost function. This approach and the associated stability proofs are scalable to the multivariate case. The main contributions of the paper are as follows:

1) Unlike previous work, we formulate the metasurface-based beam steering problem in the multi-user case.
2) We utilize a novel algorithm that combines online parameter identification techniques and the steepest ascent methods in the multivariable case.
3) We show convergence of the algorithms in the vicinity of the actual gradient.
4) We use simulations that account for the metasurface coding procedure to show the effectiveness of the proposed algorithms.

The paper is organized as follows. The structure of the hypersurface system and the metasurface coding procedure are given in Section II. The problem is formulated in Section III, followed by the design and performance evaluation of the proposed control scheme in Sections IV and V, respectively. Section VI concludes the research paper.

II. HYPERSURFACE SYSTEM STRUCTURE

A. Hypersurface Structure

The Hypersurface (HSF) technology is considered in this work to realize the IRS as it can provide autonomous operation and programmability by complementing the Software-Defined Metasurfaces (SDM) with an embedded network of

[1]IRSs are also known as Reconfigurable Intelligent Surfaces (RISs)

978-1-6654-5975- 4/22 © 2022 IEEE

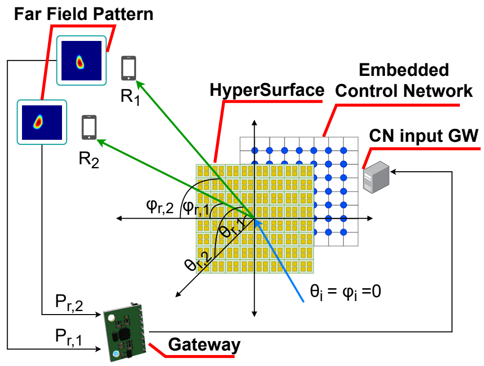

Fig. 1. The system's physical layout.

controllers. The HSF architecture is layered in structure, including the electromagnetic (EM) layer, the embedded control (ECN) layer, and the gateway (GW) layer [13]. The EM layer, comprising of a collection (usually periodic) of unit cells referred to as the *meta-atoms*, deals with the physical realization of the SDM. It allows the metasurface, through tunable actuators, such as varactors *C* and varistors *R*, to shape the properties of the impinging wave at the request of corresponding Application Programming Interface (API) calls. The ECN layer then defines the necessary protocols and hardware to facilitate the information exchange between the GW and each of the tunable elements altering the corresponding states [14]. Each distinct controller chip referred to as CN can assume a finite number of M_s states where each state is a composite of RC values. The GW layer, operating on top of the ECN layer, is responsible to connect the HSF via standardized protocols to the external world and converting the incoming software instructions into specific *metasurface codes*. The metasurface code is the set of unit cell states or *RC* values, out of the M_s available ones, by which the metasurface achieves the desired functionality with minimum error.

B. Metasurface Coding

Anomalous reflection of the impinging wavefront on the metasurface is achieved by applying a linear phase gradient Φ in the desired x and y directions such as $\Phi_x = \frac{\partial \Phi}{\partial x}$ and $\Phi_y = \frac{\partial \Phi}{\partial y}$ [15]. The relationship between the implemented phase gradients, the angle of incidence ($\{\theta_i, \varphi_i\}$ in polar coordinates), and the target reflection angle ($\{\theta_r, \varphi_r\}$) is obtained by applying the momentum conservation law for wave vectors [10]. In general, the air is the host medium for most communication applications, therefore, in this study, we consider wavelengths λ_i and λ_r for the incident and the reflected mediums, respectively, to be equal, i.e., $\lambda_i = \lambda_r = \lambda_0$. By fixing $\theta_i = \varphi_i = 0$ without loss of generality, the phase Φ_{mn} of a square unit cell, having S_u lateral size, located at the m^{th} column and n^{th} row can be expressed, by applying such gradients, as:

$$\Phi_{mn} = \frac{2\pi S_u(m \sin\theta_r \cos\varphi_r + n \sin\theta_r \sin\varphi_r)}{\lambda_0} \quad (1)$$

C. Far-Field Evaluation

In this paper, we characterize the received power at each of the receivers using a description of the radiation pattern of the reflected beam(s) resulting from the assumed planar impinging wave, as a function of the states of each of the unit cells. The model used to map the impinging wave and the controller states to the reflected radiation pattern has been shown to be accurate for beam steering applications by comparing the results with full-wave simulations [16]. The far-field pattern of the metasurface is obtained as the sum of the contributions of all unit cells by applying Huygens' principle, assuming negligible cross-talk between adjacent unit cells. Considering K numbers of receivers for the multi-beam steering case, the scattering field of the metasurface is expressed as follows [16]:

$$E(\theta, \varphi) = \sum_{m=1}^{M} \sum_{n=1}^{N} e^{jk_0 \zeta_{mn}(\theta,\varphi)} \sum_{k=1}^{K} e^{j\Phi_{mn}(\theta_{r,k}, \varphi_{r,k})} \quad (2)$$

where $\zeta_{mn}(\theta, \varphi)$ is the relative phase shift of the unit cell with respect to the radiation pattern coordinates (θ, φ), given by:

$$\zeta_{mn}(\theta, \varphi) = S_u \sin\theta[(m - \frac{1}{2})\cos\varphi + (n - \frac{1}{2})\sin\varphi] \quad (3)$$

III. Problem Formulation

The considered multi-beam steering scenario, as shown in Fig. 1, involves an HSF acting as the IRS, a single transmitter, and multiple receivers R_k (where $k = [1, 2, ..., K]$ and $K \in \mathbb{Z}^+$). The receiver R_k (either stationary or mobile) at any time t located at a specific position $l_k(t)$, in 3 dimensional (3D) space, measures the received power $P_{r,k}(t)$. Due to the absence of LOS, the transmitter attempts to establish a reliable communication channel with each of the receivers R_k via the HSF. The impinging planar wave is characterized by the angle of incidence ψ_i relative to the HSF, which, due to the 3D nature of the considered setting requires two scalar components for its description, the elevation angle θ_i and the azimuth angle φ_i. We stack both components into a vector $\psi_i = [\theta_i, \varphi_i]$ and without loss of generality we assume them to be 0. A planar wave can result from a horn antenna at the transmitter. As we aim for beam splitting to be realized on the metasurface, we consider multiple reflected beams each of which is characterized by the reflection angle $\psi_{r,k} = [\theta_{r,k}, \varphi_{r,k}]$, which is similar to the incidence beam case comprises of two scalar components. The reflection angle $\psi_{r,k}$ is dictated by the assumed state of each of the controllers c_i. We stack these into the vector $c = [c_1, c_2, ..., c_M]$. As the controller states can be changed at will by issuing appropriate command signals via the GW's API, the vector c may be considered as the control signal. The objective is to iteratively reconfigure the controller states of the HSF to appropriately shape the radiation pattern so that optimal channel quality is attained at the receivers. In this study, the channel quality is assessed merely in terms of the received power $P_{r,k}(t)$, which is thus required to be maximized.

The received power $P_{r,k}(t)$ is considered as a function of the transmitter power P_t, the incident angle ψ_i, the position

$l_k(t)$ of R_k, and the chosen c, which using the metasurface coding procedure of (1) yields the desired far-field pattern. This relationship can be expressed via a function $T_k(.)$ such that $P_{r,k}(t) = T_k(P_t, \psi_i, c, l_k(t))$. As an example of such a function, we may consider the link budget formula:

$$P_{r,k}(t) = \frac{P_t G_t G_r}{P_L L_o} \tag{4}$$

where G_t and G_r are the gains at the transmitter and receiver antennas, respectively, P_L is the path loss value that depends on the considered scenario, and L_o are the stray losses incurred during transmitter and receiver feed. The G_r at the metasurface is assumed to be 1 and the G_t of the metasurface is assumed to be proportional to the normalized radiation pattern of the incident beam. Thus, the received power $P_{r,k}(t)$ at the receiver R_k is assumed proportional to the radiation pattern of the reflected beam k, which is characterized using (2). Based on the function $T_k(.)$, one may consider c as the control variable, posing the objective of identifying the optimal c such that the $P_{r,k}(t)$ is maximized. It can be cast as a multi-input single-output (MISO) control problem, which, however, is difficult and time-consuming to solve due to the multi-dimensionality of the input space and also due to the highly nonlinear and uncertain nature of the function $T_k(.)$. However, as indicated in the previous discussion, a desired $\psi_{r,k}$ for the k^{th} reflected beam can be achieved by appropriately choosing the controller states. This can be expressed via a function C_k such that $\psi_{r,k} = C_k(c, \psi_i)$. The function $C_k(.)$ can be best described in tabular form as it is highly nonlinear.

Thus, the multi-dimensionality problem can be overcome, if we account for the function $C_k(.)$ in $T_k(.)$ and consider $\psi_{r,k}$ as the scalar control variable instead of c. This can be expressed through a composite function $W_k(.)$ such that $P_{r,k}(t) = W_k(P_t, \psi_i, l_k(t), \psi_{r,k})$. The resulting change in the input control variable also ensures the concavity of $P_{r,k}(t)$ with respect to the input variable $\psi_{r,k}$. So, the optimization problem becomes:

$$P_{r,k} : \max_{\psi_{r,k} \in [-90,90]} P_{r,k}(t) \tag{5}$$

The difficulty in solving the (5) stems from the following challenges: there exists only a single point of measurement corresponding to each of the receivers R_k, the presence of random external disturbances at different points in the considered system, the ψ_i may be unknown to the R_k and the HSF, the uncertainties in the functions $C_k(.)$ and $T_k(.)$ due to modeling errors and manufacturing-related inaccuracies, and the possibility of getting run-time errors and encountering hardware faults which can lead to abnormal behavior of both the CN and the metasurface [10].

These aforementioned challenges motivate us to consider a closed-loop implementation for the multi-beam steering problem rather than an open-loop implementation [10]. We assume the existence of a control channel for closed-loop implementation, which allows each receiver R_k to communicate the measured power $P_{r,k}$ to the input GW of the HSF. The input GW then implements a control algorithm that iteratively calculates the reflection angles $\psi_{r,k}$. The control algorithm can be expressed as:

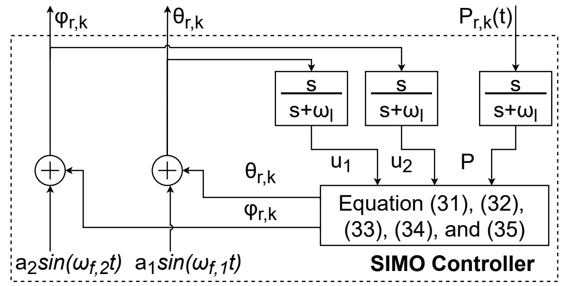

Fig. 2. Schematic diagram of the proposed SIMO controller for beam steering.

$$\begin{cases} \dot{y}_k = H_k(y_k, P_{r,k}(t)) \\ \psi_{r,k} = V_k(y_k, P_{r,k}(t)) \end{cases} \tag{6}$$

where $P_{r,k}(t)$ is the measured signal strength, y_k is a vector of controller states, $H_k(.)$ and $V_k(.)$ are possibly nonlinear functions, and $\psi_{r,k}(t)$ is the desired angle of reflection. The problem is then to design the functions $H_k(.)$ and $V_k(.)$ such that problem (5) is solved.

IV. Proposed Control Algorithm

In this section, for ease of exposition, we first present the control algorithm and its properties for the single variable case. We then present the multivariable version of the algorithm.

A. Single-Input Single-Output (SISO) Controller

Consider a continuous concave function as follows:

$$P_{r,k} = f(\theta_{r,k}) \tag{7}$$

where $\theta_{r,k}$ and $P_{r,k} \in \mathbb{R}$ are the scalar input and output of the system respectively. We assume that the function $f(\theta_{r,k})$ is smoothly differentiable and has a single extremum (maximum) $P^*_{r,k}$ at $\theta_{r,k} = \theta^*_{r,k}$. The control objective is to design an iterative algorithm that converges to the unknown extremum value $\theta^*_{r,k}$ which maximizes $P_{r,k}$.

1) Control Algorithm: We add a sinusoidal perturbation signal with an amplitude a and frequency ω_f in the system (7) about a nominal value θ_{r,k_0} such that:

$$P_{r,k}(t) = f(\theta_{r,k_0} + a\sin(\omega_f t)) \tag{8}$$

The derivative of the (8) with respect to time is given by:

$$\dot{P}_{r,k} = d_0\dot{\theta}_{r,k} + d_1\dot{\theta}_{r,k} \tag{9}$$

where $\frac{df}{d\theta_{r,k}} = d_0 + d_1(t)$, $d_0 = \left.\frac{df}{d\theta_{r,k}}\right|_{\theta_{r,k}=\theta_{r,k_0}}$, and $\dot{\theta}_{r,k} = \frac{d\theta_{r,k}}{dt} = a\omega_f\cos(\omega_f t)$. Here, $|\dot{\theta}_{r,k}| \leq a\omega_f$ and $|d_1(t)| \leq k(a)$. We then apply a first order filter in (9) such that:

$$\frac{1}{s+\omega_l}[\dot{P}_{r,k}] = d_0\frac{1}{s+\omega_l}[\dot{\theta}_{r,k}] + \frac{1}{s+\omega_l}[d_1\dot{\theta}_{r,k}] \tag{10}$$

Let $P = \frac{s}{s+\omega_l}[P_{r,k}]$, $u(t) = \frac{s}{s+\omega_l}[\theta_{r,k}]$, and $d(t) = \frac{1}{s+\omega_l}[d_1\dot{\theta}]$ Here, $h_\theta = d_0$ is the true gradient, $\omega_l > 0$ is a constant design parameter, and the signals $u(t)$ and $d(t)$ are bounded, i.e., $|u(t)| \leq \frac{a\omega_f}{\omega_l}$ and $|d(t)| \leq \frac{ak(a)\omega_f}{\omega_l}$. It follows that:

$$P = h_\theta u(t) + d(t) \tag{11}$$

Using $\hat{h}_\theta$ as an estimate of h_θ, we can predict the value of P as:

$$\hat{P} = \hat{h}_\theta u(t) \tag{12}$$

The estimation error ϵ between P and $\hat{P}$ is formed as:

$$\epsilon = P - \hat{h}_\theta u(t) \tag{13}$$

The adaptive control law to generate $\hat{h}_\theta$ can be expressed as:

$$\dot{\hat{h}}_\theta = \gamma \epsilon u, \quad \hat{h}_\theta(0) = \hat{h}_{\theta_0} \tag{14}$$

where $\gamma > 0$ is a scalar parameter known as the controller gain. We can then calculate $\theta^*_{r,k}$ using the steepest ascent algorithm shown below:

$$\theta_{r,k}[n+1] = \theta_{r,k}[n] + \alpha \hat{h}_\theta \tag{15}$$

where α is the learning rate and $\hat{h}_\theta$ is the estimated gradient from (14) at n^{th} iteration.

2) Properties: The properties of the estimation algorithm are summarized in the following theorem:

Theorem 1: The control algorithm defined by (13), (14), and (15) guarantees that the estimated value of the gradient $\hat{h}_\theta$ is bounded to a small arbitrary region in the neighborhood of the true gradient h_θ and converges exponentially fast to it within finite time, with an error that depends on the value of parameters a, $k(a)$, ω_f, ω_l, γ, κ, and ν.

Proof: The dependence of ϵ on the parameter estimation error $\tilde{h}_\theta \triangleq \hat{h}_\theta - h_\theta$ is given by:

$$\epsilon = -\tilde{h}_\theta u + d \tag{16}$$

As h_θ is constant, i.e., $\dot{h}_\theta = 0$, we can rewrite (14) using (16) in terms of the parameter estimation error, i.e., $\dot{\tilde{h}}_\theta = \dot{\hat{h}}_\theta - \dot{h}_\theta$.

$$\dot{\tilde{h}}_\theta = \dot{\hat{h}}_\theta = -\gamma u^2 \tilde{h}_\theta + \gamma d u \tag{17}$$

Here, (17) represents a linear time-varying system with input similar to the system shown in (18) whose solution is given in the closed form in (19).

$$\dot{x} = \rho(t)x(t) + \eta(t)\xi(t), \quad x(0) = x_0 \tag{18}$$

$$x(t) = x_0 e^{\int_0^t \rho(\tau)d\tau} + \eta \int_0^t e^{\int_\tau^t \rho(\sigma)d\sigma} \xi(\tau) d\tau \tag{19}$$

It is well known that the exponential convergence of $x(t)$ in (19) relies on the Persistence of Excitation (PE) condition [17]:

$$\int_t^{t+T_0} \rho(\tau)d\tau \geq \kappa_0 T_0 \tag{20}$$

for some T_0, which is equivalent to:

$$e^{\int_\tau^t \rho(\tau)d\tau} \leq \kappa e^{\nu(t-\tau)} \tag{21}$$

for all $t \geq \tau$ and some $\kappa > 0$ and $\nu < 0$. If $\rho(t)$ satisfies the condition in (20), then the solution of the system in (19), using (21) and considering the boundedness of ξ (i.e., $|\xi(t)| \leq |\delta|$), can be written as:

$$x(t) \leq x_0 \kappa e^{\nu t} + \eta\delta \int_0^t \kappa e^{\nu(t-\tau)} d\tau \tag{22}$$

$$\lim_{t\to\infty} |x(t)| \leq \left| \frac{\kappa\eta\delta}{\nu} \right| \tag{23}$$

Following (23), and the aforementioned bounds on $u(t)$ and $d(t)$, the convergence properties of (17) can be summarized by:

$$\lim_{t\to\infty} \left| \tilde{h}_\theta(t) \right| \leq \left| \frac{\gamma\kappa a^2 k(a) \omega_f^2}{\nu \omega_l^2} \right| \tag{24}$$

Therefore, we conclude that the estimated value of the gradient $\hat{h}_\theta$ converges exponentially fast to a bounded region in the neighborhood of the true gradient h_θ, and this bounded region can be defined using the parameters a, $k(a)$, ω_f, ω_l, γ, κ, and ν. This completes the proof of theorem 1.

B. Single-Input Multi-Output (SIMO) Controller

Consider a general MISO non-linear function. Without loss of generality, we consider the case of two independent variables as follows:

$$P_{r,k} = f(\theta_{r,k}, \varphi_{r,k}) \tag{25}$$

where $\theta_{r,k}$, $\varphi_{r,k} \in \mathbb{R}$ are the scalar inputs and $P_{r,k} \in \mathbb{R}$ is the scalar output of the system. We assume that the function $f(\theta_{r,k}, \varphi_{r,k})$ is smoothly differentiable and has a single extremum (maximum) $P^*_{r,k}$ at $\theta_{r,k} = \theta^*_{r,k}$ and $\varphi_{r,k} = \varphi^*_{r,k}$. The control objective is to design an iterative algorithm that converges to the unknown extremum values of $\theta^*_{r,k}$ and $\varphi^*_{r,k}$ which maximize $P_{r,k}$.

1) Control Algorithm: We add two perturbation sinusoidal signals to $\theta_{r,k}$ and $\varphi_{r,k}$, respectively, about nominal values θ_{r,k_0} and φ_{r,k_0} in the system (25). The perturbation signals for each input $\theta_{r,k}$ and $\varphi_{r,k}$ have amplitude a_1 and a_2, and frequency $\omega_{f,1}$ and $\omega_{f,2}$, respectively. After introducing the perturbation sinusoidal signals, the (25) becomes:

$$P_{r,k}(t) = f(\theta_{r,k_0} + a_1 \sin(\omega_{f,1}t), \varphi_{r,k_0} + a_2 \sin(\omega_{f,2}t)) \tag{26}$$

The derivative of the (26) with respect to time is given by:

$$\dot{P}_{r,k} = d_{11}\dot{\theta}_{r,k} + d_{21}\dot{\varphi}_{r,k} + d_{12}\dot{\theta}_{r,k} + d_{22}\dot{\varphi}_{r,k} \tag{27}$$

where

$$\frac{df}{d\theta_{r,k}} = d_{11} + d_{12}(t), \qquad \frac{df}{d\varphi_{r,k}} = d_{21} + d_{22}(t),$$
$$\dot{\theta}_{r,k} = a_1\omega_{f,1}\cos(\omega_{f,1}t), \quad \dot{\varphi}_{r,k} = a_2\omega_{f,2}\cos(\omega_{f,2}t),$$

Here, $|\dot{\theta}_{r,k}| \leq a_1\omega_{f,1}$, $|\dot{\varphi}_{r,k}| \leq a_2\omega_{f,2}$, $|d_{12}| \leq k(a_1)$ and $|d_{22}| \leq k(a_2)$. After applying first-order filters, the (27) becomes:

$$\frac{1}{s+\omega_l}[\dot{P}_{r,k}] = d_{11}\frac{1}{s+\omega_l}[\dot{\theta}_{r,k}] + d_{21}\frac{1}{s+\omega_l}[\dot{\varphi}_{r,k}] + \frac{1}{s+\omega_l}[d_{12}\dot{\theta}_{r,k}] + \frac{1}{s+\omega_l}[d_{22}\dot{\varphi}_{r,k}] \tag{28}$$

Consider the following:

$$P = \frac{s}{s+\omega_l}[P_{r,k}], \qquad h_\theta = d_{11}, \qquad h_\varphi = d_{21},$$
$$u_1(t) = \frac{s}{s+\omega_l}[\theta_{r,k}], \qquad u_2(t) = \frac{s}{s+\omega_l}[\varphi_{r,k}],$$
$$d(t) = d_1(t) + d_2(t) = \frac{1}{s+\omega_l}[d_{12}\dot{\theta}_{r,k}] + \frac{1}{s+\omega_l}[d_{22}\dot{\varphi}_{r,k}]$$

Here, h_θ and h_φ are the true gradients, $\omega_l > 0$ is a constant design parameter, and the input signals $u_1(t)$ and $u_2(t)$, and the disturbance signals $d_1(t)$ and $d_2(t)$ are bounded, i.e., $|u_1(t)| \leq \frac{a_1\omega_{f,1}}{\omega_l}$, $|u_2(t)| \leq \frac{a_2\omega_{f,2}}{\omega_l}$, $|d_1(t)| \leq \frac{a_1 k(a_1)\omega_{f,1}}{\omega_l}$, and $|d_2(t)| \leq \frac{a_2 k(a_2)\omega_{f,2}}{\omega_l}$. It follows that:

$$P = h_\theta u_1(t) + h_\varphi u_2(t) + d(t) \tag{29}$$

Using $\hat{h}_\theta$ and $\hat{h}_\varphi$ as the estimates of h_θ and h_φ, respectively, we can generate the predicted or estimated value $\hat{P}$ of P as:

$$\hat{P} = \hat{h}_\theta u_1(t) + \hat{h}_\varphi u_2(t) \tag{30}$$

The estimation error ϵ is evaluated as:

$$\epsilon = P - \hat{h}_\theta u_1(t) - \hat{h}_\varphi u_2(t) \tag{31}$$

The adaptive laws to generate $\hat{h}_\theta$ and $\hat{h}_\varphi$ can be expressed as:

$$\dot{\hat{h}}_\theta = \gamma_1 \epsilon u_1(t), \quad \hat{h}_\theta(0) = \hat{h}_{\theta_0}, \tag{32}$$
$$\dot{\hat{h}}_\varphi = \gamma_2 \epsilon u_2(t), \quad \hat{h}_\varphi(0) = \hat{h}_{\varphi_0}, \tag{33}$$

where $\gamma_1 > 0$ and $\gamma_2 > 0$ are the controlling gains, respectively, that affect the convergence properties of the system and constitute design parameters. We can then apply the steepest ascent algorithm in an attempt to converge to the extremum values $\theta^*_{r,k}$ and $\varphi^*_{r,k}$ as follows:

$$\theta_{r,k}[n+1] = \theta_{r,k}[n] + \alpha_1 \hat{h}_\theta, \tag{34}$$
$$\varphi_{r,k}[n+1] = \varphi_{r,k}[n] + \alpha_2 \hat{h}_\varphi, \tag{35}$$

where α_1 and α_2 are the learning rates, and $\hat{h}_\theta$ and $\hat{h}_\varphi$ are the estimated gradients from equations (32) and (33). The overall SIMO control scheme can be summarized by equations (31), (32), (33), (34) and (35), and a block diagram representation is shown in Fig. 2. Due to the lack of space, we will present the properties of the SIMO controller in an extension of this paper.

V. Performance Evaluation

In this section, we have evaluated the performance of the proposed SIMO controller scheme by first investigating its ability to guide a single reflected beam onto a specific target and then multiple reflected beams to multiple targets. The simulation experiments are conducted on MATLAB with the system under consideration comprising three main modules: the implementation of the controller scheme, the function which maps the $\theta_{r,k}$ and $\varphi_{r,k}$ generated by the controller to the controller states c, and the module which yields the far-field pattern as a result of the selected controller states. The far-field pattern determines the received signal $P_{r,k}(t)$ strength

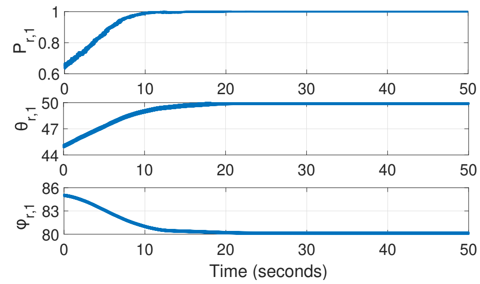

Fig. 3. Received power and the reflection angles in the case of a stationary receiver.

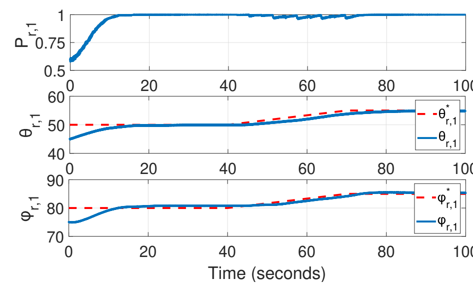

Fig. 4. Received power and reflection angles in the case of a mobile receiver.

and we have considered received power values normalized by the maximum value throughout the evaluation procedure such that $0 \leq P_{r,k}(t) \leq 1$. In all the considered scenarios, we have assumed zero-feedback delays and selected the following parameters for controller scheme: $\omega_l = 20$, $\gamma_1 = \gamma_2 = 80$, and $\alpha_1 = \alpha_2 = 0.1$, while the sampling time T_s is fixed at 0.01s for updating the estimate of $\theta_{r,k}$ and $\varphi_{r,k}$ in (34) and (35). The zero-feedback delays assumption will be relaxed in extensions of the work to be pursued in the near future.

In the first scenario, we have considered reconfiguration of the metasurface using the SIMO controller to guide the reflected single beam towards a single stationary receiver located at $\theta^*_{r,1} = 50°$ and $\varphi^*_{r,1} = 80°$. We assume the initial reflection angle components to be $\theta_{r,1} = 45°$ and $\varphi_{r,1} = 85°$. The amplitude and frequency values of the perturbation signals in this scenario are: $a_1 = 0.15$, $a_2 = 0.1$, $\omega_{f,1} = 50$, and $\omega_{f,2} = 40$. The simulation results of the considered scenario are shown in Fig. 3, where the time evolution of the normalized power $P_{r,1}$ of the received signal is depicted together with the $\theta_{r,1}$ and $\varphi_{r,1}$ of the reflected beam. It is evident from the results, where the normalized received power rises from 0.65 at $t = 0$s to 1 at $t = 12.13$s, that the proposed SIMO controller is successful in directing the reflected beam towards the desired direction. The reported convergence time though is relatively high and fine-tuning of the controller parameters to speed up the convergence will be investigated in the future.

In the second scenario, we have considered a mobile receiver which initially resides at $\theta^*_{r,1} = 50°$ and $\varphi^*_{r,1} = 80°$ from $t = 0$s till $t = 40$s and then moves from $t = 40$s till $t = 70$s to reach the new location $\theta^*_{r,1} = 55°$ and $\varphi^*_{r,1} = 85°$. Here, we have assumed $\theta_{r,1} = 45°$ and $\varphi_{r,1} = 75°$ as initial reflecting angles, whereas parameter values of the perturbation

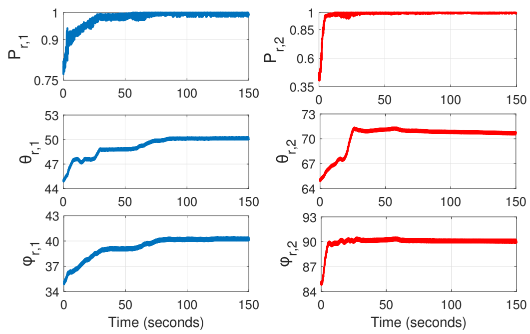

Fig. 5. Received power and reflection angles in the case of two stationary receivers.

signals are kept identical to the ones in the previous scenario. The simulation results depicting the time evolution of the normalized $P_{r,1}$, $\theta_{r,1}$, and $\varphi_{r,1}$ in this scenario are shown in Fig. 4. It is evident from the results that the proposed SIMO controller scheme can also successfully direct the reflected beam towards the desired location of the mobile receiver, even when the values of both target angles are time-varying. A major objective in future work will be to consider increasing values of the speed of the mobile receiver representing for example a car moving at a speed of 30 km/h.

Finally, in the last scenario, we have considered reconfiguration of the metasurface via two SIMO controllers in a cascaded connection to guide two reflected beams towards two stationary receivers as shown in Fig. 1. We assume the initial reflection angles of the two beams to be $\theta_{r,1} = 45°$, $\varphi_{r,1} = 35°$, $\theta_{r,2} = 65°$ and $\varphi_{r,2} = 85°$, respectively. The target reflection angles of receivers R_1 and R_2 are as follows: $\theta^*_{r,1} = 50°$, $\varphi^*_{r,1} = 40°$, $\theta^*_{r,2} = 70°$, and $\varphi^*_{r,2} = 90°$. Parameter values for the SIMO controllers of both users were selected as follows: a_1 =0.15, $a_2 = 0.2$, $\omega_{f,1} = 10$, and $\omega_{f,2} = 5$. The results in Fig. 5 indicate that the normalized received powers $P_{r,1}$ and $P_{r,2}$ rises from 0.77 and 0.44 at time $t = 0$s to 1 at $t = 57.43$s and $t = 11.81$s, respectively. Despite a steady state error of approximately 1% for parameter $\theta_{r,2}$, the reflection angles of both beams converge to a bounded region in the neighborhood of the targeted reflection angles as predicted by the outlined theorems.

VI. Conclusion

Towards autonomous operation of metasurfaces, in this paper, we consider for the first time feedback-based real-time reconfiguration of the metasurface controller states for guiding reflected beams via beam splitting towards multiple, possibly mobile users. As standard extremum seeking control strategies have proved to be ineffective, we employ direct estimation of the gradient via online parameters identification techniques, coupled with the method of steepest ascent. The latter is shown to be effective via both analysis and simulations. Future work will involve further analytical/simulative evaluation to improve the convergence time in the presence of faults on the controller network and information dissemination delays. Furthermore, we will also explore reinforcement learning-based techniques to consider the joint optimization problem and to reduce the time consumption of the resulting algorithms.

Acknowledgment

This work has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 739578 and the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation, and Digital Policy.

References

[1] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "A novel communication paradigm for high capacity and security via programmable indoor wireless environments in next generation wireless systems," *Ad Hoc Networks*, vol. 87, pp. 1–16, 2019.

[2] M. A. ElMossallamy, H. Zhang, L. Song, K. G. Seddik, Z. Han, and G. Y. Li, "Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, pp. 990–1002, 2020.

[3] G. Gradoni, M. Di Renzo, A. Diaz-Rubio, S. Tretyakov, C. Caloz, Z. Peng, A. Alu, G. Lerosey, M. Fink, V. Galdi *et al.*, "Smart radio environments," *arXiv preprint arXiv:2111.08676*, 2021.

[4] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Transactions on Communications*, vol. 69, no. 5, pp. 3313–3351, 2021.

[5] J. Y. Dai, W. Tang, M. Z. Chen, C. H. Chan, Q. Cheng, S. Jin, and T. J. Cui, "Wireless communication based on information metasurfaces," *IEEE Transactions on Microwave Theory and Techniques*, 2021.

[6] Y. Li, S. Bettinga, J. Eisenbeis, J. Kowalewski, X. Wan, X. Long, T. Li, A. Jauch, T. Cui, and T. Zwick, "Beamsteering for 5g mobile communication using programmable metasurface," *IEEE Wireless Communications Letters*, vol. 10, no. 7, pp. 1542–1546, 2021.

[7] Q. Ma, G. D. Bai, H. B. Jing, C. Yang, L. Li, and T. J. Cui, "Smart metasurface with self-adaptively reprogrammable functions," *Light: science & applications*, vol. 8, no. 1, pp. 1–12, 2019.

[8] C. Rizza, V. Loscrí, and M. Parchin, "Real-time beam steering in mmwave with reconfigurable intelligent meta-surfaces based on a genetic algorithm," in *IEEE Global Communications Conference*, 2021.

[9] C. You, B. Zheng, and R. Zhang, "Fast beam training for irs-assisted multiuser communications," *IEEE Wireless Communications Letters*, vol. 9, no. 11, pp. 1845–1849, 2020.

[10] N. Ashraf, M. Lestas, T. Saeed, H. Taghvaee, S. Abadal, A. Pitsillides, and C. Liaskos, "Extremum seeking control for beam steering using hypersurfaces," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, 2020, pp. 1–6.

[11] M. Krstic and H.-H. Wang, "Stability of extremum seeking feedback for general nonlinear dynamic systems," *Automatica*, vol. 36, no. 4, pp. 595–601, 2000.

[12] A. Ghaffari, M. Krstić, and D. Nešić, "Multivariable newton-based extremum seeking," *Automatica*, vol. 48, no. 8, pp. 1759–1767, 2012.

[13] O. Tsilipakos, A. C. Tasolamprou, A. Pitilakis, F. Liu, X. Wang, M. S. Mirmoosa, D. C. Tzarouchis, S. Abadal, H. Taghvaee, C. Liaskos *et al.*, "Toward intelligent metasurfaces: the progress from globally tunable metasurfaces to software-defined metasurfaces with an embedded network of controllers," *Advanced optical materials*, vol. 8, no. 17, 2020.

[14] A. C. Tasolamprou, A. Pitilakis, S. Abadal, O. Tsilipakos, X. Timoneda, H. Taghvaee, M. S. Mirmoosa, F. Liu, C. Liaskos, A. Tsioliaridou *et al.*, "Exploration of intercell wireless millimeter-wave communication in the landscape of intelligent metasurfaces," *IEEE access*, vol. 7, pp. 122 931–122 948, 2019.

[15] H. Taghvaee, S. Abadal, A. Pitilakis, O. Tsilipakos, A. C. Tasolamprou, C. Liaskos, M. Kafesaki, N. V. Kantartzis, A. Cabellos-Aparicio, and E. Alarcón, "Scalability analysis of programmable metasurfaces for beam steering," *IEEE access*, vol. 8, pp. 105 320–105 334, 2020.

[16] H. Taghvaee, A. Cabellos-Aparicio, J. Georgiou, and S. Abadal, "Error analysis of programmable metasurfaces for beam steering," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 10, no. 1, pp. 62–74, 2020.

[17] P. A. Ioannou and J. Sun, *Robust adaptive control*. PTR Prentice-Hall Upper Saddle River, NJ, 1996, vol. 1.