

# Generalized Firefly Algorithm for Optimal Transmit Beamforming

Tuan Anh Le and Xin-She Yang

**Abstract**—This paper proposes a generalized Firefly Algorithm (FA) to solve an optimization framework having objective function and constraints as multivariate functions of independent optimization variables. Four representative examples of how the proposed generalized FA can be adopted to solve downlink beamforming problems are shown for a classic transmit beamforming, cognitive beamforming, reconfigurable-intelligent-surfaces-aided (RIS-aided) transmit beamforming, and RIS-aided wireless power transfer (WPT). Complexity analyses indicate that in large-antenna regimes the proposed FA approaches require less computational complexity than their corresponding interior point methods (IPMs) do, yet demand a higher complexity than the iterative and the successive convex approximation (SCA) approaches do. Simulation results reveal that the proposed FA attains the same global optimal solution as that of the IPM for an optimization problem in cognitive beamforming. On the other hand, the proposed FA approaches outperform the iterative, IPM and SCA in terms of obtaining better solution for optimization problems, respectively, for a classic transmit beamforming, RIS-aided transmit beamforming and RIS-aided WPT.

**Index Terms**—Firefly algorithm, nature-inspired optimization, transmit beamforming, reconfigurable intelligent surfaces.

## I. INTRODUCTION

TRANSMIT beamforming problems are normally cast as optimization problems where beamforming vectors are optimization variables. Two fundamental optimization problems in transmit beamforming include: i) minimizing the total transmit power subject to signal-to-interference-plus-noise-ratio (SINR) constraints [1]–[4]; ii) maximizing the weakest SINR subject to a total power constraint [5], [6]. In fact, these two problems are equivalent [7], [8]. A generalized version of the second problem is introduced in [8] where the objective is to maximize an arbitrary utility function of SINRs, which is strictly increasing in every receiver’s SINR, subject to a power constraint. The other variation of the second optimization problem is the sum rate maximization [9], [10]. Furthermore, additional constraints can be introduced to these fundamental problems to capture other wireless communication applications. For instance, a soft-shaping interference constraint was added for cognitive radio scenarios [11], [12] while a power transfer constraint was included for simultaneous-wireless-information-and-power-transfer scenarios [13]. In addition, various metrics have been utilized to formulate downlink beamforming optimization problems such as secrecy capacity

[14], energy efficiency [15], data transmission reliability, data transmission security, and power transfer reliability [16].

Since the SINR is a non-convex quadratic function of the beamforming vectors, the two fundamental beamforming optimization problems are NP-hard and cannot be solved in polynomial time. Fortunately, exploiting the hidden convexity property of the SINR metric, an elegant framework was proposed in [2] to convert these two optimization problems into convex conic programming forms, which can be effectively solved by a standard interior point method (IPM). Furthermore, uplink-downlink duality was utilized to derive iterative algorithms to find optimal beamforming vectors for some power minimization problems, e.g., [1], [4], [17], [18]. An iterative algorithm was introduced in [9] to attain optimal beamforming vectors for the sum rate maximization.

Numerous transmit beamforming problems can be realized in quadratically constrained quadratic programs (QCQPs) of beamforming vectors, which are mostly non-convex [11], [19]. To solve a QCQP problem, a semidefinite relaxation technique [20] is adopted in which the original QCQP is converted to a convex semidefinite programming (SDP) with new optimization variables as beamforming matrices. If solving the transformed SDP yields a rank-one optimal beamforming matrix, then this optimal matrix is also the optimal solution to the original QCQP. Otherwise, an approximated solution to the original QCQP can be obtained by exploiting some rank-one approximations or the Gaussian randomize procedure [19]. Unfortunately, obtaining such solution requires further computational resources yet results in a sub-optimal solution.

Optimization variables for downlink beamforming problems may include different types of beamforming vectors. For example, in a reconfigurable-intelligent-surface-aided (RIS-aided) communication system, see e.g., [21], [22] and references therein, the optimization variables are active beamforming vectors for the base station (BS) and a passive beamforming vector for the RIS. The objective function and/or constraints for a RIS-aided communication system are functions of both active and passive beamforming vectors. These beamforming vectors are independent variables yet need to be jointly optimized making their problems non-convex. Widely adopted approaches for tackling such problems are to iteratively solve two sub-optimization problems, a.k.a., alternative optimization (AO) approach [21], or to approximate a non-convex using first-order Taylor expansion, a.k.a., successive convex approximation (SCA) [23]. In an AO approach, each of these two sub-optimization problems, one variable is treated as a constant while solving for the other. These sub-optimization problems themselves are mostly in QCQP forms. Due to the inherent

T. A. Le and X.-S. Yang are with the Faculty of Science and Technology, Middlesex University, London, NW4 4BT, UK. Email: {t.le; x.yang}@mdx.ac.uk.

This paper has been presented in part at the IEEE Vehicular Technology Conference (VTC 2023-Spring), Florence, Italy, June, 20-23, 2023.

non-convexity character of the original and sub-optimization problems, the resulting active and passive beamforming vectors may not be the global solutions. Whereas in a SCA approach, a lower (or upper) bounded solution is normally attained.

IPMs, a.k.a., barrier methods, are gradient based algorithms being good at exploitation,<sup>1</sup> a.k.a., intensification, hence, they are regarded as effective methods to solve convex optimization problems [25]. Unfortunately, most of transmit beamforming problems are non-convex. Solving non-convex optimization problems requires algorithms having better exploration<sup>2</sup> ability than that of the IPMs to avoid getting trapped in a local mode. Firefly algorithm (FA), i.e., a nature-inspired algorithm, possesses both exploitation and exploration abilities. Consequently, FA is a good candidate for solving non-convex downlink beamforming problems. FA is an easy-to-implement, simple, and flexible algorithm based on the flashing characters and behaviour of tropical fireflies [24]. FA was first developed and published by Xin-She Yang, respectively, in late 2007 and in 2008 [24], [26] for optimization problems with objective and constraints being functions of a single optimization variable. Although FA has been widely applied to many applications [27], there has not been any significant work investigating the application of FA in solving transmit beamforming problems. There were only two attempts to adopt FA for a throughput maximization problem in [28] and for a power minimization problem in [29]. As these two attempts only capture two fundamental transmit beamforming problems, it is not clear how FA can be adopted to solve other types of transmit beamforming problems.

This paper takes a further step on implementing FA to solve a wider range of transmit beamforming optimization problems. The contributions of the paper can be summarized as follows.

- The paper proposes a generalized FA to find the optimal solution of an optimization framework where its objective function and constraints are multivariate functions of multiple independent optimization variables. The problems in [28] and [29] are only two special cases of the proposed generalized FA while the proposed generalized FA is capable of handling a larger range of transmit beamforming problems.
- The paper shows four representative examples of how the generalized FA can be adopted for solving transmit beamforming problems, i.e., a classic transmit beamforming approach, a cognitive beamforming approach, a RIS-aided beamforming approach, and RIS-aided wireless power transfer (WPT) approach. The applications of the proposed generalized FA are beyond these four examples which are only given to showcase how different types of beamforming problems can be handled by the generalized FA.
- For the sake of completeness and comparison, the iterative closed form or SDP forms of the under investigated beamforming approaches are represented. The paper analyzes

<sup>1</sup>Exploitation is the ability of using any information from the problem of interest to form new solutions which are better than the current ones [24].

<sup>2</sup>Exploration is the ability of efficient exploring the search space to form new solutions with sufficient diversity and far from the existing ones [24].

and compares the complexities of the iterative or SDP and FA implementations of each beamforming approach.

- Simulations are carried out to evaluate the performances of the proposed FAs for the classic transmit beamforming, cognitive beamforming, RIS-aided, and RIS-aided WPT beamforming approaches.

**Notation:** Lower and upper case letter  $y$  and  $Y$ : a scalar; bold lower case letter  $\mathbf{y}$ : a column vector; bold upper case letter  $\mathbf{Y}$ : a matrix;  $\|\cdot\|$ : the Euclidean norm;  $(\cdot)^T$ : the transpose operator;  $(\cdot)^H$ : the complex conjugate transpose operator;  $\text{Tr}(\cdot)$ : the trace operator;  $\mathbf{Y} \geq \mathbf{0}$ :  $\mathbf{Y}$  is positive semidefinite;  $\mathbf{I}_x$ : an  $x \times x$  identity matrix;  $\mathcal{O}$ : the big O notation;  $\mathbb{C}^{M \times 1}$ : the set of all  $M \times 1$  vectors with complex elements;  $\mathbb{H}^{M \times M}$ : the set of all  $M \times M$  Hermitian matrices;  $y \sim \mathcal{CN}(0, \sigma^2)$ :  $y$  is a zero-mean circularly symmetric complex Gaussian random variable with variance  $\sigma^2$ ;  $\text{diag}(\mathbf{y})$ : a diagonal matrix whose diagonal elements are the entries of vector  $\mathbf{y}$ ; and finally  $\text{diag}(\mathbf{Y})$ : a vector whose entries are the diagonal elements of matrix  $\mathbf{Y}$ .

## II. GENERALIZED FIREFLY ALGORITHM FRAMEWORK

### A. Proposed Generalized Firefly Algorithm Framework

The FA was developed based on the following three idealized rules [24], [26]. First, any firefly attracts other fireflies regardless of its sex. Second, the attractiveness of any firefly to the other one is proportional to its brightness. Both attractiveness and brightness decrease as the distance between these two fireflies increases. Given two flashing fireflies, the darker firefly will move towards the brighter one. If a firefly does not find any brighter one, it will make a random move. Third, the brightness of a firefly depends on the landscape of the objective function.

In this section, we propose a generalized FA to find optimal solution for an optimization framework containing both objective and constraints as multivariate functions of independent variables. To that end, we first introduce the following optimization framework.

$$\begin{aligned} & \underset{\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z}}{\text{minimize}} && f(\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z}), \\ & \text{subject to} && g_l(\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z}) \leq 0, \quad l \in \{1, 2, \dots, L\}, \\ & && h_k(\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z}) = 0, \quad k \in \{1, 2, \dots, K\}, \end{aligned} \quad (1)$$

where  $\mathbf{A} \in \mathbb{C}^{M_a \times N_a}$ ,  $\mathbf{B} \in \mathbb{C}^{M_b \times N_b}$ ,  $\dots$ ,  $\mathbf{Z} \in \mathbb{C}^{M_z \times N_z}$ , i.e.,  $M_a, N_a, M_b, N_b, \dots, M_z, N_z \geq 1$ , are decision variables, a.k.a., optimization variables. Depending on the the values of  $\{M_a, N_a, M_b, N_b, \dots, M_z, N_z\}$ , the decision variables can be matrices, vectors, scalars, or the combination of all.

We continue by using the penalty method [24], [26] to equivalently rewrite (1) as:

$$\underset{\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z}}{\text{minimize}} \quad f(\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z}) + P(\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z}), \quad (2)$$

where  $P(\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z})$  is the penalty term defined as:

$$\begin{aligned} P(\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z}) = & \sum_{l=1}^L \lambda_l \max\{0, g_l(\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z})\}^2 \\ & + \sum_{k=1}^K \rho_k \{h_k(\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z})\}^2. \end{aligned} \quad (3)$$

In (3),  $\lambda_l > 0$ ,  $\forall l$ , and  $\rho_k > 0$ ,  $\forall k$ , are penalty constants. Let  $\{\mathbf{A}_i, \mathbf{B}_i, \dots, \mathbf{Z}_i\}$  be the  $i$ -th firefly amongst the population of  $N$  fireflies, i.e.,  $i \in \{1, 2, \dots, N\}$ . Following the second rule of the FA, the brightest firefly is the most attractive one. Since the proposed optimization framework is a minimization, we define the brightness of firefly  $i$  as:<sup>3</sup>

$$I_i(\mathbf{A}_i, \mathbf{B}_i, \dots, \mathbf{Z}_i) = \frac{1}{f(\mathbf{A}_i, \mathbf{B}_i, \dots, \mathbf{Z}_i) + P(\mathbf{A}_i, \mathbf{B}_i, \dots, \mathbf{Z}_i)}. \quad (4)$$

For any two fireflies  $i, j \in \{1, 2, \dots, N\}$ , if  $I_j(\mathbf{A}_j, \mathbf{B}_j, \dots, \mathbf{Z}_j) > I_i(\mathbf{A}_i, \mathbf{B}_i, \dots, \mathbf{Z}_i)$ , then firefly  $i$  will move towards firefly  $j$  at  $(n+1)$ -th generation as:

$$\mathbf{A}_i^{(n+1)} = \mathbf{A}_i^{(n)} + \beta_{a,0} e^{-\gamma_a (r_{a,ij}^{(n)})^2} (\mathbf{A}_j^{(n)} - \mathbf{A}_i^{(n)}) + \alpha_a^{(n)} \mathbf{\Lambda}_{a,i}^{(n)}, \quad (5)$$

$$\mathbf{B}_i^{(n+1)} = \mathbf{B}_i^{(n)} + \beta_{b,0} e^{-\gamma_b (r_{b,ij}^{(n)})^2} (\mathbf{B}_j^{(n)} - \mathbf{B}_i^{(n)}) + \alpha_b^{(n)} \mathbf{\Lambda}_{b,i}^{(n)}, \quad (6)$$

⋮

$$\mathbf{Z}_i^{(n+1)} = \mathbf{Z}_i^{(n)} + \beta_{z,0} e^{-\gamma_z (r_{z,ij}^{(n)})^2} (\mathbf{Z}_j^{(n)} - \mathbf{Z}_i^{(n)}) + \alpha_z^{(n)} \mathbf{\Lambda}_{z,i}^{(n)}, \quad (7)$$

where  $r_{a,ij}^{(n)} = \|\mathbf{A}_j^{(n)} - \mathbf{A}_i^{(n)}\|$ ,  $r_{b,ij}^{(n)} = \|\mathbf{B}_j^{(n)} - \mathbf{B}_i^{(n)}\|$ ,  $\dots$ ,  $r_{z,ij}^{(n)} = \|\mathbf{Z}_j^{(n)} - \mathbf{Z}_i^{(n)}\|$  are the Cartesian distances which are not necessary Euclidean distances yet they can be any measure effectively characterized the quantities of interest in the optimization problem;  $\beta_{a,0}, \beta_{b,0}, \dots, \beta_{z,0}$  are, respectively, the attractiveness at  $r_{a,ij}^{(n)} = 0, r_{b,ij}^{(n)} = 0, \dots, r_{z,ij}^{(n)} = 0$ ; finally  $\gamma_a, \gamma_b, \dots, \gamma_z$  present the variations of the attractiveness. The second terms in (5), (6), and (7) capture the attractions. The third terms in (5), (6), and (7) are randomizations with randomization factors  $\alpha_a^{(n)}, \alpha_b^{(n)}, \dots, \alpha_z^{(n)}$  and  $\mathbf{\Lambda}_{a,i}^{(n)} \in \mathbb{C}^{M_a \times N_a}, \mathbf{\Lambda}_{b,i}^{(n)} \in \mathbb{C}^{M_b \times N_b}, \dots, \mathbf{\Lambda}_{z,i}^{(n)} \in \mathbb{C}^{M_z \times N_z}$  being matrices of random numbers drawn from a Gaussian or an uniform distribution. The proposed generalized FA for solving the optimization framework (1) is summarized in Algorithm 1, where  $T$  is the maximum generation of the algorithm. For any particular optimization problem subsumed under the framework, the corresponding FA will have the same steps as those in Algorithm 1 except the input, step 3, step 16, step 18, step 19, and the return value.

### B. Asymptotic Convergence and Optimality

Since the firefly algorithm, like quite a few other nature-inspired algorithms, is a metaheuristic algorithm, there is no rigorous proof of convergence so far in the current literature, despite many applications of such metaheuristic algorithms. In this section, we provide some intuitive discussions on the optimality and convergence of the FA framework.<sup>4</sup>

1) *Asymptotic Optimality*: Without loss of generality, let  $\gamma_a = \gamma_b = \dots = \gamma_z = \gamma$ , we consider two special cases of the variations of the attractiveness when  $\gamma \rightarrow 0$  and  $\gamma \rightarrow \infty$ . When  $\gamma \rightarrow 0$ , it is clear that  $e^{-\gamma (r_{a,ij}^{(n)})^2} \rightarrow 1, e^{-\gamma (r_{b,ij}^{(n)})^2} \rightarrow 1, \dots, e^{-\gamma (r_{z,ij}^{(n)})^2} \rightarrow 1$ . Therefore the attractivenesses in (5), (6), and (7) are constant and, respectively, equal to  $\beta_{a,0}, \beta_{b,0}$ , and  $\beta_{z,0}$ . Equivalently, it is an idealized sky scenario where the

<sup>3</sup>Note that if (1) is a maximization problem, then (2) can be expressed as: minimize  $-f(\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z}) + P(\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z})$ .

<sup>4</sup>Mathematical analysis of the FA's optimality and convergence deserves an important research topic. Such analysis is postponed to future research due to the space constraint.

### Algorithm 1 Generalized Firefly Algorithm for solving (1)

---

```

1: Input: FA parameters:  $N, T, \lambda_l, \rho_k, \beta_{a,0}, \beta_{b,0}, \dots, \beta_{z,0}, \gamma_a, \gamma_b, \dots, \gamma_z$ ; Optimization data: the structures/parameters of functions  $f(\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z}), g_l(\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z}), h_k(\mathbf{A}, \mathbf{B}, \dots, \mathbf{Z})$ ;
2: Randomly generate  $N$  populations  $\{\{\mathbf{A}_1, \mathbf{B}_1, \dots, \mathbf{Z}_1\}, \{\mathbf{A}_2, \mathbf{B}_2, \dots, \mathbf{Z}_2\}, \dots, \{\mathbf{A}_N, \mathbf{B}_N, \dots, \mathbf{Z}_N\}\}$ ;
3: Evaluate the light intensities of  $N$  population as (4);
4: Rank the fireflies in a descending order of  $I_i(\mathbf{A}_i, \mathbf{B}_i, \dots, \mathbf{Z}_i)$ ;
5: Define the current best solution:  $I^* := I_1(\mathbf{A}^*, \mathbf{B}^*, \dots, \mathbf{Z}^*); \{\mathbf{A}^*, \mathbf{B}^*, \dots, \mathbf{Z}^*\} := \{\mathbf{A}_1, \mathbf{B}_1, \dots, \mathbf{Z}_1\}$ ;
6: for  $n = 1 : T$  do
7:   for  $i = 1 : N$  do
8:     for  $j = 1 : N$  do
9:       if  $I_i(\mathbf{A}_i, \mathbf{B}_i, \dots, \mathbf{Z}_i) > I^*$  then
10:         $I^* := I_i(\mathbf{A}_i, \mathbf{B}_i, \dots, \mathbf{Z}_i); \{\mathbf{A}^*, \mathbf{B}^*, \dots, \mathbf{Z}^*\} := \{\mathbf{A}_i, \mathbf{B}_i, \dots, \mathbf{Z}_i\}$ ;
11:      end if
12:      if  $I_j(\{\mathbf{A}_j, \mathbf{B}_j, \dots, \mathbf{Z}_j\}) > I^*$  then
13:         $I^* := I_j(\mathbf{A}_j, \mathbf{B}_j, \dots, \mathbf{Z}_j); \{\mathbf{A}^*, \mathbf{B}^*, \dots, \mathbf{Z}^*\} := \{\mathbf{A}_j, \mathbf{B}_j, \dots, \mathbf{Z}_j\}$ ;
14:      end if
15:      if  $I_j(\mathbf{A}_j, \mathbf{B}_j, \dots, \mathbf{Z}_j) > I_i(\mathbf{A}_i, \mathbf{B}_i, \dots, \mathbf{Z}_i)$  then
16:        Move firefly  $i$  towards firefly  $j$  as (5)-(7);
17:      end if
18:      Attractiveness varies with distances via  $e^{-\gamma_a (r_{a,ij}^{(n)})^2}, e^{-\gamma_b (r_{b,ij}^{(n)})^2}, \dots, e^{-\gamma_z (r_{z,ij}^{(n)})^2}$ ;
19:      Evaluate new solutions and update light intensity as (4);
20:    end for
21:  end for
22:  Rank the fireflies in a descending order of  $I_i(\mathbf{A}_i, \mathbf{B}_i, \dots, \mathbf{Z}_i)$ ;
23:  Update the current best solution:  $I^* := I_1(\mathbf{A}^*, \mathbf{B}^*, \dots, \mathbf{Z}^*); \{\mathbf{A}^*, \mathbf{B}^*, \dots, \mathbf{Z}^*\} := \{\mathbf{A}_1, \mathbf{B}_1, \dots, \mathbf{Z}_1\}$ ;
24: end for
25: return  $\{\mathbf{A}^*, \mathbf{B}^*, \dots, \mathbf{Z}^*\}$ .

```

---

brightness of each firefly does not change over the distance, which can be seen everywhere. Consequently, a global optimum can be obtained.

On the other hand, when  $\gamma \rightarrow \infty$ , it is obvious that  $e^{-\gamma (r_{a,ij}^{(n)})^2} \rightarrow 0, e^{-\gamma (r_{b,ij}^{(n)})^2} \rightarrow 0, \dots, e^{-\gamma (r_{z,ij}^{(n)})^2} \rightarrow 0$ , indicating that the attractiveness of each firefly is zero. Equivalently, each firefly is randomly in a heavily foggy region and cannot be seen by the others. Each will randomly move and the optimality is not always guaranteed. In this case, FA is equivalent to a random search approach.

In fact, the attractiveness is in between these two extreme cases, i.e.,  $0 < \gamma < \infty$ . The value of  $\gamma^{-0.5}$  defines the average distance of a herd of fireflies being seen by its adjacent herds. Hence, the entire population can be separated into number of herds. This automatic division property provides FA suitable

ability of handling highly nonlinear and multimodal optimization problems. By controlling the attractiveness  $\gamma_a, \gamma_b, \dots, \gamma_z$  and the roaming randomness  $\alpha_a, \alpha_b, \dots, \alpha_z$ , it has been shown in previous studies that FA can outperform both Particle Swarm Optimization (PSO), see, e.g., [30]–[33], and random search approaches, see e.g., [24], [26].

2) *Asymptotic Convergence*: When  $\gamma \rightarrow 0$ , the convergence of FA is similar to that of PSO where the convergence was analyzed by Clerc and Kennedy in 2002 in [34]. When  $\gamma \rightarrow \infty$ , the FA may act like a random search, though its behaviour is similar to that of Simulated Annealing (SA) because the FA's solution is perturbed or modified in the similar way as that in the SA in this limiting case. The SA was shown to be convergent under the right-cooling conditions [35]. The reduction of the roaming randomness, i.e.,  $\alpha_a, \alpha_b, \dots, \alpha_z$ , in the FA can be considered as a type of cooling schedule, and thus it can be expected that FA can converge in this case.

Let us now investigate the case when  $0 < \gamma < \infty$ . Given a very large number of firefly population  $N$ , it can be assumed that  $N$  is much greater than the number of local optima. The initial locations of  $N$  fireflies should be uniformly distributed over the whole search space. As the iterations of Algorithm 1 progress, i.e.,  $n$  increases, these initial  $N$  fireflies should converge into all locally brighter ones, i.e., the local optima including the global ones, in a stochastic manner due to the third term in (5), (6), and (7). By comparing the brightest fireflies amongst the locally brighter ones, i.e., the best solutions amongst the local optima, the global optima can be attained. Theoretically, these fireflies will reach the global optimal when  $N \rightarrow \infty$  and  $n \gg 1$ . However, it has been reported in the related literature that the FA converges with less than 50 to 100 generations [24], [26].

In sections IV, V, and VI, we present how the proposed FA can be adopted to solve optimization problems for transmit beamforming designs.<sup>5</sup> Hereafter, “min” and “s. t.” are, respectively, used to represent “minimize” and “subject to”.

### III. TRANSMIT BEAMFORMING

In this section we consider a classic transmit beamforming problem with a well-known iterative method based on uplink-downlink duality. We then introduce our FA solution to the problem.

#### A. Problem Formulation

1) *Problem Formulation*: Consider an  $M_t$ -antenna BS serving  $U$  single-antenna mobile users. Let  $\mathbf{h}_i^H \in \mathbb{C}^{1 \times M_t}$ ,  $\mathbf{w}_i \in \mathbb{C}^{M_t \times 1}$  and  $s_i$ , respectively, be the channel between the  $i$ -th user and the BS, the information-beamforming vector and the data symbol for the  $i$ th user. The overall signal received by the  $i$ th user is  $y_i = \sum_{j=1}^U \mathbf{h}_i^H \mathbf{w}_j s_j + n_i$  where  $n_i$  is a zero mean circularly symmetric complex Gaussian noise with variance  $\sigma^2$ , i.e.,  $n_i \sim \mathcal{CN}(0, \sigma^2)$ , at the user. Let  $\mathbf{R}_i = \mathbf{h}_i \mathbf{h}_i^H$  represent the instantaneous channel state information (CSI) or  $\mathbf{R}_i = \mathbb{E}(\mathbf{h}_i \mathbf{h}_i^H)$  denote the statistical CSI,  $\{\mathbf{w}_i\} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_U\}$  be the set

of candidate information-beamforming vectors for all users. Assuming that  $\mathbb{E}(|s_i|^2) = 1$ , the SINR at the  $i$ -th user is

$$\text{SINR}_i = \frac{\mathbf{w}_i^H \mathbf{R}_i \mathbf{w}_i}{\sum_{j=1, j \neq i}^U \mathbf{w}_j^H \mathbf{R}_i \mathbf{w}_j + \sigma^2}. \quad (8)$$

We design the set of beamforming vectors  $\{\mathbf{w}_i\}$  such that the BS's total transmit power is minimized while maintaining the SINR level at each user above the required threshold. To that end, the problem is formulated as follows:

$$\begin{aligned} \min_{\mathbf{w}_i} \quad & \sum_{i=1}^U \mathbf{w}_i^H \mathbf{w}_i \\ \text{s. t.} \quad & \frac{\mathbf{w}_i^H \mathbf{R}_i \mathbf{w}_i}{\sum_{j=1, j \neq i}^U \mathbf{w}_j^H \mathbf{R}_i \mathbf{w}_j + \sigma_i^2} \geq \gamma_i, \quad \forall i \in \{1, \dots, U\}, \end{aligned} \quad (9)$$

where  $\gamma_i$  is the required SINR level for the  $i$ -th user. Problem (9) is known as non-convex due to the SINR constraint.

2) *Iterative Approach*: An elegant approach to solve (9) was introduced in [1] based on uplink-downlink duality where the optimal solution of the downlink problem can be sought via solving the following dual-uplink problem:<sup>6</sup>

$$\begin{aligned} \min_{p_i} \quad & \sum_{i=1}^U p_i \\ \text{subject to} \quad & \mathbf{p} \geq \mathbf{\Gamma} \mathbf{t}(\mathbf{p}), \end{aligned} \quad (10)$$

where  $\mathbf{p} = [p_1 \ p_2 \ \dots \ p_U]^T$ ,  $\mathbf{\Gamma} = \text{diag}[\gamma_1, \gamma_2, \dots, \gamma_U]$ ,  $\mathbf{t}(\mathbf{p}) = [t_1(\mathbf{p}) \ t_2(\mathbf{p}) \ \dots \ t_U(\mathbf{p})]^T$ ,

$$t_i(\mathbf{p}) = \arg \min_{\hat{\mathbf{w}}_i} \frac{\hat{\mathbf{w}}_i^H \mathbf{Q}_i(\mathbf{p}) \hat{\mathbf{w}}_i}{\hat{\mathbf{w}}_i^H \mathbf{R}_i \hat{\mathbf{w}}_i}, \quad (11)$$

$\mathbf{Q}_i(\mathbf{p}) = \left( \sum_{j=1, j \neq i}^U p_j \mathbf{R}_j + \sigma_i^2 \mathbf{I} \right)$ ,  $p_i = \lambda_i \sigma_i^2$  is the dual-uplink power for  $i$ -th user,  $\lambda_i$  is the  $i^{\text{th}}$  Lagrange multiplier associated with the  $i^{\text{th}}$  constraint in (9), and  $\hat{\mathbf{w}}_i$ , i.e.,  $\hat{\mathbf{w}}_i^H \hat{\mathbf{w}}_i = 1$ , is the dual-uplink beamforming vector for  $i$ -th user. Starting from any positive initial value of  $\mathbf{p}(0)$ , the solution for the dual-uplink problem (10) can be found iteratively as  $\mathbf{p}(n+1) = \mathbf{\Gamma} \mathbf{t}(\mathbf{p}(n))$ . The iterative downlink algorithm to find optimal solutions for (9) is summarised in algorithm 2.

#### B. Proposed Firefly Algorithm

We rewrite (9) as

$$\begin{aligned} \min_{\mathbf{W}} \quad & f(\mathbf{W}) \\ \text{s. t.} \quad & d_i(\mathbf{W}) \leq 0, \quad \forall i, \end{aligned} \quad (12)$$

where  $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_U] \in \mathbb{C}^{M_t \times U}$ ,  $f(\mathbf{W}) = \sum_{i=1}^U \mathbf{w}_i^H \mathbf{w}_i$ ,  $d_i(\mathbf{W}) = -\mathbf{w}_i^H \mathbf{R}_i \mathbf{w}_i + \gamma_i \sum_{j=1, j \neq i}^U \mathbf{w}_j^H \mathbf{R}_i \mathbf{w}_j + \gamma_i \sigma_i^2$ . Using the penalty method, we recast (22) into an unconstrained problem as:

$$\min_{\mathbf{W}} \quad f(\mathbf{W}) + P(\mathbf{W}), \quad (13)$$

<sup>5</sup>The original FA has been discretized to solve various discrete or combinatorial optimization problems [36]. For example, Osaba et al. [37] used a discrete FA to solve rich vehicle routing problems.

<sup>6</sup>This approach was also adopted for transmit beamforming problems in coordinated multi-point (CoMP) transmissions, see e.g., [38] and [39].

**Algorithm 2** Iterative algorithm for problem (9)

- 
- 1: **Input:**  $\mathbf{\Gamma} = \text{diag}[\gamma_1, \gamma_2, \dots, \gamma_U]$ ,  $\mathbf{R}_i$ ,  $\forall i$ , number of iterations  $T$ .
  - 2: Initialize  $\mathbf{p}(1) \geq 0$ .
  - 3: **for**  $n = 1 : T$  **do**
  - 4:   **for**  $i = 1 : U$  **do**
  - 5:     Find  $\hat{\mathbf{w}}_i(n)$  as the dominant eigenvector of the matrix  $\mathbf{G}_i(n) = p_i(n) \mathbf{Q}_i^{-1}(\mathbf{p}(n)) \mathbf{R}_i$
  - 6:     Calculate  $t_i(\mathbf{p}(n)) = \frac{\hat{\mathbf{w}}_i^H(n) \mathbf{Q}_i(\mathbf{p}(n)) \hat{\mathbf{w}}_i(n)}{\hat{\mathbf{w}}_i^H(n) \mathbf{R}_i \hat{\mathbf{w}}_i(n)}$ .
  - 7:   **end for**
  - 8:   Update  $\mathbf{p}(n+1) = \mathbf{\Gamma} \mathbf{t}(\mathbf{p}(n))$ .
  - 9: **end for**
  - 10:  $p_i^* = \mathbf{p}(n+1)$  and  $\hat{\mathbf{w}}_i^* = \hat{\mathbf{w}}_i(n+1)$ .
  - 11: **Output:**  $\mathbf{w}_i^* = \sqrt{p_i^*} \hat{\mathbf{w}}_i^*$ .
- 

where  $P(\mathbf{W})$  is the penalty term given as:

$$P(\mathbf{W}) = \sum_{i=1}^U \lambda_i \max\{0, d_i(\mathbf{W})\}^2, \quad (14)$$

with  $\lambda_i > 0$  is the penalty constant.

Let  $\{\mathbf{W}_i\} = \{\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_U\}\}$  be the  $i$ -th firefly. We initialize a population of  $N$  fireflies  $\{\mathbf{W}_i\}$ ,  $i \in \{1, 2, \dots, N\}$ , and define the light density of the firefly  $\{\mathbf{W}_i\}$  as:

$$I_i(\mathbf{W}_i) = \frac{1}{f(\mathbf{W}_i) + P(\mathbf{W}_i)}. \quad (15)$$

For any two fireflies  $i$  and  $j$  in the population, if  $I_j(\mathbf{W}_j) > I_i(\mathbf{W}_i)$  then the firefly  $i$  will move toward the firefly  $j$  as:

$$\mathbf{W}_i^{(n+1)} = \mathbf{W}_i^{(n)} + \beta_0 e^{-\gamma(r_{ij}^{(n)})^2} (\mathbf{W}_j^{(n)} - \mathbf{W}_i^{(n)}) + \alpha^{(n)} \mathbf{V}, \quad (16)$$

where  $r_{ij}^{(n)} = \|\mathbf{W}_j^{(n)} - \mathbf{W}_i^{(n)}\|$  is the Cartesian distance,  $\beta_0$  is the attractiveness at  $r_{ij}^{(n)} = 0$ ,  $\gamma$  presents the variation of of the attractiveness. The second term of (16) represent the attraction. The third term of (16) is a randomization comprised of a randomization factor  $\alpha^{(n)}$  and a matrix of random numbers  $\mathbf{V} \in \mathbb{C}^{M_t \times U}$ . The random factor  $\alpha^{(n)}$  and the elements of  $\mathbf{V}$  are drawn from either a Gaussian or an uniform distribution.

It can be seen that problem (12) is a special case of the proposed framework (1) where the objective and constraints are functions of optimization variable  $\mathbf{W}$ . Hence, the proposed FA has the same steps as those in Algorithm 1 except steps 3, 16, 18 and 19 given in Algorithm 3.

**Algorithm 3** Modified generalized FA for solving (12)

- 
- Input:** FA parameters:  $N, T, \lambda_i, \beta_0$ ; Optimization data:  $\mathbf{R}_i, \sigma_i^2, \gamma_i$ ;
- Step 3: Evaluate the light intensities of  $N$  fireflies as (15);
- Step 16: Move firefly  $i$  towards firefly  $j$  as (16);
- Step 18: Attractiveness varies with distance via  $e^{-\gamma(r_{ij}^{(n)})^2}$ ;
- Step 19: Evaluate new solutions; update  $I_i(\mathbf{W}_i)$  as (15);
- return**  $\mathbf{W}^*$ .
- 

*C. Complexity Analysis*

The complexity of algorithm 2 is described in the following lemma.

*Lemma 1:* The computational complexity of algorithm 2 is on the order of  $T [U(M_t^3 + M_t^2 + M_t \log M_t) + U]$ .

*Proof:* The proof is based on the observation that complexities of steps 5, 6 and 8 are, respectively, on the order of  $M_t^3 + M_t \log M_t$ ,  $M_t^2$  and  $U$ . ■

*Lemma 2:* The computational complexity of Algorithm 3 is on the order of:

$$TN^2 [M_t^2 + NUM_t(1 + UM_t)] + TN \log N + NM_t U + NUM_t(1 + UM_t) + N \log N. \quad (17)$$

*Proof:* Due to space limitation, we provide main observations to derive (17) as follows. The dominant terms of the computational complexity of Algorithm 3 are at steps 2, 3, 4, 16, 19, and 22. The complexity of generating  $N$  matrices, each matrix of size  $M_t \times U$ , in step 2 is on the order of  $NM_t U$ . The complexity of evaluating each  $d_i(\mathbf{W})$  is on the order of  $UM_t^2$ , while the complexity of evaluating  $\sum_{i=1}^U \mathbf{w}_i^H \mathbf{w}_i$  is on the order of  $UM_t$ .<sup>7</sup> Hence the complexity of calculating the light density for  $N$  fireflies, i.e., steps 3 and 19, is on the order of  $N(UM_t + U^2 M_t^2) = NUM_t(1 + UM_t)$ . The complexity of ranking  $N$  firefly in steps 4 and 22 is  $N \log N$ . Finally, the complexity of moving a firefly in step 16 is on the order of  $M_t^2$ . Assuming a worst case when step 16 is executed in every inner loop of the algorithm, after some manipulations, one can arrive at (17). ■

## IV. COGNITIVE BEAMFORMING

*A. Problem Formulation*

1) *Problem Formulation:* Consider a cognitive wireless communication system consisting of an  $M_t$ -antenna cognitive base station (BS),  $U$  active single-antenna secondary users (SUs) and  $K$  single-antenna primary users (PUs). The cognitive BS is allowed to communicate with its SUs in the same frequency band owned by the primary system if its interference imposed on each PU is less than a predefined tolerable threshold of  $I_{to,k}$ . The received signal at the  $t$ -th SU,  $t \in \{1, \dots, U\}$ , is:

$$y_t = \mathbf{h}_{s,t}^H \mathbf{w}_t s_t + \sum_{j=1, j \neq t}^U \mathbf{h}_{s,t}^H \mathbf{w}_j s_j + n_t, \quad (18)$$

where  $\mathbf{h}_{s,t}^H \in \mathbb{C}^{1 \times M_t}$  is the channel coefficient of the wireless link between the  $t$ -th SU and the cognitive BS;  $\mathbf{w}_t \in \mathbb{C}^{M_t \times 1}$  and  $s_t \sim \mathcal{CN}(0, 1)$  are, respectively, the beamforming vector and the data symbol associated to the  $t$ -th SU; and  $n_t \sim \mathcal{CN}(0, \sigma_t^2)$  is a zero mean circularly symmetric complex Gaussian noise with variance  $\sigma_t^2$ , at the  $t$ -th SU. Let  $\mathbf{R}_{s,t} = \mathbb{E}(\mathbf{h}_{s,t} \mathbf{h}_{s,t}^H)$  for the statistical CSI and  $\mathbf{R}_{s,t} = \mathbf{h}_{s,t} \mathbf{h}_{s,t}^H$  for the instantaneous CSI. The SINR at the  $t$ -th SU can be expressed as:

$$\text{SINR}_t = \frac{\mathbf{w}_t^H \mathbf{R}_{s,t} \mathbf{w}_t}{\sum_{j=1, j \neq t}^U \mathbf{w}_j^H \mathbf{R}_{s,t} \mathbf{w}_j + \sigma_t^2}. \quad (19)$$

<sup>7</sup>Here, we adopt the schoolbook iterative algorithm to evaluate complexity of the multiplication of two matrices of sizes  $n \times m$  and  $m \times p$  as the order of  $nmp$ .

Let  $\mathbf{h}_{p,k}^H \in \mathbb{C}^{1 \times M_t}$  be the channel coefficient of the wireless link between the  $k$ -th PU,  $k \in \{1, \dots, K\}$ , and the cognitive BS,  $\mathbf{R}_{p,k} = \mathbb{E}(\mathbf{h}_{p,k} \mathbf{h}_{p,k}^H)$  for the statistical CSI and  $\mathbf{R}_{p,k} = \mathbf{h}_{p,k} \mathbf{h}_{p,k}^H$  for the instantaneous CSI. The total interference power imposed on the  $k$ -th PU by the cognitive BS is  $\sum_{j=1}^U \mathbf{w}_j^H \mathbf{R}_{p,k} \mathbf{w}_j$ .

Our objective is to design downlink beamforming vectors for the SUs that minimize the cognitive BS transmit power while maintaining the required SINR level for every SU and keeping the interference level imposed at each PU receiver below the predefined tolerable threshold. The optimization problem to design beamforming vectors is cast as:

$$\begin{aligned} \min_{\mathbf{w}_t} \quad & \sum_{t=1}^U \mathbf{w}_t^H \mathbf{w}_t \\ \text{s. t.} \quad & \frac{\mathbf{w}_t^H \mathbf{R}_{s,t} \mathbf{w}_t}{\sum_{j=1, j \neq t}^U \mathbf{w}_j^H \mathbf{R}_{s,t} \mathbf{w}_j + \sigma_t^2} \geq \eta_t, \quad \forall t \in \{1, \dots, U\}, \\ & \sum_{j=1}^U \mathbf{w}_j^H \mathbf{R}_{p,k} \mathbf{w}_j \leq I_{\text{to},k}, \quad \forall k \in \{1, \dots, K\}, \end{aligned} \quad (20)$$

where  $\eta_t$  is the required SINR level for the  $t$ -th SU. Due to the SINR constraint, problem (20) is non-convex.

2) *SDP Approach*: For the sake of completeness, we provide a review on a traditional approach to solve (20) using semidefinite programming (SDP). We first form a new optimization variable  $\mathbf{F}_t = \mathbf{w}_t \mathbf{w}_t^H$  where  $\mathbf{F}_t \geq \mathbf{0}$ ,  $\mathbf{F}_t \in \mathbb{H}^{M_t \times M_t}$ , and  $\mathbf{F}_t$  is a rank-one matrix.<sup>8</sup> We then utilize the identity  $\mathbf{x}^H \mathbf{X} \mathbf{x} = \text{Tr}(\mathbf{X} \mathbf{x} \mathbf{x}^H)$  to rewrite (20) as:

$$\begin{aligned} \min_{\mathbf{F}_t \in \mathbb{H}^{M_t \times M_t}} \quad & \sum_{t=1}^U \text{Tr}(\mathbf{F}_t) \\ \text{s. t.} \quad & \left(1 + \frac{1}{\eta_t}\right) \text{Tr}(\mathbf{R}_{s,t} \mathbf{F}_t) - \sum_{j=1}^U \text{Tr}(\mathbf{R}_{s,t} \mathbf{F}_j) - \sigma_t^2 \geq 0, \quad \forall t, \\ & I_{\text{to},k} - \sum_{j=1}^U \text{Tr}(\mathbf{R}_{p,k} \mathbf{F}_j) \geq 0, \quad \forall k, \\ & \mathbf{F}_t \geq \mathbf{0}, \quad \forall t, \end{aligned} \quad (21)$$

where  $t \in \{1, \dots, U\}$ ,  $k \in \{1, \dots, K\}$ .

Problem (21) is in a standard SDP form. Hence, its optimal solution can be obtained in a polynomial time by using a general purpose IPM, e.g., CVX which is a Matlab based modeling system for constructing and solving disciplined convex programs [40]. In arriving at (21), we have relaxed the rank-one constraint on  $\mathbf{F}_t$ ,  $\forall t$ . If the solution of (21) does not have rank-one, then further computation resources are required to derive a sub-optimal solution via some rank-one approximations or the Gaussian randomize procedure [19].

### B. Proposed Firefly Algorithm

Here, we adopt the generalized FA in Algorithm 1 to solve (20). Rearranging the constraint, we rewrite (20) as:

<sup>8</sup>A matrix is rank-one if and only if it has only one linearly independent column/row.

$$\begin{aligned} \min_{\mathbf{W}} \quad & f(\mathbf{W}) \\ \text{s. t.} \quad & \phi_t(\mathbf{W}) \leq 0, \quad \forall t \in \{1, \dots, U\}, \\ & \varphi_k(\mathbf{W}) \leq 0, \quad \forall k \in \{1, \dots, K\}, \end{aligned} \quad (22)$$

where  $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_U] \in \mathbb{C}^{M_t \times U}$ ,  $f(\mathbf{W}) = \sum_{t=1}^U \mathbf{w}_t^H \mathbf{w}_t$ ,  $\phi_t(\mathbf{W}) = \eta_t \sum_{j=1, j \neq t}^U \mathbf{w}_j^H \mathbf{R}_{s,t} \mathbf{w}_j + \eta_t \sigma_t^2 - \mathbf{w}_t^H \mathbf{R}_{s,t} \mathbf{w}_t$  and  $\varphi_k(\mathbf{W}) = \sum_{j=1}^U \mathbf{w}_j^H \mathbf{R}_{p,k} \mathbf{w}_j - I_{\text{to},k}$ . Using the penalty method, we first transform (22) into an unconstrained problem as:

$$\min_{\mathbf{W}} f(\mathbf{W}) + P(\mathbf{W}), \quad (23)$$

where  $P(\mathbf{W})$  is the penalty term given as:

$$P(\mathbf{W}) = \sum_{t=1}^U \lambda_t \max\{0, \phi_t(\mathbf{W})\}^2 + \sum_{k=1}^K \rho_k \max\{0, \varphi_k(\mathbf{W})\}^2, \quad (24)$$

with  $\lambda_t > 0$  and  $\rho_k > 0$  are penalty constants.

Let  $\mathbf{W}_i = [\mathbf{w}_1^i, \mathbf{w}_2^i, \dots, \mathbf{w}_U^i] \in \mathbb{C}^{M_t \times U}$  be the firefly  $i$ . We initialize a population of  $N$  fireflies  $\mathbf{W}_i$ ,  $i \in \{1, 2, \dots, N\}$ , and define the light density of the firefly  $\mathbf{W}_i$  as:

$$I_i(\mathbf{W}_i) = \frac{1}{f(\mathbf{W}_i) + P(\mathbf{W}_i)}. \quad (25)$$

For any two fireflies  $i$  and  $j$  in the population, if  $I_j(\mathbf{W}_j) > I_i(\mathbf{W}_i)$  then the firefly  $i$  will move toward the firefly  $j$  as:

$$\mathbf{W}_i^{(n+1)} = \mathbf{W}_i^{(n)} + \beta_0 e^{-\gamma(r_{ij}^{(n)})^2} (\mathbf{W}_j^{(n)} - \mathbf{W}_i^{(n)}) + \alpha^{(n)} \mathbf{V}, \quad (26)$$

where  $r_{ij}^{(n)} = \|(\mathbf{W}_j^{(n)} - \mathbf{W}_i^{(n)})\|$  is the Cartesian distance,  $\beta_0$  is the attractiveness at  $r_{ij}^{(n)} = 0$ ,  $\gamma$  presents the variation of of the attractiveness. The second term of (26) captures the attraction. The third term of (26) is a randomization comprised of a randomization factor  $\alpha^{(n)}$  and a matrix of random numbers  $\mathbf{V} \in \mathbb{C}^{M_t \times U}$ . The random factor  $\alpha^{(n)}$  and the elements of  $\mathbf{V}$  are drawn from either a Gaussian or an uniform distribution.

It can be seen that problem (22) is a special case of the proposed framework (1) where the objective and constraints are functions of only one optimization variable  $\mathbf{W}$ . Hence, the proposed FA has the same steps as those in Algorithm 1 except steps 3, 16, 18 and 19 given in Algorithm 4.

---

#### Algorithm 4 Modified generalized FA for solving (20)

---

**Input:** FA parameters:  $N, T, \lambda_t, \rho_k, \beta_0, \gamma$ ; Optimization data:  $\mathbf{R}_{s,t}, \mathbf{R}_{p,k}, \sigma_t^2, \eta_t, I_{\text{to},k}$ ;  
 Step 3: Evaluate the light intensities of  $N$  fireflies as (25);  
 Step 16: Move firefly  $i$  towards firefly  $j$  as (26);  
 Step 18: Attractiveness varies with distance via  $e^{-\gamma(r_{ij}^{(n)})^2}$ ;  
 Step 19: Evaluate new solutions; update  $I_i(\mathbf{W}_i)$  as (25);  
**return**  $\mathbf{W}^*$ .

---

### C. Complexity Analysis

We investigate the complexity of solving (21) in a worst-case runtime of the IPM followed by the complexity analysis of the proposed FA. We start by the following definition.

*Definition 1:* At a given  $\varepsilon > 0$ , the set of  $\{\mathbf{F}_i^\varepsilon\}$  is an  $\varepsilon$ -solution to problem (21), i.e., an acceptable solution with the accuracy of  $\varepsilon$ , if

$$\sum_{i=1}^U \text{Tr}(\mathbf{F}_i^\varepsilon) \leq \min_{\mathbf{F}_i \in \mathbb{H}^{M_i \times M_i}} \sum_{i=1}^U \text{Tr}(\mathbf{F}_i) + \varepsilon. \quad (27)$$

The number of decision variables of (21) is  $M_i^2$ . The complexity of (21) is described in the following lemma.

*Lemma 3:* The computational complexity to attain  $\varepsilon$ -solution to (21) is on the order of:

$$\ln(\varepsilon^{-1}) \sqrt{U(M_i + 1) + K} [(M_i^2 + 1)(U + K) + UM_i^2(M_i^2 + M_i) + M_i^4] M_i^2. \quad (28)$$

*Proof:* We sketch some main steps to arrive at the lemma due to space limitation. It can be observed that (21) has  $(U+K)$  linear-matrix-inequality (LMI) constraints of size 1 and  $U$  LMI constraints of size  $M_i$ . One can follow the same steps as in [41, Section V-A] to derive the following facts: (i) the iteration complexity is on the order of  $\ln(\varepsilon^{-1}) \sqrt{U(M_i + 1) + K}$ , and (ii) the per-iteration complexity is on the order of  $[(M_i^2 + 1)(U + K) + UM_i^2(M_i^2 + M_i) + M_i^4] M_i^2$ . ■

*Lemma 4:* The computational complexity of Algorithm 4 is on the order of:

$$TN^2 [M_i^2 + NUM_i(1 + UM_i + KM_i)] + TN \log N + NM_i U + NUM_i(1 + UM_i + KM_i) + N \log N. \quad (29)$$

*Proof:* Due to space limitation, we provide main observations to derive (29) as follows. The dominant terms of the computational complexity of Algorithm 4 are at steps 2, 3, 4, 16, 19, and 22. The complexity of generating  $N$  matrices, each matrix of size  $M_i \times U$ , in step 2 is on the order of  $NM_i U$ . The complexity of evaluating each  $\phi_i(\mathbf{W})$  or  $\varphi_k(\mathbf{W})$  is on the order of  $UM_i^2$ , while the complexity of evaluating  $\sum_{i=1}^U \mathbf{w}_i^H \mathbf{w}_i$  is on the order of  $UM_i$ . Hence the complexity of calculating the light density for  $N$  fireflies, i.e., steps 3 and 19, is on the order of  $N(UM_i + U^2 M_i^2 + KUM_i^2) = NUM_i(1 + UM_i + KM_i)$ . The complexity of ranking  $N$  firefly in steps 4 and 22 is  $N \log N$ . Finally, the complexity of moving a firefly in step 16 is on the order of  $M_i^2$ . Assuming a worst case when step 16 is executed in every inner loop of the algorithm, after some manipulations, one can arrive at (29). ■

## V. RECONFIGURABLE INTELLIGENT SURFACE-AIDED BEAMFORMING

### A. Problem Formulation

1) *Problem Formulation:* Consider a communication system comprising of an  $M_i$ -antenna BS communicating with  $U$  single-antenna mobile users in which the direct communication links between the BS and its mobile users are blocked, e.g., because of high building etc., [42]. To circumvent the problem, an  $N_i$ -reflective-element RIS is utilized to support the communication. Let  $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_{N_i}] \in \mathbb{C}^{M_i \times N_i}$  represent the channel coefficients between the BS and the RIS and  $\mathbf{g}_i = [g_{i1}, \dots, g_{iN_i}]^T \in \mathbb{C}^{N_i \times 1}$  be the channel coefficients between the RIS and the  $i$ -th user.

Let  $x_i$ , i.e.,  $\mathbb{E}[|x_i|^2] = 1$ , and  $\mathbf{w}_i \in \mathbb{C}^{M_i \times 1}$ , respectively, represent the data symbol and the active beamforming vector

for the  $i$ -th user. Each reflective element of the RIS generates a phase shift to support the communication between the BS and the mobile users. Let  $\theta_k$  be the phase shift at the  $k$ -th reflective element and let  $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_{N_i}]^T$  denote the phase-shift coefficients generated by the RIS with  $|\theta_k| \leq 1$  and  $\arg(\theta_k) \in [-\pi, \pi], \forall k = 1, \dots, N_i$ . Vector  $\boldsymbol{\theta}$  is the passive beamforming vector for the RIS. The signal arrived at the  $i$ -th user is:

$$\begin{aligned} y_i &= \mathbf{g}_i^H \text{diag}(\boldsymbol{\theta})^H \mathbf{H}^H \mathbf{w}_i x_i + \mathbf{g}_i^H \text{diag}(\boldsymbol{\theta})^H \mathbf{H}^H \sum_{j=1, j \neq i}^U \mathbf{w}_j x_j + n_i, \\ &= \boldsymbol{\theta}^H \mathbf{G}_i^H \mathbf{w}_i x_i + \boldsymbol{\theta}^H \mathbf{G}_i^H \sum_{j=1, j \neq i}^U \mathbf{w}_j x_j + n_i, \end{aligned} \quad (30)$$

where  $\mathbf{G}_i^H = \text{diag}(\mathbf{g}_i^*) \mathbf{H}^H \in \mathbb{C}^{N_i \times M_i}$  and  $n_i \sim \mathcal{CN}(0, \sigma^2)$  represents the additive noise measured at the  $i$ -th user. Furthermore, let  $\{\mathbf{w}_i\} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_U\}$  denote the set of active beamforming vectors, and  $\text{SINR}_i(\{\mathbf{w}_i\}, \boldsymbol{\theta})$  be the SINR at the  $i$ -th user. One can write:

$$\text{SINR}_i(\{\mathbf{w}_i\}, \boldsymbol{\theta}) = \frac{|\boldsymbol{\theta}^H \mathbf{G}_i^H \mathbf{w}_i|^2}{\sum_{j=1, j \neq i}^U |\boldsymbol{\theta}^H \mathbf{G}_i^H \mathbf{w}_j|^2 + \sigma_i^2}. \quad (31)$$

The optimization is posed as follows:

$$\begin{aligned} \min_{\{\mathbf{w}_i\}, \boldsymbol{\theta}} \quad & \sum_{i=1}^U \mathbf{w}_i^H \mathbf{w}_i \\ \text{s. t.} \quad & \text{SINR}_i(\{\mathbf{w}_i\}, \boldsymbol{\theta}) \geq \eta_i, \forall i, \\ & |\theta_k| \leq 1, \forall k, \end{aligned} \quad (32)$$

where  $\eta_i$  is the required SINR level measured at the  $i$ -th user. Since the SINR constraint is a function of two optimization variables  $\mathbf{w}_i$  and  $\boldsymbol{\theta}$ , problem (32) is non-convex.

2) *Alternative Optimization Approach:* For the sake of completeness, the widely-adopted AO approach [21], [42]–[44] is represented here as a baseline to solve (32). Let  $\mathbf{F}_i = \mathbf{w}_i \mathbf{w}_i^H$ , and  $\boldsymbol{\Theta} = \boldsymbol{\theta} \boldsymbol{\theta}^H$ , i.e.,  $\text{rank}(\mathbf{F}_i) = 1$  and  $\text{rank}(\boldsymbol{\Theta}) = 1$ . As  $\mathbf{F}_i$  and  $\boldsymbol{\Theta}$  are two independent variables, they can be alternatively solved [21], [42]–[44]. To that end, relaxing the rank-one constraint on  $\mathbf{F}_i$  and beginning with any initial value of the reflecting coefficient matrix  $\boldsymbol{\Theta}^{(0)}$ , the following sub-problem will be solved at the  $p$ -th iteration:

$$\begin{aligned} \min_{\{\mathbf{F}_i\}} \quad & \text{Tr} \left( \sum_{i=1}^U \mathbf{F}_i \right) \\ \text{s. t.} \quad & \text{Tr} \frac{\mathbf{G}_i \boldsymbol{\Theta}^{(p-1)} \mathbf{G}_i^H \mathbf{F}_i}{\eta_i \sigma_i^2} - \sum_{j=1, j \neq i}^U \text{Tr} \frac{\mathbf{G}_i \boldsymbol{\Theta}^{(p-1)} \mathbf{G}_i^H \mathbf{F}_j}{\sigma_i^2} - 1 \geq 0, \forall i, \\ & \mathbf{F}_i \geq \mathbf{0}, \forall i \in \{1, \dots, U\}. \end{aligned} \quad (33)$$

The reflecting coefficients  $\boldsymbol{\Theta}^{(p)}$  is then updated from the optimal solution of (33) at  $p$ -th iteration, i.e.,  $\{\mathbf{F}_i^{(p)}\}$ , by solving

the following sub-problem [42]:

$$\begin{aligned}
 & \min_{\Theta} \quad \text{Tr}(\Theta) \\
 & \text{s. t.} \quad \text{Tr} \frac{\Theta \mathbf{G}_i^H \mathbf{F}_i^{(p)} \mathbf{G}_i}{\eta_i \sigma_i^2} - \sum_{j=1, j \neq i}^U \text{Tr} \frac{\Theta \mathbf{G}_i^H \mathbf{F}_j^{(p)} \mathbf{G}_i}{\sigma_i^2} - 1 \geq 0, \forall i, \\
 & \quad \text{diag}(\text{diag}(\Theta)) \leq \mathbf{I}_N, \\
 & \quad \Theta \geq \mathbf{0}.
 \end{aligned} \tag{34}$$

The AO approach repetitively solves two SDPs (33) and (34) in  $n_0$  iterations to obtain the solution for (32).

*Remark 1:* It is worth noticing that the AO approach approximates the originally non-convex optimization (32) by two sub-problems (33) and (34). Although (33) and (34) are convex, the solutions to these sub-problems can be regarded as the upper bounds of the original problem (32) as these solutions may not be the global solution. Furthermore, the AO approach adopts the so-called semidefinite relaxation technique [20] in which the rank-one constraints on  $\mathbf{F}_i$  and  $\Theta$  are relaxed. If solving (33) and/or (34) does not return rank-one matrices  $\mathbf{F}_i$  and/or  $\Theta$ , then a rank-one approximation or a Gaussian randomize procedure [19] is required to extract approximated rank-one solutions. Extracting the approximated solutions requires further computational resources yet only results in sub-optimal solutions.

Motivated by the above observations, we introduce a novel FA approach to simultaneously solve  $\mathbf{w}_i$  and  $\theta$  for the original problem (32) in the following section.

### B. Proposed Firefly Algorithm

The optimization (32) can be expressed as

$$\begin{aligned}
 & \min_{\{\mathbf{W}, \theta\}} \quad f(\mathbf{W}) \\
 & \text{s. t.} \quad \phi_i(\{\mathbf{W}, \theta\}) \leq 0, \forall i, \\
 & \quad \varphi_k(\theta_k) \leq 0, \forall k,
 \end{aligned} \tag{35}$$

where  $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_U] \in \mathbb{C}^{M_t \times U}$ ,  $f(\mathbf{W}) = \sum_{i=1}^U \mathbf{w}_i^H \mathbf{w}_i$ ,

$$\begin{aligned}
 \phi_i(\mathbf{W}, \theta) = \eta_i & \frac{\sum_{j=1}^U \mathbf{w}_j^H \mathbf{G}_i \theta \theta^H \mathbf{G}_i^H \mathbf{w}_j}{\sigma_i^2} + \eta_i \\
 & - (1 + \eta_i) \frac{\mathbf{w}_i^H \mathbf{G}_i \theta \theta^H \mathbf{G}_i^H \mathbf{w}_i}{\sigma_i^2},
 \end{aligned} \tag{36}$$

and  $\varphi_k(\theta_k) = |\theta_k| - 1$ . Adopting the penalty method, (35) can be written as:

$$\min_{\{\mathbf{W}, \theta\}} \quad f(\mathbf{W}) + P(\mathbf{W}, \theta), \tag{37}$$

where  $P(\mathbf{W}, \theta)$  is the penalty term given as:

$$P(\mathbf{W}, \theta) = \sum_{i=1}^U \lambda_i \max\{0, \phi_i(\{\mathbf{W}, \theta\})\}^2 + \sum_{k=1}^{N_t} \rho_k \max\{0, \varphi_k(\theta_k)\}^2, \tag{38}$$

with  $\lambda_i > 0$  and  $\rho_k > 0$  are penalty constants.

Let  $\{\mathbf{W}_t, \theta_t\} = \{\mathbf{w}_1^t, \mathbf{w}_2^t, \dots, \mathbf{w}_U^t\}, \theta_t$  be the firefly  $t$ . We initialize a population of  $N$  fireflies  $\{\mathbf{W}_t, \theta_t\}$ ,  $t \in \{1, 2, \dots, N\}$  and define the light density, i.e., the brightness, of the firefly  $t$   $\{\mathbf{W}_t, \theta_t\}$  as:

$$I_t(\mathbf{W}_t, \theta_t) = \frac{1}{f(\mathbf{W}_t) + P(\mathbf{W}_t, \theta_t)}. \tag{39}$$

For any fireflies  $t$  and  $l$  amongst the population, if  $I_t(\mathbf{W}_t, \theta_t) > I_l(\mathbf{W}_l, \theta_l)$  then the firefly  $l$  will move toward the firefly  $t$  as:

$$\begin{aligned}
 \mathbf{W}_l^{(n+1)} &= \mathbf{W}_l^{(n)} + \beta_0 e^{-\gamma(r_{w,tl}^{(n)})^2} (\mathbf{W}_t^{(n)} - \mathbf{W}_l^{(n)}) + \alpha^{(n)} \mathbf{V}, \tag{40} \\
 \theta_l^{(n+1)} &= \theta_l^{(n)} + \beta_0 e^{-\gamma(r_{\theta,tl}^{(n)})^2} (\theta_t^{(n)} - \theta_l^{(n)}) + \alpha^{(n)} \mathbf{v}, \tag{41}
 \end{aligned}$$

where  $r_{w,tl}^{(n)} = \|\mathbf{W}_t^{(n)} - \mathbf{W}_l^{(n)}\|$  and  $r_{\theta,tl}^{(n)} = \|\theta_t^{(n)} - \theta_l^{(n)}\|$  are the Cartesian distances,  $\beta_0$  is the attractiveness at  $r_{w,tl}^{(n)} = 0$  and  $r_{\theta,tl}^{(n)} = 0$ ,  $\gamma$  presents the variation of of the attractiveness. The second terms of (40) and (41) capture the attractions while the third terms of (40) and (41) are randomization comprised of randomization factor  $\alpha^{(n)}$ ,  $\mathbf{V} \in \mathbb{C}^{M_t \times U}$  and  $\mathbf{v} \in \mathbb{C}^{M_t \times 1}$ . The factor  $\alpha^{(n)}$ , the elements of  $\mathbf{V}$  and  $\mathbf{v}$  are drawn from either a uniform or a Gaussian distribution.

It can be observed that problem (35) is a special case of the proposed framework (1) where the objective and constraints are functions of optimization variables  $\mathbf{W}$  and  $\theta$ . The proposed FA for RIS has the same steps as those in Algorithm 1 except steps 3, 16, 18 and 19 given in Algorithm 5.

---

### Algorithm 5 Modified generalized FA for solving (32)

---

**Input:** FA parameters:  $N, T, \lambda_i, \rho_n, \beta_0, \gamma$ ; Optimization data:  $\mathbf{H}, \mathbf{g}_i, \sigma_i^2, \eta_i, I_{t_0}$ ;  
 Step 3: Evaluate the light intensities of  $N$  fireflies as (39);  
 Step 16: Move firefly  $i$  towards firefly  $j$  as (40) and (41);  
 Step 18: Attractiveness varies with distances via  $e^{-\gamma(r_{w,ji}^{(n)})^2}$  and  $e^{-\gamma(r_{\theta,ji}^{(n)})^2}$ ;  
 Step 19: Evaluate new solutions; update  $I_i(\mathbf{W}_i, \theta_i)$  as (39);  
**return**  $\mathbf{W}^*, \theta^*$ .

---

### C. Complexity Analysis

Here, we analyze the computational complexities of the AO and the proposed FA for RIS-aided beamforming problem.

*Lemma 5:* The complexity of the AO approach is on the order of:

$$n_o (\tau_1 + \tau_2), \tag{42}$$

where

$$\begin{aligned}
 \tau_1 &= \ln(\varepsilon^{-1}) \sqrt{U(M_t + 1)} [(M_t^2 + 1)U + UM_t^2(M_t^2 + M_t) \\
 & \quad + M_t^4] M_t^2,
 \end{aligned} \tag{43}$$

$$\tau_2 = \ln(\varepsilon^{-1}) \sqrt{U + 2N_t} [(N_t^2 + 1)(U + 2N_t^2) + N_t^4] N_t^2 \tag{44}$$

*Proof:* We first give some hints to derive the computational complexity of obtaining optimal solution to problems (33) and (34). With the observation that (33) has  $U$  LMI constraints of size 1 and  $U$  LMI constraints of size  $M_t$ , one can follow the same steps as in [41, Section V-A] to derive the complexity of solving (33) as  $\tau_1$  given in (43).

At a given  $\varepsilon > 0$ ,  $\Theta^\varepsilon$  is called an  $\varepsilon$ -solution to problem (34) if  $\text{Tr}(\Theta^\varepsilon) \leq \min \text{Tr}(\Theta) + \varepsilon$ . The number of decision variables of



(34) is  $N_t^2$ . Observing that (34) has  $U$  linear-matrix-inequality (LMI) constraints of size 1 and 2 LMI constraints of size  $N_t$ , one can derive the computational complexity to attain  $\varepsilon$ -solution to (34) as the order of  $\tau_2$  given in (44).

Since the AO approach iteratively solves (33) and (34) in  $n_o$  iterations, the complexity of AO approach is on the order of  $n_o(\tau_1 + \tau_2)$ . ■

*Lemma 6:* The computational complexity of Algorithm 5 is on the order of

$$\begin{aligned} & TN^2 \left[ M_t^2 + N_t + N(UM_t + U(N_t^2 + M_t N_t) + N_t) \right] \\ & + TN \log N + NM_t U + N_t N + N \log N \\ & + N(UM_t + U(N_t^2 + M_t N_t) + N_t). \end{aligned} \quad (45)$$

*Proof:* The proof is based on the following observations. The dominant terms of the computational complexity of Algorithm 5 are at steps 2, 3, 4, 16, 19, and 22. The complexity of generating  $N$  fireflies in step 2 is on the order of  $NM_t U + N_t N$ . The complexities of evaluating  $\phi_i(\mathbf{W}, \boldsymbol{\theta})$ ,  $\varphi_k(\theta_k)$ , and  $\sum_{i=1}^U \mathbf{w}_i^H \mathbf{w}_i$  are, respectively, on the order of  $U(N_t^2 + M_t N_t)$ ,  $N_t$ , and  $UM_t$ . Hence, the complexity of calculating the light density for  $N$  fireflies, i.e., steps 3 and 19, is on the order of  $N(UM_t + U(N_t^2 + M_t N_t) + N_t)$ . The complexity of ranking  $N$  firefly in steps 4 and 22 is  $N \log N$ . Finally, the complexity of moving a firefly in step 16 is on the order of  $M_t^2 + N_t$ . Assuming a worst case when step 16 is executed in every inner loop of the algorithm, after some manipulations, one can arrive at (45). ■

## VI. RIS-AIDED WIRELESS POWER TRANSFER

### A. Problem Formulation

1) *Problem Formulation:* Consider a similar communication system in V-A, however, the users are energy harvesting receivers (EHRs) instead of information decoding receivers. Using the same notations as in V-A, the power arrived at the  $i$ -th user is:

$$E_i = \left| \mathbf{g}_i^H \text{diag}(\boldsymbol{\theta})^H \mathbf{H}^H \sum_{j=1}^U \mathbf{w}_j \right|^2 = \sum_{j=1}^U \mathbf{w}_j^H \mathbf{G}_i \boldsymbol{\theta} \boldsymbol{\theta}^H \mathbf{G}_i^H \mathbf{w}_j \quad (46)$$

where  $\mathbf{w}_j$  is the active energy beamforming vector for the  $j$ -th user. We are interested in maximizing a total weighted sum power received at the EHRs obtained via the following optimization problem:

$$\begin{aligned} & \max_{\{\mathbf{w}_i\}, \boldsymbol{\theta}} \sum_{i=1}^U \sum_{j=1}^U \alpha_i \mathbf{w}_j^H \mathbf{G}_i \boldsymbol{\theta} \boldsymbol{\theta}^H \mathbf{G}_i^H \mathbf{w}_j \\ & \text{s. t.} \quad \sum_{j=1}^U \mathbf{w}_j^H \mathbf{w}_j \leq P, |\theta_k| = 1, \forall k, \end{aligned} \quad (47)$$

where  $P$  is the maximum transmit power of the BS and  $\alpha_i \geq 0$  is the weighting factor for the  $i$ -th EHR.

2) *Successive Convex Approximation:* According to [23], for any fix  $\boldsymbol{\theta}$ , only one common energy beam is sufficient. Using a successive convex approximation (SCA) technique, [23] proposed an iterative algorithm to find optimal active and passive beamforming vectors for problem (47) as follows. Starting with an initialized value  $\boldsymbol{\theta}^{(0)}$ , the optimal active

beamforming vector at the  $l$ -th iterations is calculated as  $\mathbf{w}^{(l)} = \sqrt{P} \text{eig}_{\max} \left( \sum_{i=1}^U \alpha_i \mathbf{G}_i \boldsymbol{\theta}^{(l-1)} \boldsymbol{\theta}^{(l-1)H} \mathbf{G}_i^H \right)$  where  $\text{eig}_{\max}(\mathbf{X})$  is the maximum eigenvalue of matrix  $\mathbf{X}$ . The  $k$ -th coefficient of the RIS's phase shift vector at the  $l$ -th iterations is calculated as  $[\boldsymbol{\theta}^{(l)}]_k = 1$  if  $\mu_k = 0$  and  $[\boldsymbol{\theta}^{(l)}]_k = \frac{\mu_k}{|\mu_k|}$  if  $\mu_k \neq 0$ , where  $\mu_k = \left[ \sum_{i=1}^U \alpha_i \mathbf{G}_i^H \mathbf{w}^{(l)} \mathbf{w}^{(l)H} \mathbf{G}_i \boldsymbol{\theta}^{(l-1)} \right]_k$ .

### B. Proposed Firefly Algorithm

The optimization (47) can be expressed as

$$\begin{aligned} & \min_{\{\mathbf{W}, \boldsymbol{\theta}\}} -f(\mathbf{W}, \boldsymbol{\theta}) \\ & \text{s. t.} \quad \phi(\{\mathbf{W}, \boldsymbol{\theta}\}) \leq 0, \\ & \quad \varphi_k(\theta_k) = 0, \forall k, \end{aligned} \quad (48)$$

where  $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_U] \in \mathbb{C}^{M_t \times U}$ ,  $f(\mathbf{W}, \boldsymbol{\theta}) = \sum_{i=1}^U \sum_{j=1}^U \alpha_i \mathbf{w}_j^H \mathbf{G}_i \boldsymbol{\theta} \boldsymbol{\theta}^H \mathbf{G}_i^H \mathbf{w}_j$ ,  $\phi(\mathbf{W}, \boldsymbol{\theta}) = \sum_{j=1}^U \mathbf{w}_j^H \mathbf{w}_j - P$ , and  $\varphi_k(\theta_k) = |\theta_k| - 1$ . Adopting the penalty method, (35) can be written as:

$$\min_{\{\mathbf{W}, \boldsymbol{\theta}\}} -f(\mathbf{W}, \boldsymbol{\theta}) + P(\mathbf{W}, \boldsymbol{\theta}) \quad (49)$$

where  $P(\mathbf{W}, \boldsymbol{\theta}) = \lambda \max\{0, \phi(\{\mathbf{W}, \boldsymbol{\theta}\})\}^2 + \sum_{k=1}^{N_t} \rho_k \{\varphi_k(\theta_k)\}^2$ , with  $\lambda > 0$  and  $\rho_k > 0$  are penalty constants.

Let  $\{\mathbf{W}_t, \boldsymbol{\theta}_t\} = \{[\mathbf{w}_1^t, \mathbf{w}_2^t, \dots, \mathbf{w}_U^t], \boldsymbol{\theta}_t\}$  be the firefly  $t$ . We initialize a population of  $N$  fireflies  $\{\mathbf{W}_t, \boldsymbol{\theta}_t\}$ ,  $t \in \{1, 2, \dots, N\}$  and define the light density, i.e., the brightness, of the firefly  $t$   $\{I_t, \mathbf{W}_t, \boldsymbol{\theta}_t\}$  as:

$$I_t(\mathbf{W}_t, \boldsymbol{\theta}_t) = \frac{1}{-f(\mathbf{W}_t) + P(\mathbf{W}_t, \boldsymbol{\theta}_t)}. \quad (50)$$

It can be observed that problem (48) is a special case of the proposed framework (1) where the objective and constraints are functions of optimization variables  $\mathbf{W}$  and  $\boldsymbol{\theta}$ . Utilizing the firefly movements define in (40) and (41) in Section V-B, the proposed FA for RIS has the same steps as those in Algorithm 1 except steps 3, 16, 18 and 19 given in Algorithm 6.

---

#### Algorithm 6 Modified generalized FA for solving (47)

---

**Input:** FA parameters:  $N, T, \lambda, \rho_k, \beta_0; \gamma$ ; Optimization data:  $\mathbf{H}, \mathbf{g}_i, \alpha_i, P$ ;

Step 3: Evaluate the light intensities of  $N$  fireflies as (50);

Step 16: Move firefly  $i$  towards firefly  $j$  as (40) and (41);

Step 18: Attractiveness varies with distances via  $e^{-\gamma \left(\frac{r_{w,i}^{(m)}}{r_{w,j}^{(m)}}\right)^2}$  and  $e^{-\gamma \left(\frac{r_{\theta,i}^{(m)}}{r_{\theta,j}^{(m)}}\right)^2}$ ;

Step 19: Evaluate new solutions; update  $I_i(\mathbf{W}_i, \boldsymbol{\theta}_i)$  as (50);

**return**  $\mathbf{W}^*, \boldsymbol{\theta}^*$ .

---

### C. Complexity Analysis

Here, we analyze the complexities of the SCA approach and the proposed FA for the RIS-aided WPT beamforming. We start by introducing the following lemma.

*Lemma 7:* The complexity of the SCA approach is on the order of:

$$m_0 \left( UM_t (M_t + N_t) + M_t^3 + M_t \log M_t + N_t^3 + N_t^2 M_t \right), \quad (51)$$

where  $m_0$  is the number of iterations of the SCA approach.

*Proof:* At each iteration, the complexity of evaluating  $\alpha_i \mathbf{G}_i \boldsymbol{\theta}^{(l-1)} \boldsymbol{\theta}^{(l-1)H} \mathbf{G}_i^H$  is on the order of  $U(M_t^2 + M_t N_t)$ . The complexities of finding a maximum eigenvalue of the  $M_t \times M_t$  matrix  $\alpha_i \mathbf{G}_i \boldsymbol{\theta}^{(l-1)} \boldsymbol{\theta}^{(l-1)H} \mathbf{G}_i^H$  based on the SVD method is on the order of  $M_t^3 + M_t \log M_t$ . Hence, the complexity of finding  $\mathbf{w}^{(l)}$  is on the order of  $UM_t(M_t + N_t) + M_t^3 + M_t \log M_t$ . Furthermore, the complexity of calculating  $\mu_k$  is on the order of  $N_t^2 + M_t N_t$ . Therefore, the complexity of finding  $\boldsymbol{\theta}^{(l)}$  is on the order of  $N_t(N_t^2 + M_t N_t)$ . Consequently,  $m_0$  iterations of evaluating  $\mathbf{w}^{(l)}$  and  $\boldsymbol{\theta}^{(l)}$  lead to (51). ■

*Lemma 8:* The complexity of the Algorithm 6 is on the order of:

$$\begin{aligned} & TN^2 \left[ M_t^2 + N_t + N(UM_t + U(N_t^2 + M_t N_t) + N_t) \right] \\ & + TN \log N + NM_t U + N_t N + N \log N \\ & + N(UM_t + U(N_t^2 + M_t N_t) + N_t). \end{aligned} \quad (52)$$

*Proof:* Noticing that the complexities of evaluating  $\phi(\mathbf{W}, \boldsymbol{\theta})$ ,  $\varphi_k(\theta_k)$ , and  $f(\mathbf{W}, \boldsymbol{\theta})$  are, respectively, on the order of  $UM_t$ ,  $N_t$ , and  $U(N_t M_t + N_t^2)$ . One can easily show that the complexity of the Algorithm 6 is the same as that of the Algorithm 5. ■

## VII. NUMERICAL RESULTS

In this section, we perform simulations to evaluate the performances of the proposed FA approaches, i.e., FA approaches for transmit beamforming, cognitive cognitive beamforming, RIS-aided transmit beamforming, and RIS-aided WPT, and compare them with their iterative, SDP, and SCA counterparts.

CVX package [40] is utilized to obtain the solution for the cognitive SPD approach, i.e., problem (21), and the AO approach for the RIS-aided transmit beamforming. In the AO approach, two SDPs (33) and (34) are alternatively solved in  $n_0 = 10$  iterations. The setup parameters for FAs are as follows. The variation of the attractiveness  $\gamma$  is set at 1. The penalty constants are set equal but they dynamically vary as  $\lambda_i = \rho_k = n^2$ ,  $\forall i, k$  where  $n$  is the generation index in Algorithm 1. The attractiveness at zero distance is  $\beta_0 = 1$ . Finally, the initial randomization factor is  $\alpha^{(0)} = 0.9$  and its value at the  $n$ -th generation is  $\alpha^{(n)} = \alpha^{(0)} 0.9^n$ .

### A. Evaluation on Transmit Beamforming

We simulate a scenario of two users, i.e.,  $U = 2$ , randomly distributed within 2 km from their BS. The array antenna gain at the BS is 15dBi. The noise power spectral density, noise figure at each user and the subcarrier bandwidth are, respectively,  $-174$  dBm/Hz, 5 dB and 15 kHz wide. The path loss model is  $35 + 34.5 \log_{10}(l)$ , where  $l$  is in kilometers. A log-normal shadowing with a standard deviation of 8 dB is assumed. Furthermore, a complex Gaussian distribution is set with the variance of 1/2 on each of its real and imaginary components for the downlink channel fading coefficients. Monte Carlo simulations have been carried out over 1000 channel realizations.

Fig. 1 illustrates the total transmit power of the proposed FA approach and its iterative counterpart versus the required SINR level with different numbers of BS's antennas. The

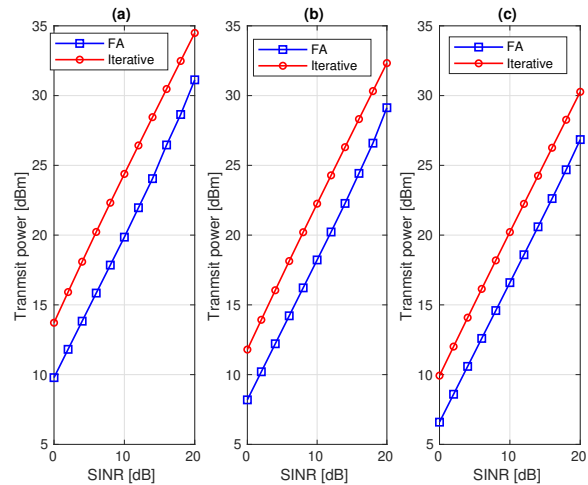


Fig. 1: The total BS's transmit power versus the required SINR level with different numbers of BS's antennas: (a) 4 antennas; (b) 6 antennas; (c) 8 antennas. The firefly population is  $N = 30$ . The number of maximum generations  $T = 30$ .

results on Fig. 1 clearly show that the proposed FA approach outperforms the iterative method in obtaining lower required transmit power, i.e., around 3 to 4 dB lower, for all simulated setups. The results in Fig. 1 confirm the ability of the proposed FA in handling highly nonlinear and multimodal optimization problems. This power saving gain, however, comes at the price of a higher complexity. Using the parameter setup for Fig. 1 in Lemmas 1 and 2, i.e.,  $U = 2$ ,  $T = N = 30$ ,  $M_t = 4, 6, 8$ , one can find the complexities of the Iterative and FA approaches are, respectively, in the order of  $O(10^4)$  and  $O(10^8)$ . When the number of antennas elements are large, letting  $T = N = M_t$ , it can be shown that the dominant terms of the complexities of the Iterative and FA approach are in the order of  $O(M_t^4)$  and  $O(M_t^6)$ , respectively. The trade off between the power saving gain and computational complexity of the proposed FA approach in comparison with the Iterative method should be considered by the network designer/operator.

Fig. 2 shows the total BS's transmit power of the Iterative and proposed FA versus the number of iteration/generations with different numbers of BS's antennas. The results indicate that the Iterative approach converges after just 5 iterations/generations while the proposed FA requires about 20 generations/iterations to level off.

Fig. 3 shows the total BS's transmit power of the proposed FA approach versus the number of population  $N$  with different BS's antenna elements. It can be seen that the observed curves converge after  $N = 30$ . Our simulations indicate that the proposed FA approach performs well with at least 30 fireflies to solve (12) under the investigated SINR range.

### B. Evaluations on Cognitive Transmit Beamforming

We first reproduce the result of the experiment described in Example 1 of [3] to compare the proposed FA approach with the SDP approach. In that experiment, three SUs are located at  $-5^\circ$ ,  $10^\circ$ ,  $25^\circ$ , and two PUs are located at  $30^\circ$  and  $50^\circ$ ,

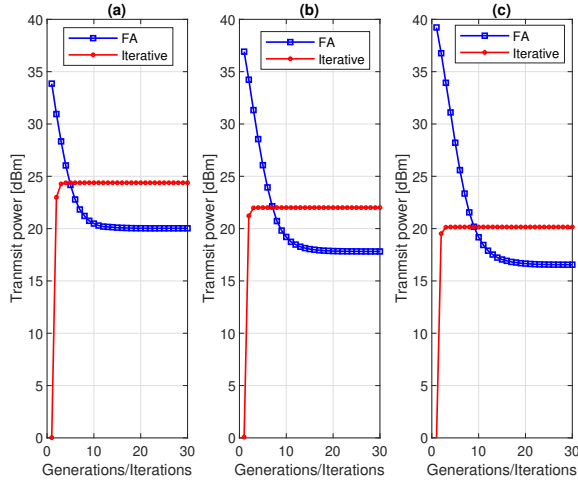


Fig. 2: The total BS's transmit power versus the generations/iteration with different numbers of BS's antennas: (a) 4 antennas; (b) 6 antennas; (c) 8 antennas. The firefly population is  $N = 30$ . The required SINR level at each user is 10 dB.

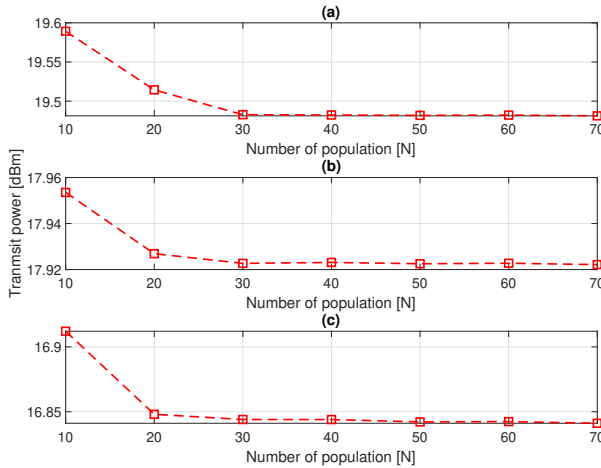


Fig. 3: The total BS's transmit power versus the number of population with different numbers of BS's antennas: (a) 4 antennas; (b) 6 antennas; (c) 8 antennas. The number of generation is  $T = 30$ . The required SINR level at each user is 10 dB.

relative to the BS's array broadside. The tolerable interference level two PUs are  $I_{to,1} = 0.001$  and  $I_{to,2} = 0.0001$ . The noise variance is set to 0.1 while the required SINR values are set to 1 for the SUs.

The channel covariance matrices from the secondary BS to SU  $t$ , i.e.,  $\mathbf{R}_{s,t} = \mathbf{R}(\zeta_{s,t}, \delta_a)$ , and to PU  $k$ , i.e.,  $\mathbf{R}_{p,k} = \mathbf{R}(\zeta_{p,k}, \delta_a)$ , are the function of the angle of departure, i.e.,  $\zeta_{s,t}$  or  $\zeta_{p,k}$ , and the standard deviation of the angular spread, i.e.,  $\delta_a$ . The  $(m, n)$ th entry of  $\mathbf{R}(\zeta, \delta_a)$  is, [20]:

$$e^{\frac{j2\pi\Delta}{\psi}[(n-m)\sin\zeta]} e^{-2\left[\frac{\pi\Delta\delta_a}{\psi}((n-m)\cos\zeta)\right]^2}, \quad (53)$$

where  $\psi$  is the carrier wavelength,  $\sigma_a = 2^\circ$ , and the antenna spacing at the BS is set as  $\Delta = \psi/2$ .

Fig. 4 (a) illustrates the radiation patterns at the BS of the SDP approach as described in (21), which is the reproduction

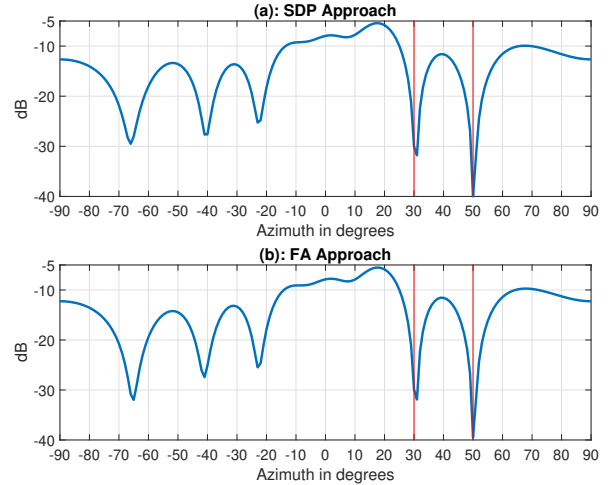


Fig. 4: The radiation pattern of the BS with 8 antennas: (a) The reproduction of [3, Fig. 3]; (b) The proposed FA approach with the number of population  $N = 100$ .

of Fig. 3 in [3], while Fig. 4 (b) shows the radiation patterns at the BS of the FA approach proposed in Algorithm 4. The results clearly indicate that the FA obtains the same radiation pattern as the SDP approach does. Both approaches are able to form nulls to the locations/angles where the PUs are located. In other words, the proposed FA can obtain the same optimal solution as the IPM does for the SDP counterpart. This confirms the ability of the proposed FA in handling highly nonlinear and multimodal optimization problems.

With the setup in Fig. 4, i.e.,  $M_t = 8$ ,  $U = 3$ ,  $K = 2$ ,  $N = 100$  and,  $T = 80$ , one can easily verify from Lemmas 3 and 4 that the proposed FA approach requires higher computational complexity than the SDP approach does when it returns rank-one optimal solution. When the number of antennas is large, one can show that the dominant term of (28) is  $M_t^{0.5}$ . On the other hand, assuming  $T = N = M_t$ , the dominant term of (29) is  $M_t^6$ . Hence, the complexity of an IPM to solve (21) is slightly higher than the complexity of the proposed FA in Algorithm 4, i.e.,  $O(M_t^{6.5})$  in comparison with  $O(M_t^6)$ .

Fig. 5 shows the transmit power of the proposed FA approach versus the number of population with different numbers of transmit antennas. The results indicate that the proposed FA converges with all number of antenna setups as all the observed curves level off after the maximum size of population of  $N = 50$ . However, the higher of the antenna elements is, the larger the size of the population is required for a converged transmit power. For example, with  $M = 8, 16$ , and  $32$ , the proposed FA approach, respectively, obtains a stable transmit power at  $N = 30, 40$  and  $50$ . This is due to the fact that the size of the system increases with a higher number of antenna elements, i.e., a higher degree of freedom. As a result, it requires a larger size of the population to provide a sufficient diversification for the exploration of the FA. The results also show that the required transmit power decreases when the number of antennas increase as the result of having higher degree of freedom.

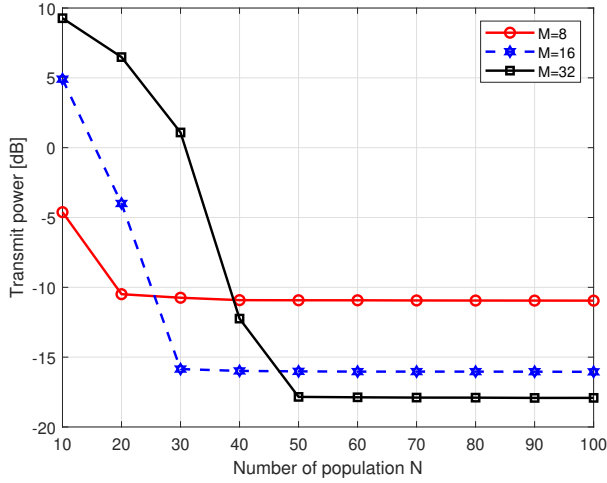


Fig. 5: The total transmit power of the proposed FA approach versus the number of population with different numbers of transmit antennas. The number of maximum generation  $T = 150$ .

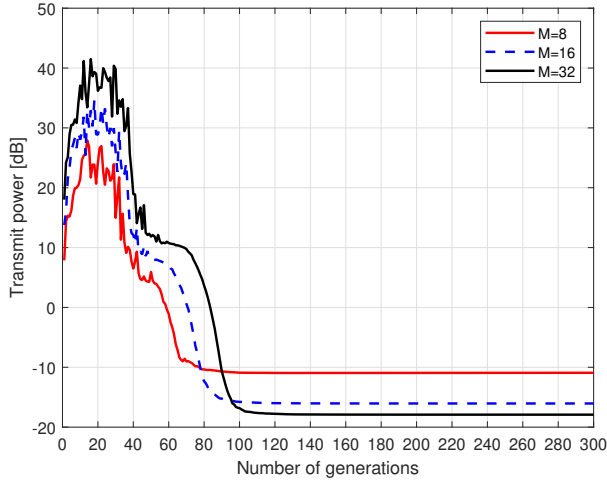


Fig. 6: The total transmit power of the proposed FA approach versus the number of maximum generations with different numbers of transmit antennas. The number of population  $N = 70$ .

Fig. 6 depicts the transmit power of the proposed FA approach versus the number of maximum generations with different numbers of transmit antennas. A similar trend as in Fig. 5 is also observed in this figure. The transmit power attained by the proposed FA approach converges with all numbers of antenna setups. The higher number of antennas is, the higher number of generations is needed as a result of higher exploitation required for the increase of the problem dimension. For instance, the transmit power levels off at around 90, 100, and 120 generations, respectively, for  $M = 8$ , 16, and 32.

### C. Evaluations on RIS-aided Transmit Beamforming

We simulate a RIS-aided communication system which consists of one BS, one RIS, and two users, i.e.,  $U = 2$ . The distance between the BS and the RIS is 10 m. Users

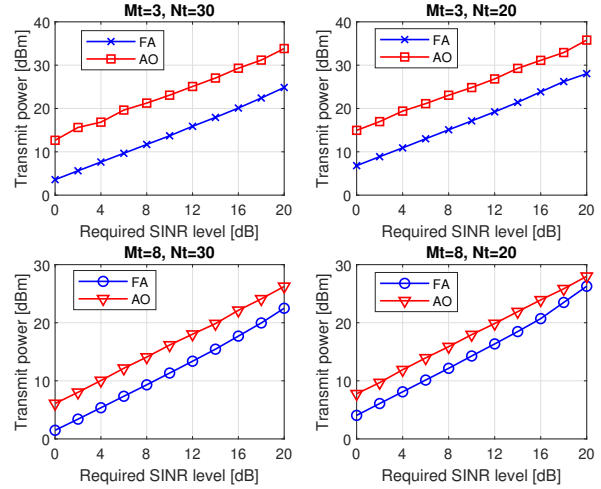


Fig. 7: The total BS's transmit power versus the required SINR level with different numbers of BS's antennas and RIS's reflective elements. The firefly population is  $N = 120$ . The number of maximum generations  $T = 50$ .

are randomly distributed with a distance of 6 m from the RIS. The pathloss exponents of both wireless links from the BS to the RIS and from the RIS to users are set to be 2.2 with the signal attenuation at the reference distance of 1 m being 30 dB [23], i.e., the large-scale fading coefficient is modeled as  $-30 - 22 \log_{10}(d)$  dB where  $d$  is the distance between the BS to RIS or RIS to a user. The noise variance at each user is  $-124$  dBm. Monte Carlo simulations are carried over 100 channel realizations. Each channel realization is associated with a random user location and a random fading coefficient.

Fig. 7 illustrates the total BS's transmit power versus the required SINR level with different numbers of BS's antennas and RIS's reflective elements. The results indicate that the proposed FA prevails the AO approach in terms of lower power consumption. The superior performance of the FA approach over its AO counterpart can be explained as follows. As the AO approach approximates non-convex problem (32) by two convex sub-problems (33) and (34), the solution obtained by the AO approach is not necessary the global optimal solution of the original problem (32). On the other hand, the proposed FA possessing both exploitation and exploration abilities can effectively handle such non-convex problem and obtain much better solution than its counterpart. The results shown on Fig. 7 verify the ability of the proposed FA in handling highly nonlinear and multimodal optimization problems.

It can be observed from Fig. 7 that at a given number of RIS's reflective elements, the performance gap between the proposed FA and the AO decreases when the number of BS's antennas increases. For example, when  $N_t = 20$ , the gaps are, respectively, around 7.5 dB and 3.5 dB with  $M_t = 3$  and  $M_t = 8$ . Fortunately, at a given number of BS's antennas, the performance gap improves when the number of RIS's elements increases. For instance, with  $M_t = 8$ , the performance gap increases from around 3.5 dB to 4.5 dB when  $N_t$  increases from 20 to 30. Interestingly, the FA performs especially well with a relatively high ratio of  $N_t/M_t$ , i.e., the performance gap

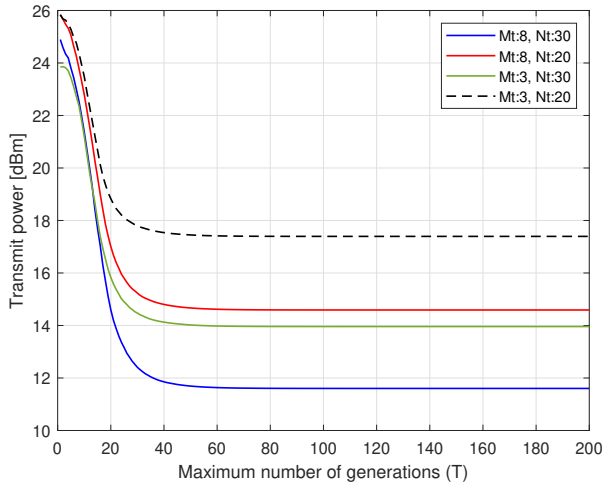


Fig. 8: The total BS's transmit power versus the number of maximum generations with different numbers of BS's antennas and RIS's reflective elements. The firefly population is  $N = 120$ . The required SINR level is 10 dB.

is around 9.5 dB with the ratio of 30/3 while it is around 3.5 with the ratio of 20/8. The results can be explained as follows. A higher number of RIS's reflective elements gives more degree of freedom for the FA to perform. Moreover, the channel between the RIS and these users plays a higher role than that between the BS and the RIS does as the former is closer to these users. Last but not least, the performance gaps slightly decrease at relatively high SINR level especially when the  $N_t/M_t$  ratio is relatively low. For example with the ratio of 20/8, the performance gap is around 1.8 dB at SINR of 20 dB compared with around 3.5 dB at the other SINR levels, i.e., see the bottom-right corner figure of Fig. 7. This is because of a fact that the FA has reached its limit of exploration with  $N = 120$  fireflies, at a stricter constraint condition.

We now compare the computational complexities of the AO and FA approaches for the experiments presented on Fig. 7. As  $N_t$  is larger than  $M_t$ , from Lemma 5 one can show that the dominant term of the complexity of the AO approach is  $n_0 N_t^{6\frac{1}{2}}$ . Similarly, from Lemma 6 one can conclude that the dominant term of the complexity of the FA approach is  $T N^3 N_t^2$ . Substituting for  $N_t = 30$ ,  $n_0 = 10$ ,  $N = 120$  and  $T = 50$ , we can arrive at the fact that the computational complexities of the AO and FA approaches are on the same order of  $\mathcal{O}(10^{10})$ . When the numbers of antennas  $M_t$  and  $N_t$  are large, letting  $N_t = n_0 = M_t$  in (42), one can show that the dominant term of the complexity to attain  $\varepsilon$ -solution to (32) is  $M_t^{7\frac{1}{2}}$ . On the other hand, one can derive the dominant term of (45) as  $M_t^6$  when assuming  $T = N = N_t = M_t$ . Hence, the complexity of an IPM to solve (32) is higher than the complexity of the proposed FA in Algorithm 5, i.e.,  $\mathcal{O}(M_t^{7\frac{1}{2}})$  in comparison with  $\mathcal{O}(M_t^6)$ .

In Fig. 8, the total BS's transmit power is plotted versus the maximum of generation  $T$  used in the FA in Algorithm 5 with different BS's antennas and RIS's reflective elements. The results indicate that the proposed FA requires around 50 to 60

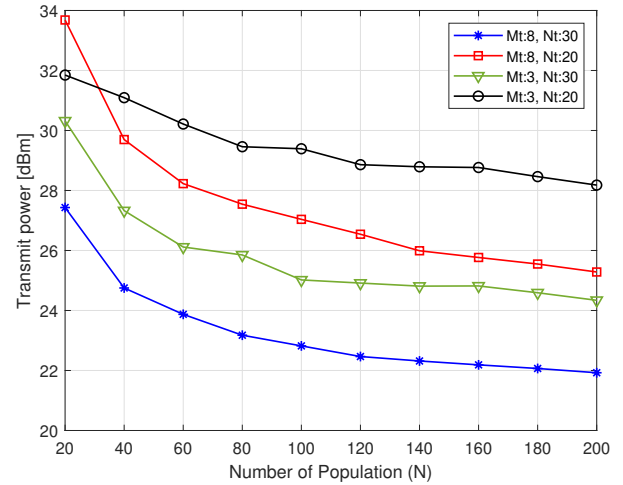


Fig. 9: The total transmit power versus the number of populations with different numbers of BS's antennas and RIS's reflective elements. The number of maximum generations  $T = 50$ . The required SINR level is 20 dB.

generations to attain the optimal solution for all setups.

Fig. 9 illustrates the total transmit power versus the number of population  $N$  with different BS's antennas and RIS's elements. The results show that increasing the size of the firefly population enables the FA to obtain better solution. For example, the total transmit power decreases around 7 dB, 5.4 dB, 5 dB, and 3 dB, respectively, for the setups of  $(M_t = 8, N_t = 20)$ ,  $(M_t = 3, N_t = 30)$ ,  $(M_t = 8, N_t = 20)$ , and  $(M_t = 3, N_t = 20)$  when the firefly population increases from 20 to 120. The performance gap at the 20 dB SINR level observed in Fig. 7 for  $(M_t = 8, N_t = 20)$  can be improved 1 dB further when the population size is enlarged from 120 to 200. These total-transmit-power curves converge after  $N = 180$  as the reduction in the total transmit power is negligible when the population increases to  $N = 200$  for all setups.

#### D. Evaluations on RIS-aided WPT

Here, we use the same setup for the RIS-aided communication system as considered in the previous section, i.e., Section VII-C. However, the EHRs are randomly placed with the distance of 2 m from the RIS. We run  $m_0 = 10$  iterations to obtain the solution for the SCA approach.

Fig. 10 shows the sum-power received at EHRs versus BS's maximum transmit power with different numbers of BS's antennas and RIS's reflective elements. It is clear from the figure that the proposed FA approach outperforms the SCA approach in [23] in offering higher sum-power at EHRs. The performance gaps are, respectively, around 18 dB, 17 dB, 15 dB, and 14 dB for the setups of  $(M_t = 3, N_t = 30)$ ,  $(M_t = 8, N_t = 30)$ ,  $(M_t = 3, N_t = 20)$ , and  $(M_t = 8, N_t = 20)$ . The superior performance of the proposed FA over the SCA is due to the advantage of having exploitation and exploration abilities to handle non-convex optimization problems. On the other hand, the SCA employs the first-order Taylor expansion to approximate the optimization problem resulting in a lower-bounded solution. Furthermore, the FA approach allocates

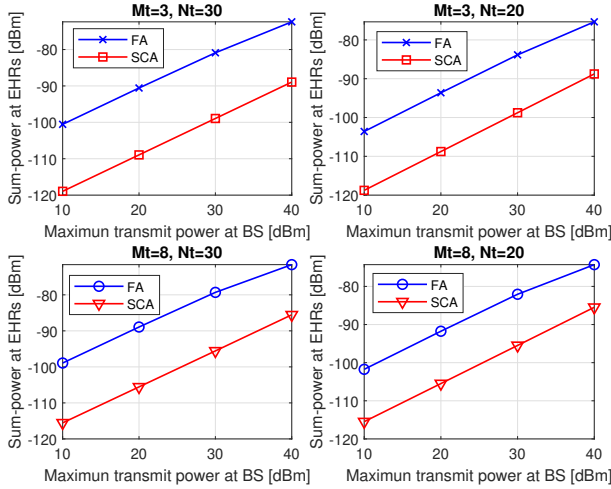


Fig. 10: Sum-power received at EHRs versus BS's maximum transmit power with different numbers of BS's antennas and RIS's reflective elements. The firefly population is  $N = 100$ . The number of maximum generations  $T = 50$ .

one active beamforming vector for each EHR whereas the SCA only uses one active beamforming vector for all EHRs. The results shown on Fig. 10 again verify the ability of the proposed FA in handling highly nonlinear and multimodal optimization problems.

Comparing Figs. 7 and 10, it can be observed that the FA behaves in a similar manner for both power minimization problem (35) and sum-power maximization problem (48). For instance, at the same value of  $M_t$ , the higher the value of  $N_t$ , the larger the performance gap is. At the same value of  $N_t$ , the lower the value of  $M_t$ , the bigger the performance gap is. The results also recommend to maintain a relatively high ratio of  $N_t/M_t$  to attain the best performance of the FA. Slight declines in the performance gaps are also observed at the stricter constraint of BS's transmit power, i.e., 40 dBm, as the FA's population reach their limit of exploration.

We proceed by comparing the computational complexities of the SCA and FA approaches for the experiments shown on Fig. 10. As  $N_t$  is larger than  $M_t$ , from Lemmas 7 and 8, it is clear that the dominant terms of the complexities of the SCA and the FA approaches are, respectively,  $m_0 N_t^3$  and  $T N^3 N_t^2$ . Substituting for  $N_t = 30$ ,  $m_0 = 10$ ,  $N = 100$  and  $T = 50$ , we can arrive at the fact that the computational complexities of the SCA and FA approaches are, respectively, on the orders of  $O(10^5)$  and  $O(10^{10})$ . When the numbers of antennas  $M_t$  and  $N_t$  are large, letting  $N_t = m_0 = M_t$  in (51), one can show that the dominant term of the complexity of the SCA is  $M_t^4$ . On the other hand, the dominant term of (52) is  $M_t^6$  when assuming  $T = N = N_t = M_t$ . Hence, the complexity of the SCA approach is lower than that of the proposed FA in Algorithm 6, i.e.,  $O(M_t^4)$  in comparison with  $O(M_t^6)$ .

Sum-power received at EHRs are shown versus the number of maximum generations with different numbers of BS's antennas and RIS's reflective elements in Fig. 11. The figure reveals that the proposed FA converges after around 50 to 60 generations for all observed setups.

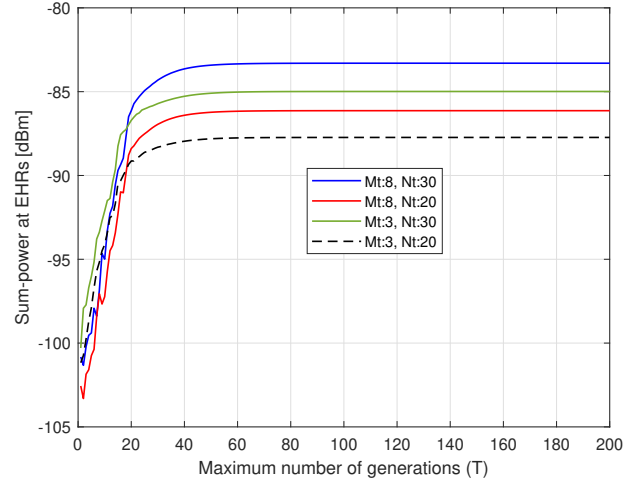


Fig. 11: Sum-power received at EHRs versus the number of maximum generations with different numbers of BS's antennas and RIS's reflective elements. The firefly population is  $N = 100$ . The required SINR level is 10 dB.

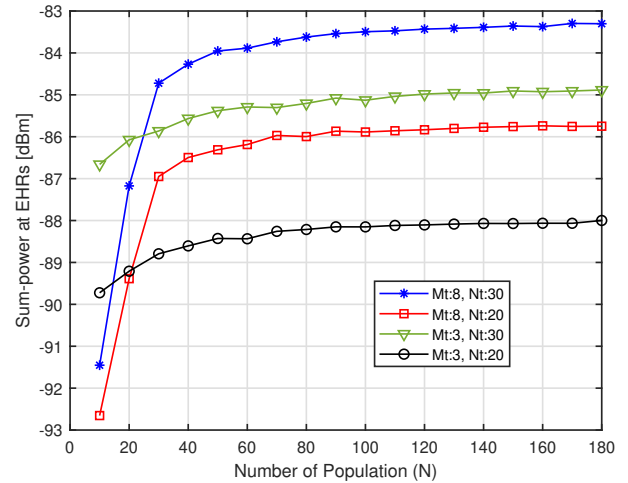


Fig. 12: Sum-power received at EHRs versus the number of populations with different numbers of BS's antennas and RIS's reflective elements. The number of maximum generations  $T = 50$ . The required SINR level is 20 dB.

The effect of the firefly population on the sum-power received at EHRs is illustrated on Fig. 12. The figure shows that all the curves converge after the population size of 80. However the difference between the EHRs' sum-power offered by 80 fireflies and that offered by 40 fireflies is no more than 0.7 dB for all observed setups. This indicates that the complexity of the proposed FA for the RIS-aided WPT sum-power maximization problem in (48) can be reduced with an acceptable tradeoff in the optimality.

## VIII. CONCLUSION

We have proposed a generalized FA to find optimal solution for an optimization framework containing objective function and constraints as multivariate functions of independent optimization variables. We have adopted the proposed generalized

FA to solve four representative examples of classic transmit beamforming, cognitive beamforming, RIS-aided transmit beamforming, and RIS-aided wireless power transfer. Our analyzes have indicated that the computational complexities of proposed FA approaches are less than those of their IPM counterparts, i.e., the SDP and the AO approaches, yet higher than that of the iterative and SCA approaches in large-antenna scenarios. Simulation results have revealed the fact that the proposed FA attains the same optimal solution as the IMP does for the under-investigated cognitive beamforming problem. Interestingly, the proposed FA outperforms the iterative, AO, and SCA approaches for the under-investigated classic transmit beamforming, RIS-aided transmit beamforming, and wireless power transfer problems, respectively. This confirms the effectiveness of the proposed generalized FA in handling multivariate and non-convex problems.

## REFERENCES

- [1] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1437–1450, Oct. 1998.
- [2] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via Conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.
- [3] Y. Huang and D. P. Palomar, "Rank-constrained separable Semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 644–678, Feb. 2010.
- [4] H. Dahrouj and W. Yu, "Multicell interference mitigation with joint beamforming and common message decoding," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2264–2273, Aug. 2011.
- [5] W. Yang and G. Xu, "Optimal downlink power assignment for smart antenna systems," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process., ICASSP '98*, vol. 6, 1998, pp. 3337–3340.
- [6] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.
- [7] D. Hammarwall, M. Bengtsson, and B. Ottersten, "On downlink beamforming with indefinite shaping constraints," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3566–3580, Sep. 2006.
- [8] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure [lecture notes]," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142–148, Jul. 2014.
- [9] G. Zheng, K.-K. Wong, and T.-S. Ng, "Throughput maximization in linear multiuser MIMO-OFDM downlink systems," *IEEE Trans. Veh. Technol.*, vol. 57, no. 3, pp. 1993–1998, May 2008.
- [10] R. Bhagavatula and R. W. Heath, "Adaptive limited feedback for sum-rate maximizing beamforming in cooperative multicell systems," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 800–811, Feb. 2011.
- [11] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664–678, Feb. 2010.
- [12] Y. Huang, Q. Li, W.-K. Ma, and S. Zhang, "Robust multicast beamforming for spectrum sharing-based cognitive radios," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 527–533, Jan. 2012.
- [13] B. Clerckx, R. Zhang, R. Schober, D. W. K. Ng, D. I. Kim, and H. V. Poor, "Fundamentals of wireless information and power transfer: From RF energy harvester models to signal and system designs," *IEEE J. Sel. Areas in Commun.*, vol. 37, no. 1, pp. 4–33, Jan. 2019.
- [14] D. W. K. Ng, E. S. Lo, and R. Schober, "Robust beamforming for secure communication in systems with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4599–4615, Aug. 2014.
- [15] —, "Wireless information and power transfer: Energy efficiency optimization in OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 6352–6370, Dec. 2013.
- [16] T. A. Le, Q.-T. Vien, H. X. Nguyen, D. W. K. Ng, and R. Schober, "Robust chance-constrained optimization for power-efficient and secure SWIPT systems," *IEEE Trans. Green Commun. and Netw.*, vol. 1, no. 3, pp. 333–346, Sep. 2017.
- [17] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, Jun. 2007.
- [18] T. A. Le and K. Navaie, "Downlink beamforming in underlay cognitive cellular networks," *IEEE Trans. Commun.*, vol. 62, no. 7, pp. 2212–2223, Jul. 2014.
- [19] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.
- [20] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using Semidefinite optimization," in *Proc. 37th Annu. Allerton Conf. Commun., Control, and Comput.*, 1999, pp. 987 – 996.
- [21] Z. Peng, Z. Chen, C. Pan, G. Zhou, and H. Ren, "Robust transmission design for RIS-aided communications with both transceiver hardware impairments and imperfect CSI," *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 528–532, Mar. 2022.
- [22] S. Gong, C. Xing, P. Yue, L. Zhao, and T. Q. S. Quek, "Hybrid analog and digital beamforming for RIS-assisted mmWave communications," *IEEE Trans. Wireless Commun.*, vol. 22, no. 3, pp. 1537–1554, Mar. 2023.
- [23] Q. Wu and R. Zhang, "Weighted sum power maximization for intelligent reflecting surface aided SWIPT," *IEEE Wireless Commun. Letters*, vol. 9, no. 5, pp. 586–590, May 2020.
- [24] X.-S. Yang, *Nature-Inspired Metaheuristic Algorithms*. Luniver Press, 2008.
- [25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

- [26] X.-S. Yang, *Engineering optimisation: an introduction with metaheuristic applications*. Wiley, 2009.
- [27] —, “Chapter 13: Firefly algorithm: Variants and applications,” in *Swarm Intelligence Algorithms*. CRC Press, 2020, pp. 175–186.
- [28] T. Yamanaka and K. Higuchi, “Transmitter beamforming control based on firefly algorithm for massive MIMO systems with per-antenna power constraint,” in *2017 23rd Asia-Pacific Conf. Commun. (APCC)*, 2017, pp. 1–6.
- [29] T. A. Le and X.-S. Yang, “Firefly algorithm for beamforming design in RIS-aided communications systems,” in *Proc. IEEE Veh. Techno. Conf. (VTC 2023-Spring)*, Jun. 2023, pp. 1–5.
- [30] I. Fister, I. F. Jr., X.-S. Yang, and J. Brest, “A comprehensive review of firefly algorithms,” *Swarm and Evolutionary Computation*, vol. 13, pp. 34–46, Dec. 2013.
- [31] X.-S. Yang and X.-S. He, *Why the Firefly Algorithm Works?* Cham: Springer International Publishing, 2018, pp. 245–259.
- [32] W. Windarto and E. Eridani, “Comparison of particle swarm optimization and firefly algorithm in parameter estimation of Lotka-Volterra,” *AIP Conference Proceedings*, vol. 2268, no. 1, p. 050008, 09 2020.
- [33] R. Ezzeldin, M. Zelenakova, H. F. Abd-Elhamid, K. Pietrucha-Urbanik, and S. Elabd, “Hybrid optimization algorithms of firefly with ga and pso for the optimal design of water distribution networks,” *Water*, vol. 15, no. 10, 2023.
- [34] M. Clerc and J. Kennedy, “The particle swarm - explosion, stability, and convergence in a multidimensional complex space,” *IEEE Trans. Evolutionary Computation*, vol. 6, no. 1, pp. 58–73, 2002.
- [35] D. Bertsimas and J. Tsitsiklis, “Simulated annealing,” *Statistical Science*, vol. 8, no. 1, pp. 10–15, 1993.
- [36] X.-S. Yang, *Cuckoo Search and Firefly Algorithm: Theory and Applications*. Studies in Computational Intelligence, 2014.
- [37] E. Osaba, X.-S. Yang, F. Diaz, E. Onieva, A. D. Masegosa, and A. Perallos, “A discrete firefly algorithm to solve a rich vehicle routing problem modelling a newspaper distribution system with recycling policy,” *Soft Computing*, vol. 21, pp. 5295–5308, 2017.
- [38] H. Dahrouj and W. Yu, “Coordinated beamforming for the multicell multi-antenna wireless system,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1748–1759, May 2010.
- [39] T. A. Le and M. R. Nakhai, “Downlink optimization with interference pricing and statistical CSI,” *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2339–2349, Jun 2013.
- [40] CVX Research Inc., “CVX: Matlab software for disciplined convex programming, academic users,” <http://cvxr.com/cvx>, 2015.
- [41] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, “Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization,” *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5690–5705, Nov. 2014.
- [42] T. A. Le, T. V. Chien, and M. D. Renzo, “Robust probabilistic-constrained optimization for IRS-Aided MISO communication systems,” *IEEE Wireless Commun. Lett.*, vol. 10, no. 1, pp. 1–5, Jan. 2021.
- [43] H. Yu, H. D. Tuan, A. A. Nasir, T. Q. Duong, and H. V. Poor, “Joint design of reconfigurable intelligent surfaces and transmit beamforming under proper and improper gaussian signaling,” *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2589–2603, Nov. 2020.
- [44] N. S. Perović, L.-N. Tran, M. D. Renzo, and M. F. Flanagan, “Optimization of RIS-aided MIMO systems via the cutoff rate,” *IEEE Wireless Communications Letters*, vol. 10, no. 8, pp. 1692–1696, Aug. 2021.



**Tuan Anh Le (S'10-M'13-SM'19)** received the Ph.D. degree in telecommunications research from King's College London, The University of London, U.K., in 2012. He was a Post-Doctoral Research Fellow with the School of Electronic and Electrical Engineering, University of Leeds, Leeds, U.K.

He is a Senior Lecturer at Middlesex University, London, U.K. His current research interests include integrated sensing and communication (ISAC), RIS-aided communication, RF energy harvesting and wireless power transfer, physical-layer security, nature-inspired optimization, and applied machine learning for wireless communications. He served as a Technical Program Chair for 26th International Conference on Telecommunications (ICT 2019). He was an Exemplary Reviewer of IEEE Communications Letters in 2019.



**Xin-She Yang** obtained his DPhil in Applied Mathematics from the University of Oxford. He then worked at Cambridge University and National Physical Laboratory (UK) as a Senior Research Scientist. Now he is Reader at Middlesex University London, and a co-Editor of the Springer Tracts in Nature-Inspired Computing. He is also an elected Fellow

of the Institute of Mathematics and its Applications. He was the IEEE Computational Intelligence Society (CIS) chair for the Task Force on Business Intelligence and Knowledge Management (2015 to 2020). He has published more than 300 peer-reviewed research papers with more than 84,000 citations, and he has been on the prestigious list of highly-cited researchers (Web of Sciences) for eight consecutive years (2016-2023).