

Cross-Entropy Method for Electromagnetic Optimization With Constraints and Mixed Variables

Maria Kovaleva, *Student Member, IEEE*, David Bulger, Basit Ali Zeb, *Member, IEEE*,
and Karu P. Esselle, *Fellow, IEEE*

Abstract—An elegant and simple approach is presented for electromagnetic (EM) optimizations, especially when mixed variables and/or constraints are involved. In mixed-variable optimization, some variables are continuous (can take any value within a range) and others are discrete (can take only values from a database). An example constraint is when the total length of a device under optimization is specified. Our approach can handle such optimization problems and is based on an abstract probabilistic evolutionary optimization algorithm, called the cross-entropy (CE) method. We believe that this is the first application of CE with full-wave EM simulations. A quick performance benchmarking on two test functions was performed to compare convergence of CE and two other established optimization algorithms. Then, the advantages of the CE method when simultaneously optimizing a mix of discrete and continuous variables and imposing geometric constraints are illustrated. Finally, six resonant cavity antennas (RCAs) were optimized, and one was prototyped and tested to verify predicted results. This one-layer-superstrate RCA prototype has a measured peak directivity of 17.6 dBi with a 3 dB directivity bandwidth of 51% and lower sidelobes, outperforming all such prototypes in the literature.

Index Terms—Bandwidth, covariance matrix adaptation (CMA), covariance matrix adaptation evolutionary strategies (CMA-ES), cross-entropy (CE), directivity, electromagnetic (EM), evolutionary, Fabry–Perot cavity antenna, high-frequency, high-gain, microwave, particle swarm optimization (PSO), radio frequency (RF), resonant cavity antenna (RCA), wideband.

I. INTRODUCTION

PROGRESS in evolutionary optimization has led to a dramatic change in modern design approaches in the field of electromagnetics (EM). The areas where optimization methods have advanced our knowledge of EM design problems include microstrip antennas [1]–[3], antenna arrays of arbitrary shapes [4]–[6], and electromagnetic bandgap (EBG) structures and EBG-based antennas [7]–[9]. Manual design methods became unfeasible, which led to the active application of optimization methods in EM engineering. Since

mid-1990s, when genetic algorithms (GAs) were studied for application to EM design problems [10]–[12], the number of optimization methods known to the EM community has significantly increased [13]–[16]. The collection of evolutionary optimization algorithms, or broadly speaking computational intelligence area, is evolving so rapidly that decades of research will be needed in order to utilize its potential in EM engineering [17].

Electronic and EM engineering design problems are often constrained and include discrete parameters. This is because the designs are frequently realized using commercially available parts, e.g., dielectric materials available in only specific thickness and permittivity values and resistors available in specific resistance values. Overall dimensions of many modern designs are constrained by space limitations. Therefore, optimization methods are often required to handle mixed-variable (continuous and discrete) design parameters and constraints. In terms of design parameters, GAs and ant-colony optimization (ACO) handle *discrete* variables intrinsically, while most of the other algorithms, including popular particle swarm optimization (PSO), differential evolution (DE), cuckoo search, invasive weed optimization, and evolutionary strategies (ES), were originally developed for problems with a *continuous* domain. In these methods, the constraints are usually handled outside the algorithm update procedures by special restrictive approaches, which lead to inefficient optimization schemes. Recently, the demand for practical mixed-variable optimization has been recognized, and hybrid real-binary techniques have been developed, such as hybrid real-binary PSO [18], mixed-integer GA [19], mixed-integer covariance matrix adaptation ES (CMA-ES) [20], and hybrid real-binary DE [21].

In this paper, we present a new, elegant, yet simple approach for the optimization of EM structures with mixed variables and constraints. The method is based on a cross-entropy (CE) algorithm introduced by Rubinstein and Kroese [22] in 1997. Its first application to EM was in synthesizing antenna arrays using closed-form expressions in 2008 [23], and a little research has been done since then on its application to other EM design problems [24]–[28], including those requiring numerical simulations. The advantage of the CE approach is its adaptive update procedure, which makes it inherently capable of optimizing mixed-variable and constrained problems in a much simpler way than hybrid real-binary techniques.

Manuscript received September 10, 2016; revised April 27, 2017; accepted July 26, 2017. Date of publication August 17, 2017; date of current version October 5, 2017. This work was supported in part by the Australian Government Research Training Program Scholarship and in part by the Australian Research Council. (*Corresponding author: Maria Kovaleva.*)

The authors are with the Centre for Collaboration in Electromagnetics and Antenna Engineering, Department of Engineering, Faculty of Science and Engineering, Macquarie University, Sydney, NSW 2109, Australia (e-mail: maria.kovaleva@students.mq.edu.au; karu@ieee.org).

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

Digital Object Identifier 10.1109/TAP.2017.2740974

Previously, only Gaussian and Bernoulli sampling distribution families were implemented in EM applications of the CE method, and the advantages of using other distributions were not explored. We have interfaced CE with a full-wave EM solver in order to study the performance of a new class of compact resonant cavity antennas (RCAs), which employ a single-slab all-dielectric superstrate with a transverse permittivity gradient (TPG) [29]. To the best of our knowledge, this is the first time the CE method has been applied with full-wave simulations. This is also the first time it has been used to optimize mixed-variable constrained EM problems.

This paper is organized as follows. In Section II, the concept of the CE method is briefly described, and the algorithm is verified by applying to two classic benchmark problems. The computational performance of CE is compared with PSO and CMA-ES to assess its effectiveness. In Section III, one example application of CE to a real-world antenna problem, i.e., an optimal RCA with a specified diameter, is exploited. This optimization example demonstrates the capabilities of the new method, because it involves discrete and continuous design parameters with constraints. One of the optimized antennas was prototyped, and the measurement results are presented in Section IV. Section V concludes this paper.

II. CROSS ENTROPY METHOD

The CE algorithm is a stochastic optimization technique based on minimizing the CE (or Kullback–Leibler divergence) between probability distributions to solve difficult multimodal, multiobjective optimization problems. The idea of the optimizer evolved from a method to estimate the probability of rare events. Detailed explanations and possible modifications of the CE method can be found in the literature [30]. Recent studies noted some similarities of the CE approach with ACO and estimation of distribution algorithms, all of which belong to a *model-based search paradigm*. These methods differ from traditional heuristics in the way that they adjust the solution reproducing mechanism instead of manipulating solutions. Recent stochastic runtime analyses show excellent runtime results and asymptotic convergence properties of the CE method [31], [32].

A. Basic Principle of the CE Method

In statistics, the probability distribution of a random variable x is a description of the probability of each possible value, or range of values, that x could take. This probability is specified by the probability distribution function (pdf) $p(x; v)$, where v is a vector of shape parameters identifying the distributions within the family. For instance, the well-known normal (Gaussian) distribution of a random variable $x \in \mathbb{R}$ is characterized by its mean μ and variance σ^2 , and its pdf is expressed as

$$p(x; v) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad v = (\mu, \sigma^2). \quad (1)$$

In each iteration, CE produces new solutions by sampling from the pdf. In any specific implementation of the CE method, only a finite-dimensional family of probability distributions is used, and it is often called a *model*.

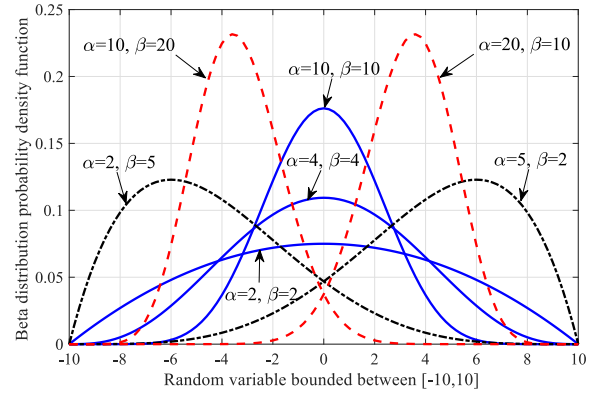


Fig. 1. Beta distribution pdf of a random variable bounded between $[-10, 10]$ for various shape parameters (α, β) .

CE aims to produce a sequence of sampling distributions that are increasingly concentrated around the optimal design. The choice of a probability distribution family in the CE method strongly depends on the nature of the design variables. For optimization problems with continuous variables, normal and beta distribution families can be used to generate possible solutions. Beta distribution, which has pdf given by

$$p(x; v) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx}, \quad v = (\alpha, \beta) \quad (2)$$

is particularly useful in describing continuous bounded variables. Fig. 1 shows how shape parameters affect the pdf of a variable enclosed in an interval $[-10, 10]$. For optimization problems with discrete variables, such discrete probability distribution families as binomial, Poisson, Bernoulli, discrete uniform, and geometric can be applied. For optimization problems with mixed variables, consisting of both discrete and continuous parameters, a combination of appropriate distribution families can be used. Thus, optimization of mixed-variable problems is intrinsic in the CE method.

Let us consider an engineering optimization problem where we wish to determine the values of several *design variables* in order to optimize some user-defined performance measure called the *fitness function*. As any other evolutionary optimization method, the CE evaluates batches of designs at a time, called *populations*. Each population consists of N designs, called *candidates*. Each candidate is characterized by the vector of design variables $X = (x_1, x_2, \dots, x_d)$, where d is the number of variables or the problem dimensionality. Each candidate should be evaluated, usually by simulation, to calculate its fitness function. Thus, the total number of function evaluations in a complete optimization process equals the number of candidates N multiplied by the number of populations that is required to achieve the defined goal.

The basic CE principle is summarized in the flowchart in Fig. 2, and its application to find the optimum of the univariate (1-D) Ackley test function is shown in Fig. 3. The algorithm is initialized by selecting the first population of N candidates from the initial sampling distribution $p(x; v^0)$. In Fig. 3, pdf is shown on the left-hand side. It can be seen that the initial beta distribution with $v^0 = (1, 1)$ is used to characterize the continuous variable x bounded

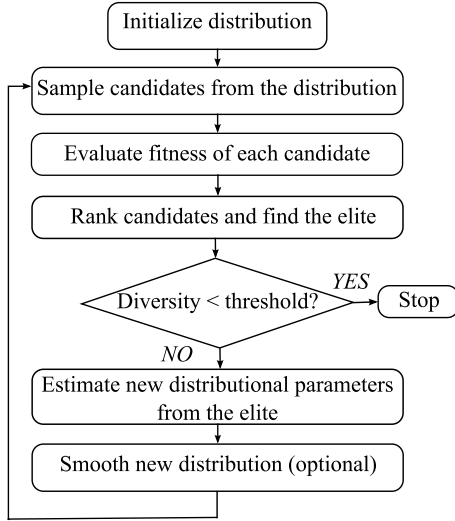


Fig. 2. Flowchart of the CE method algorithm with an optional stopping criterion defined using the population diversity.

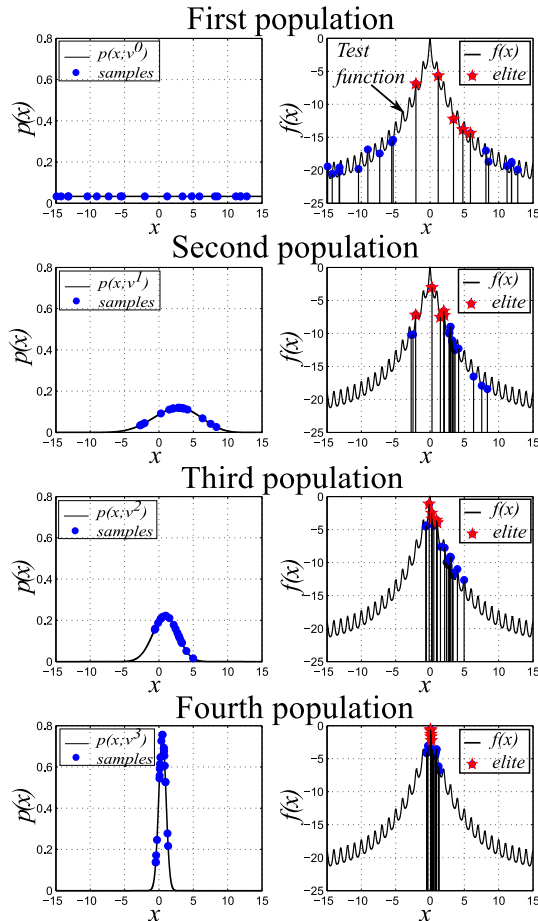


Fig. 3. Illustration of selection and evaluation procedure in the CE method. The figure shows the first four iterations in the optimization of univariate Ackley test function with pdf (left) and fitness (right).

between -15 and 15 . Subsequently, the initial population is created by randomly sampling N candidates from $p(x; v^0)$. The performance of every candidate is evaluated according to the fitness function $f(x)$, and N_{el} best candidates are selected [Fig. 3 (right)] as the elite subpopulation. Depending on the type of optimization, whether it is a maximization or

minimization problem, the fitness values are sorted in descending or ascending order, respectively. The *elite* subpopulation of the N_{el} best performing candidates is used to reconstruct a new distribution $p(x; v^1)$. In particular, the algorithm estimates distributional parameters v^1 that best describe the current *elite* subpopulation according to the Kullback–Leibler distance or the CE minimization [22]. These empirical shape parameters describe the distribution $p(x; v^1)$, which is used to generate the next population. Selection, evaluation, and updating procedures are repeated until the stopping criterion is satisfied. In the illustrated example, the elite subpopulation of the fourth population reached the global solution.

A number of techniques can be used to avoid premature convergence of the CE method. First, the initial sampling distribution can be very wide [as $p(x; (2, 2))$ in Fig. 1] or even uniform [as $p(x; (1, 1))$ in Fig. 3] to cover all plausible designs. Second, the population size N can be defined large enough to maximize the chance that good designs will appear in a random sample. Third, the smoothing parameter α can be applied to distributional parameters in order to slow down the convergence

$$v_s^{t+1} = \alpha * v_{un}^{t+1} + (1 - \alpha)v_s^t \quad (3)$$

where v_{un}^{t+1} is a vector of reconstructed distributional parameters at the current population and v_s^t is a vector of smoothed distributional parameters from a previous population. Finally, if the best fitness value has not been improved for a number of iterations, a mutation operator [24] can be used to add small random perturbations to the shape parameters.

B. Performance Benchmarking

In order to test our implementation of the CE algorithm, we optimized two classical multidimensional test functions, i.e., unimodal sum of squares and multimodal Ackley function using CE [33]. Multimodal functions have multiple local optima, whereas unimodal functions have only one. For comparison with CE, we selected two popular optimization methods in EM: classical PSO with the decreasing inertia weight and the state-of-the-art CMA-ES [34], [35]. We calculated the mean number of function evaluations (MNFE) required for each algorithm to converge to the threshold (δ) of -0.01 over 200 Monte–Carlo simulations. Both the test functions were restricted to the domain $x_i = [-10, 10]$ for all $i = 1, \dots, d$, where d is the problem dimensionality, which was varied to take the values of 5, 10, 15, 20, and 30.

The results are summarized in Table I. For each problem, the internal parameters of CE and PSO were adjusted to ensure that the algorithms perform with 100% success rate. Presented PSO results were acquired with population sizes from 20 (for low-dimensional cases) to 300 (for high-dimensional cases), absorbing the wall boundary condition and the inertia weight [36] with $w_{min} = 0.01$ for low-dimensional cases and $w_{min} = 0.4$ for high-dimensional cases. CE results were achieved with Gaussian distribution and population sizes from 50 to 140, elite sizes from 10 to 15, and a smoothing parameter α within the recommended range of 0.4–0.7. It can be seen that CE and CMA-ES converge much faster than PSO,

TABLE I
COMPARISON OF MNFE FOR CE, PSO, AND CMA-ES
ON TWO TEST FUNCTIONS

Test functions	Dimension, d	CE	PSO	CMA-ES
UNIMODAL:	5	567	573	623
	10	1771	2325	1706
	15	3255	5155	2919
	20	4864	12413	4157
	30	7059	62340	6711
MULTIMODAL:	5	853	1384	994
	10	1908	6636	2281
	15	3090	13708	3531
	20	4628	20310	4704
	30	6512	78380	6810

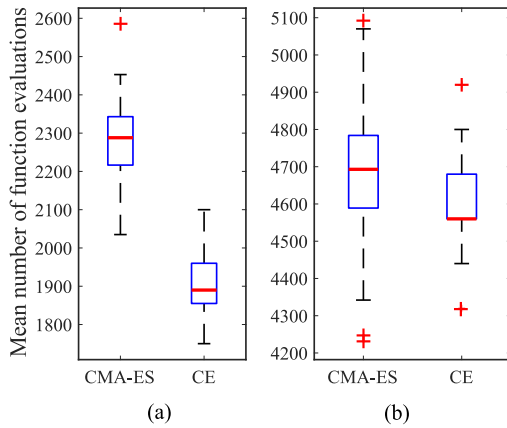


Fig. 4. Box-and-whisker plots of MNFEs for 60 optimization runs by CMA-ES and CE on a d -dimensional Ackley function. (a) $d = 10$. (b) $d = 20$. The boxes are drawn around the 25th and 75th percentiles and divided (by the red thick line) at the 50th percentile. The whiskers extend out to the 10th and 90th percentiles.

especially for high-dimensional and multimodal problems. The convergence rates of CE and CMA-ES are comparable. This can be explained by the similarities in the nature of their updating rules on continuous problems, i.e., both the methods manipulate the parameters of a Gaussian distribution. CMA-ES outperformed CE on the unimodal test problem, and CE outperformed CMA-ES on the multimodal function. This is expected, because CMA-ES is similar to a great extent to a gradient descent local optimizer [37]. The comparison of MNFE for CE and CMA-ES on 10-D and 20-D Ackley functions is shown in Fig. 4 using a box-and-whisker plot. This graph is useful for illustrating variations in large data sets, and we use it to demonstrate that the performance of CE is steadily faster than that of CMA-ES. The results for CMA-ES in Table I agree well with those reported in [16]. Our comparison study is not explicit, but it demonstrates that implemented PSO, CE, and CMA-ES algorithms converge to the global optimum.

In summary, the appealing features of the CE method are: 1) versatility of updating rules and 2) exponential convergence rate. These properties are especially useful when fitness functions are evaluated through computationally expensive full-wave simulations, and thus, NFEs determine the overall

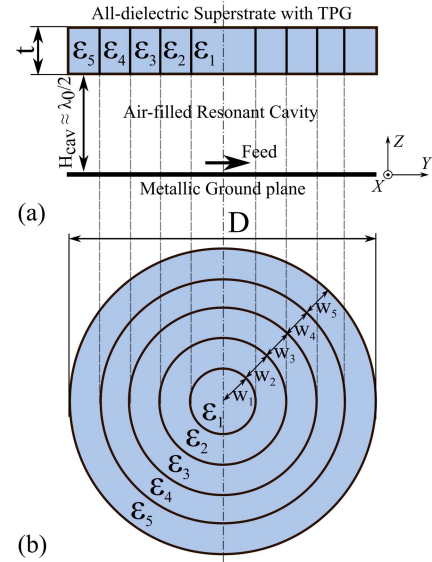


Fig. 5. (a) Side view and (b) top view of an RCA with a TPG superstrate. The superstrate is made out of five dielectric segments.

optimization time. It is noted that CE requires larger samples of candidates to keep the balance between exploration and exploitation. Yet, CE consistently requires less total time compared with CMA-ES and significantly less compared with PSO for multimodal problems. The convergence rate of CE can be controlled by manipulating the size of elite subpopulation and α parameter.

III. DESIGN EXAMPLE

In this section, we describe a practical application of the CE method to optimize the performance of a new class of compact RCAs (see Fig. 5). This example allows us to illustrate the advantages of CE, in particular its ability to simultaneously handle both mixed variables and design constraints without complexity. The constraint in this design example is the total diameter of the antenna. Although the widths w_i of dielectric sections shown in Fig. 5 can take any positive value, their sum ($\sum_{i=1}^N w_i$), i.e., the total radius is limited by specifications. Some design variables, such as w_i and t , are continuous, but the section dielectric constants (ϵ_i) are discrete, i.e., they can take only commercially available values.

In this RCA, a partially reflective superstrate is placed above a metal ground plane to form a resonant cavity in between. The fields in the cavity experience multiple reflections, and when appropriate conditions are satisfied, a directive beam is generated at boresight. In earlier RCAs, resonant cavities with a high- Q factors were used, which resulted in narrow directivity bandwidths in the order of 0.1%–3% [7], [38]. Over the last decade, several techniques have been proposed to overcome this narrowband limitation [39]–[41]. Recently, it was shown that all-dielectric superstructures with TPGs, such as the one shown in Fig. 5, which is made of concentric dielectric rings, provide extremely large directivity bandwidths in the order of 50% [29].

The original RCA [29] of this class was designed using the grid search method to study the effects of varying superstrate

thickness and the number of superstrate segments on antenna performance. Due to the *curse of dimensionality* [42], it is impracticable to search for the optimal design of this RCA manually. Therefore, we aim to explore the potential of this class of antennas by applying CE optimization while making use of its natural ability to deal with constraints and mixed variables simultaneously.

A. Implementation Details

The given RCA with TPG superstrate is a combinatorial and constrained optimization problem. We approach it by combining discrete and continuous variables in a single parameter vector. Discrete variables describe those parameters of the superstrate that can only take the values from a given catalog, for example, readily available permittivity and thickness values of Rogers's materials. Continuous variables describe the width of each section (w_i) in the superstrate under the constraint that they sum to $D/2$, where D is the specified antenna diameter.

In a general case, the permittivities ε_i and section widths w_i of the superstrate form the design vector

$$x = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n, w_1, w_2, \dots, w_n\} \quad (4)$$

where

$$(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) \in S \quad (5)$$

and the section widths w_i are constrained by

$$w_1 + w_2 + \dots + w_n = D/2. \quad (6)$$

S is a catalog of discrete commercially available permittivities and n is the number of sections in the TPG superstrate.

The uniqueness of the proposed CE optimization approach lies in the use of the general discrete distribution family for the discrete variables and the Dirichlet distribution family for the continuous constrained variables. The sampling distribution for ε_i is described by a stochastic vector of length q_p , which is the number of elements in the catalog. These vectors are stored in an $n \times q_p$ matrix. During initialization, this matrix is uniform, meaning that every permittivity has the equal probability to occur in each section. Then, based on fitness function evaluations, those permittivities that lead to better result earn a higher probability of occurrence in the next population. Specifically, each row in the parameter matrix is replaced by the sample proportion vector of that section's permittivities in the elite subpopulation.

The widths w_i cannot be sampled independently, due to the constraint that they sum to $D/2$. For fabrication simplicity and to avoid physically meaningless designs, we assume $w_i > 0.1\lambda_0$ for each section. Thus, the vector $(w_1 - 0.1\lambda_0, \dots, w_n - 0.1\lambda_0)/(D/2 - 0.1n\lambda_0)$ [abbreviated as (x_1, \dots, x_n)] must be elementwise nonnegative and sum to 1, and therefore can be sampled from a multivariate Dirichlet distribution, with the parameter vector $(\alpha_1, \dots, \alpha_n)$. For convenience, we used a simple method-of-moments update of the Dirichlet parameter vector at each iteration [43], rather than minimizing the Kullback–Leibler distance, replacing the

TABLE II
PARAMETERS FOR THE OPTIMIZATION OF RCA

	Case I	Case II
Number of sections, n	3	6
Catalogue of available permittivity values, q_p	5	10
Catalogue of available thickness values, q_t	1	5
Number of candidates in a population, N	45	70
Number of candidates in elite subpopulation, N_{el}	10	10
Smoothing parameter, α	0.5	0.6
Stopping threshold, $[\delta_{perm}, \delta_{width}]$	[0, 0.009 λ_0]	

parameter α_i at each iteration with

$$\alpha_i = \bar{x}_i \left(\frac{1 - \sum_{j=1}^n \bar{x}_j^2}{\sum_{j=1}^n \text{Var}[x_j]} \right) - 1$$

where the sample mean and the variance denoted by the overbar and Var are taken over the elite subpopulation.

The cavity height (H_{cav}) is 13 mm. First, we optimized an RCA with three-sectional superstrate of fixed thickness t . The design variables in this case are $\{\varepsilon_1, \varepsilon_2, \varepsilon_3, w_1, w_2, w_3\}$. This was repeated for three values of t . The catalog of available materials included five laminates. The overall diameter of TPG superstrate and ground plane was constrained to $D = 48$ mm, which is $1.85\lambda_0$, where λ_0 is a free-space wavelength corresponding to the first resonance frequency of the cavity (f_0), which is 11.5 GHz. Then, we optimized an RCA with six-sectional superstrate of variable thickness t with design variables of $\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6, w_1, w_2, w_3, w_4, w_5, w_6, t\}$. Here the catalog of substrates had ten permittivity and five thickness values. Design parameters for both the cases are summarized in Table II.

B. Fitness Function

The aim of optimization is to achieve a high directivity bandwidth product (DBP) and low sidelobe levels (SLLs) in the range from 10 to 20 GHz. Hence, the product of peak directivity Dir_{peak} (in dBi) and 3 dB directivity percentage bandwidth BW (%) given by

$$F_{\text{DBP}} = \text{Dir}_{\text{peak}} * \text{BW} \quad (7)$$

has been considered. As 40% bandwidth is sufficient for most applications, we restricted the largest BW to BW_{lim} as follows:

$$\text{BW} = \begin{cases} \text{BW}, & \text{if } \text{BW} < \text{BW}_{\text{lim}} \\ \text{BW}_{\text{lim}}, & \text{otherwise.} \end{cases} \quad (8)$$

This technique, found successful in previous optimizations of wideband antennas [44], is applied here to obtain a design with a high peak directivity Dir_{peak} , as well as a large percentage BW. In this example, we have limited desirable BW to 40%; however, it can be changed to any other value, which would result in a different superstrate profile and performance.

SLLs in E-plane are much higher those in H-plane; therefore, fitness of SLLs in the E-plane was included in the objective. Sidelobe fitness was defined as the average difference between the desired and current value of SLLs over the

TABLE III
OPTIMIZATION CASE-I RESULTS FOR THREE PREDEFINED
SUPERSTRATE THICKNESS VALUES ($D = 48$ mm)

Superstrate thickness, mm	$t_1=6.35$ ($0.24\lambda_0$)	$t_2=6.99$ ($0.27\lambda_0$)	$t_3=7.62$ ($0.29\lambda_0$)
Superstrate section parameters			
ε_1, w_1	9.8, 7.92	9.8, 10.8	9.8, 8.40
ε_2, w_2	9.2, 8.16	6, 5.28	6, 6.96
ε_3, w_3	4.5, 7.92	4.5, 7.92	4.5, 8.64
Dir_{peak} , dBi	18.51	18.32	19.38
BW, %	45.07	41.46	42.39
F_{SLL} , dB	5.79	4.95	6.16
Fitness function, eq.(10)	624.60	633.80	651.80

frequency range from f_l to f_h

$$F_{SLL} = \text{mean}(\max(0, SLL_{E,(f_l < f < f_h)} - SLL_{E,obj})) \quad (9)$$

where the desired SLL in the E-plane is $SLL_{E,obj} = -20$ dB, and $SLL_{E,(f_l < f < f_h)}$ is the current SLL in the range between f_l and f_h in steps of 0.5 GHz. These multiple objectives were combined into a single objective by a weighted sum

$$f(x) = F_{DBP} - k * F_{SLL} \quad (10)$$

where k was set to 20 in our design to reflect the relative importance. Impedance matching was not included in the objective function as it was beyond the scope of this paper.

Since the superstrate is nonperiodic and has small lateral dimensions (diameter $< 2\lambda_0$ at the lowest frequency), classic design approaches involving unit-cell optimizations [45] or transmission line modeling [46] cannot be used for the analysis of this compact RCA. Instead, full-wave analysis of the complete antenna had to be performed to accurately predict boresight directivity and SLLs. Thus, CE algorithm, implemented in MATLAB, was linked to the time-domain solver in CST Microwave Studio using macroprogramming. To reduce the simulation time, we exploited the bilateral x - and y -symmetries of the RCA superstrate by applying appropriate boundary conditions in two planes.

C. Optimization Case-I: RCA With Three-Sectional TPG

The RCA with a three-sectional superstrate was optimized for three fixed thickness values t_1 , t_2 , and t_3 , and the best result for each thickness is summarized in Table III. The catalog of relative permittivities contained a list of Rogers TMM laminates: [3.27; 4.5; 6; 9.2; 9.8]. Evaluation of one population (45 candidates) took approximately 20 min on an Intel Core i7-4790 CPU@3.6 GHz processor with a 32 GB RAM. The stopping criterion was a measure of elite subpopulation diversity, defined as the maximum variation between the candidates in each elite subpopulation. For stopping, permittivity variation (δ_{perm}) was 0 and width variation (δ_{width}) was $0.009\lambda_0 = 0.2$ mm. Fig. 6(a) shows the convergence plots for Case-I optimization. With parameters $N = 45$, $N_{el} = 10$, and $\alpha = 0.5$, the stopping criterion was reached after 20 iterations for t_1 . Approximately, 14 h were required to complete this optimization. With a less strict stopping criterion, optimization would have been stopped after ten iterations with a solution that is almost as good as this one.

TABLE IV
CATALOG OF RF LAMINATES FOR CASE-II OPTIMIZATION

Dielectric constant	[1.96 2.2 2.94 3.27 3.5 4.5 6 9.2 9.8 10.2]
Thickness, mm	[3.81 5.08 6.35 6.99 7.62]

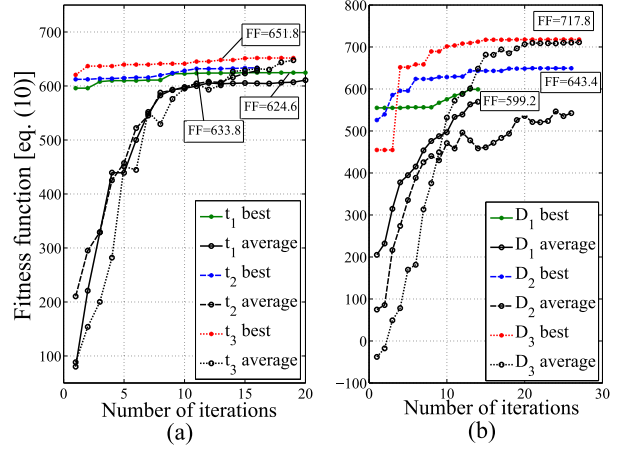


Fig. 6. Convergence results of the CE method for (a) Case-I and (b) Case-II optimization of six compact RCAs with TPG superstrates, where t is the superstrate thickness and D is its diameter. For each optimization run, “best” is the best result observed so far, and “average” is the mean of all results observed at each iteration.

It is interesting to note that without any enforcement, optimal permittivity profiles have a gradual decrease of permittivity from the center toward the edges. This agrees well with theoretical arguments made in [29]. It can also be seen that performance of the antennas obtained by our method is better than that in [29].

D. Optimization Case-II: RCA With Six-Sectional TPG

The RCA with a six-sectional superstrate was optimized for three different total diameter values D_1 , D_2 , and D_3 , with an extended catalog of materials shown in Table IV. Population size N for Case-II optimization was increased to 70 due to the higher dimensionality of the problem, and α was changed to 0.6. The best result for each diameter is summarized in Table V. It can be seen that the highest peak directivity of 19.4 dBi and the lowest SLLs were obtained for the largest diameter (D_3) of the antenna. For this design, the stopping criterion was reached after 27 iterations in approximately 48 h. Convergence plots for Case-II optimization are shown in Fig. 6(b).

To the best of our knowledge, no classical PSO, GA, or CMA-ES algorithm can solve this optimization problem with those constraints and mixed variables. Therefore, performance comparison with other algorithms is not possible.

IV. EXPERIMENTAL RESULTS

The optimized RCA with a three-sectional superstrate of thickness $t_1 = 6.35$ mm was fabricated to validate the results. The prototype, shown in Fig. 7 (inset), was built by subtractive manufacturing using TMM10i, TMM10, and TMM4 Rogers laminates. Two nylon spacers, which supported the superstrate

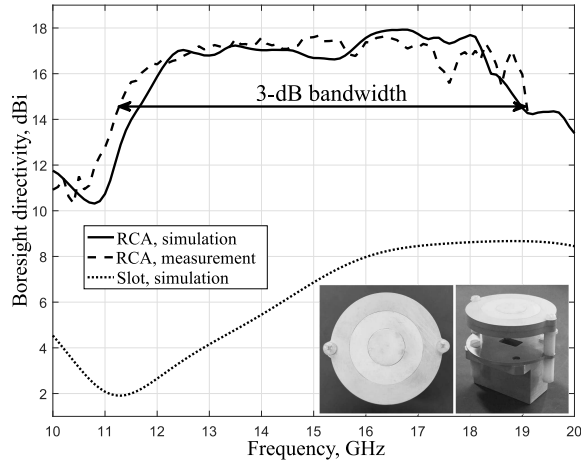


Fig. 7. Measured boresight directivity of the RCA prototype. Measured peak directivity is 17.6 dBi, and measured 3 dB directivity bandwidth (51%) extends from 11.3 to 19.1 GHz. Predicted directivity of the RCA and the slot feed antenna are also shown for comparison. Inset: the RCA prototype.

TABLE V
OPTIMIZATION CASE-II RESULTS FOR THREE PREDEFINED
SUPERSTRATE DIAMETER VALUES ($\lambda_0 = 26$ mm)

Antenna diameter, mm	$D_1=38$ ($1.5\lambda_0$)	$D_2=48$ ($1.85\lambda_0$)	$D_3=58$ ($2.2\lambda_0$)
Superstrate section parameters			
ϵ_1, w_1	9.2, 2.59	9.2, 2.54	10.2, 7.92
ϵ_2, w_2	6, 1.93	9.2, 4.65	9.2, 6.62
ϵ_3, w_3	9.2, 2.57	9.2, 3.37	6, 4.47
ϵ_4, w_4	6, 2.35	6, 4.57	4.5, 2.92
ϵ_5, w_5	4.5, 7.25	4.5, 3.71	3.5, 3.39
ϵ_6, w_6	3.5, 2.31	3.5, 5.15	3.5, 3.68
t	6.99	6.99	6.99
Dir_{peak} , dBi	17.8	18.5	19.4
BW, %	41	43	40
F_{SLL} , dB	5.62	4.9	2.9
Fitness function, eq.(10)	599.2	643.4	717.8

above the aluminum ground plane, introduced 0.5 dB decrease in simulated peak directivity. A rectangular slot feed antenna in a ground plane was fed by a coaxial-to-waveguide adapter WR-75 to excite the resonant cavity. The slot dimensions were 12×7.5 mm² after fine-tuning to provide good impedance matching. The diameter of the ground plane was equal to the diameter of the superstrate (48 mm).

The radiation characteristics of the prototype were measured in a spherical near-field measurement system at the Australian Antenna Measurement Facility. Fig. 7 shows the predicted and measured boresight directivity versus frequency. Overall, good agreement is noted between predicted and measured results. Discrepancy between 17 and 18 GHz is attributed to the use of an approximate simulation model of the WR-75 coaxial-to-waveguide adapter. The measured peak directivity is 17.6 dBi, and 3 dB directivity bandwidth extends from 11.3 to 19.1 GHz, featuring a 51% percentage bandwidth. The comparison of predicted and measured patterns in H- and E-planes at three selected frequencies within the operating band is shown in Fig. 8. The predicted and measured beamwidths agree very well at all frequencies. Overall, measured SLLs are less than -15 dB in H-plane, and less than -10 dB in E-plane, except for 18 GHz, where SLL in

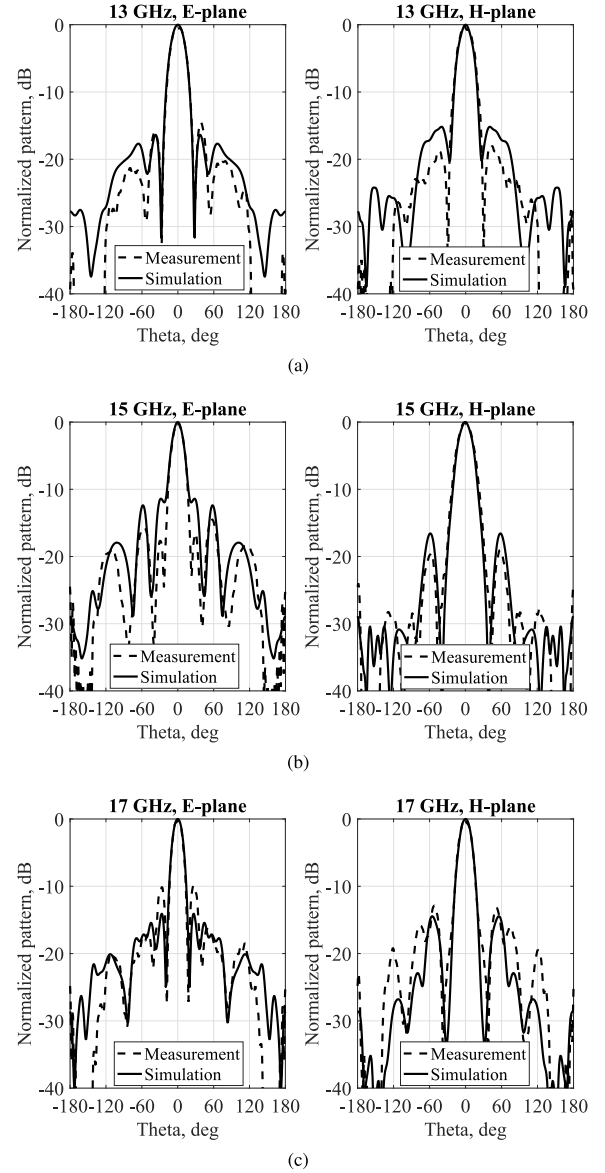


Fig. 8. Predicted and measured radiation patterns of the optimal antenna with superstrate thickness $t_1 = 6.35$ mm and fitness function $f(x) = 624.6$ from Case-I study. (a) 13 GHz. (b) 15 GHz. (c) 17 GHz.

E-plane rises to -7 dB. The predicted and measured front-to-back ratio is greater than 24 dB at all measured frequencies.

V. CONCLUSION

Mixed variables and constraints are commonly found in EM engineering and other disciplines. At present, there is no elegant and simple optimization method to accommodate them when designing antennas and radio-frequency (RF) circuits. This paper shows that the CE method can handle both these challenges simultaneously and efficiently. When applied to two benchmarks, CE requires much fewer function evaluations than PSO and as many as the state-of-the-art CMA-ES. In multi-modal optimization problems, which are common in EM engineering, CE outperforms CMA-ES. By using specific sampling probability distributions in CE updating rules, we designed compact RCAs with improved performance as compared to the initial design. The measurements of an antenna prototype validated the method and predicted results.

As CE is a universal optimization method, the new technique described in this paper is applicable to any EM and RF design problems with both continuous and discrete parameters and/or constraints (for example, the total length of a device has a specified value). In EM and RF engineering, such design requirements are common. Some examples where this technique can be applied include lens antennas, dielectric resonator antennas, microwave filters, EM absorbers, cloaks, and antenna arrays.

REFERENCES

- [1] A. A. Minasian and T. S. Bird, "Particle swarm optimization of microstrip antennas for wireless communication systems," *IEEE Trans. Antennas Propag.*, vol. 61, no. 12, pp. 6214–6217, Dec. 2013.
- [2] A. Mitchell, M. Lech, D. M. Kokotoff, and R. B. Waterhouse, "Search for high-performance probe-fed stacked patches using optimization," *IEEE Trans. Antennas Propag.*, vol. 51, no. 2, pp. 249–255, Feb. 2003.
- [3] D. W. Boeringer and D. H. Werner, "Particle swarm optimization versus genetic algorithms for phased array synthesis," *IEEE Trans. Antennas Propag.*, vol. 52, no. 3, pp. 771–779, Mar. 2004.
- [4] M. G. Bray, D. H. Werner, D. W. Boeringer, and D. W. Machuga, "Optimization of thinned aperiodic linear phased arrays using genetic algorithms to reduce grating lobes during scanning," *IEEE Trans. Antennas Propag.*, vol. 50, no. 12, pp. 1732–1742, Dec. 2002.
- [5] R. L. Haupt, "Optimized element spacing for low sidelobe concentric ring arrays," *IEEE Trans. Antennas Propag.*, vol. 56, no. 1, pp. 266–268, Jan. 2008.
- [6] E. Afacan, "Location optimization for square array antennas using differential evolution algorithm," *Int. J. Antennas Propag.*, vol. 2016, May 2016, Art. no. 1956047. [Online]. Available: <http://dx.doi.org/10.1155/2016/1956047>
- [7] Y. Ge, K. P. Esselle, and Y. Hao, "Design of low-profile high-gain EBG resonator antennas using a genetic algorithm," *IEEE Antennas Wireless Propag. Lett.*, vol. 6, pp. 480–483, 2007.
- [8] Y. Ge and K. P. Esselle, "GA/FDTD technique for the design and optimisation of periodic metamaterials," *IET Microw., Antennas Propag.*, vol. 1, no. 1, pp. 158–164, Feb. 2007.
- [9] F. Yang and Y. Rahmat-Samii, *Electromagnetic Band Gap Structures in Antenna Engineering*. Cambridge, U.K.: Cambridge Univ. Press, 2008.
- [10] R. L. Haupt, "Thinned arrays using genetic algorithms," *IEEE Trans. Antennas Propag.*, vol. 42, no. 7, pp. 993–999, Jul. 1994.
- [11] J. M. Johnson and Y. Rahmat-Samii, "Genetic algorithm optimization and its application to antenna design," in *Antennas Propag. Soc. Int. Symp. Dig.*, 1994, pp. 326–329.
- [12] D. S. Weile and E. Michielssen, "Genetic algorithm optimization applied to electromagnetics: A review," *IEEE Trans. Antennas Propag.*, vol. 45, no. 3, pp. 343–353, Mar. 1997.
- [13] S. Karimkashi and A. A. Kishk, "Invasive weed optimization and its features in electromagnetics," *IEEE Trans. Antennas Propag.*, vol. 58, no. 4, pp. 1269–1278, Apr. 2010.
- [14] A. Foudazi and A. R. Mallahzadeh, "Pattern synthesis for multi-feed reflector antennas using invasive weed optimisation," *IET Microw., Antennas Propag.*, vol. 6, no. 14, pp. 1583–1589, Nov. 2012.
- [15] Z. Bayraktar, M. Komurcu, J. A. Bossard, and D. H. Werner, "The wind driven optimization technique and its application in electromagnetics," *IEEE Trans. Antennas Propag.*, vol. 61, no. 5, pp. 2745–2757, May 2013.
- [16] M. D. Gregory, S. V. Martin, and D. H. Werner, "Improved electromagnetics optimization: The covariance matrix adaptation evolutionary strategy," *IEEE Antennas Propag. Mag.*, vol. 57, no. 3, pp. 48–59, Jun. 2015.
- [17] B. Xing and W.-J. Gao, *Introduction to Computational Intelligence*. Cham, Switzerland: Springer, 2014, pp. 3–17.
- [18] N. Jin and Y. Rahmat-Samii, "Hybrid real-binary particle swarm optimization (HPSO) in engineering electromagnetics," *IEEE Trans. Antennas Propag.*, vol. 58, no. 12, pp. 3786–3794, Dec. 2010.
- [19] R. L. Haupt, "Antenna design with a mixed integer genetic algorithm," *IEEE Trans. Antennas Propag.*, vol. 55, no. 3, pp. 577–582, Mar. 2007.
- [20] E. Boudaher and A. Hoorfar, "Electromagnetic optimization using mixed-parameter and multiobjective covariance matrix adaptation evolution strategy," *IEEE Trans. Antennas Propag.*, vol. 63, no. 4, pp. 1712–1724, Apr. 2015.
- [21] L. Xie and Y.-C. Jiao, "Hybrid real-binary differential evolution algorithm applied to antenna optimization," *Microw. Opt. Technol. Lett.*, vol. 54, no. 6, pp. 1460–1463, 2012.
- [22] R. Y. Rubinstein and D. P. Kroese, *The Cross Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation* (Information Science and Statistics). Secaucus, NJ, USA: Springer-Verlag, 2004.
- [23] J. D. Connor, S. Y. Foo, and M. H. Weatherspoon, "Synthesizing antenna array sidelobe levels and null placements using the cross entropy method," in *Proc. 34th Annu. Conf. IEEE Ind. Electron. (IECON)*, Nov. 2008, pp. 1937–1941.
- [24] S. Ho and S. Yang, "The cross-entropy method and its application to inverse problems," *IEEE Trans. Magn.*, vol. 46, no. 8, pp. 3401–3404, Aug. 2010.
- [25] P. Minvielle, E. Tantar, A.-A. Tantar, and P. Berisset, "Sparse antenna array optimization with the cross-entropy method," *IEEE Trans. Antennas Propag.*, vol. 59, no. 8, pp. 2862–2871, Aug. 2011.
- [26] L. Bian, X. Che, and G. Yang, "Pattern null steering using the cross entropy method by controlling the current phases," *Int. J. Adv. Comput. Technol.*, vol. 3, no. 4, pp. 129–136, 2011.
- [27] L. H. Abderrahmane and B. Boussouar, "New optimisation algorithm for planar antenna array synthesis," *AEUE-Int. J. Electron. Commun.*, vol. 66, no. 9, pp. 752–757, 2012.
- [28] M. H. Weatherspoon, J. D. Connor, and S. Y. Foo, "Shaped beam synthesis of phased arrays using the cross entropy method," *Int. J. Numer. Model., Electron. Netw., Devices Fields*, vol. 26, no. 6, pp. 630–642, 2013.
- [29] R. M. Hashmi and K. P. Esselle, "A class of extremely wideband resonant cavity antennas with large directivity-bandwidth products," *IEEE Trans. Antennas Propag.*, vol. 64, no. 2, pp. 830–835, Feb. 2016.
- [30] P.-T. de Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A tutorial on the cross-entropy method," *Ann. Oper. Res.*, vol. 134, no. 1, pp. 19–67, 2005.
- [31] A. Costa, O. D. Jones, and D. Kroese, "Convergence properties of the cross-entropy method for discrete optimization," *Oper. Res. Lett.*, vol. 35, no. 5, pp. 573–580, 2007.
- [32] Z. Wu and M. Kolonko, "Asymptotic properties of a generalized cross-entropy optimization algorithm," *IEEE Trans. Evol. Comput.*, vol. 18, no. 5, pp. 658–673, Oct. 2014.
- [33] S. Surjanovic and D. Bingham. (2013). *Virtual Library of Simulation Experiments: Test Functions and Datasets*. [Online]. Available: <http://www.sfu.ca/~ssurjano>
- [34] N. Jin and Y. Rahmat-Samii, "Particle swarm optimization for antenna designs in engineering electromagnetics," *J. Artif. Evol. Appl.*, vol. 2008, Jan. 2008, Art. no. 9.
- [35] M. D. Gregory, Z. Bayraktar, and D. H. Werner, "Fast optimization of electromagnetic design problems using the covariance matrix adaptation evolutionary strategy," *IEEE Trans. Antennas Propag.*, vol. 59, no. 4, pp. 1275–1285, Apr. 2011.
- [36] X. Hu, Y. Shi, and R. Eberhart, "Recent advances in particle swarm," in *Proc. Congr. Evol. Comput. (CEC)*, vol. 1, Jun. 2004, pp. 90–97.
- [37] N. Hansen and A. Auger, *Principled Design of Continuous Stochastic Search: From Theory to Practice*. Berlin, Germany: Springer, 2014, pp. 145–180.
- [38] A. R. Weily, K. P. Esselle, B. C. Sanders, and T. S. Bird, "High-gain 1D EBG resonator antenna," *Microw. Opt. Technol. Lett.*, vol. 47, no. 2, pp. 107–114, 2005.
- [39] A. P. Feresidis and J. C. Vardaxoglou, "A broadband high-gain resonant cavity antenna with single feed," in *Proc. 1st Eur. Conf. Antennas Propag. (EuCAP)*, Nov. 2006, pp. 1–5.
- [40] N. Wang, Q. Liu, C. Wu, L. Talbi, Q. Zeng, and J. Xu, "Wideband Fabry-Pérot resonator antenna with two complementary FSS layers," *IEEE Trans. Antennas Propag.*, vol. 62, no. 5, pp. 2463–2471, May 2014.
- [41] R. M. Hashmi, B. A. Zeb, and K. P. Esselle, "Wideband high-gain EBG resonator antennas with small footprints and all-dielectric superstructures," *IEEE Trans. Antennas Propag.*, vol. 62, no. 6, pp. 2970–2977, Jun. 2014.
- [42] R. B. Marimont and M. B. Shapiro, "Nearest neighbour searches and the curse of dimensionality," *IMA J. Appl. Math.*, vol. 24, no. 1, pp. 59–70, 1979.
- [43] T. P. Minka, "Estimating a Dirichlet distribution," Tech. Rep., 2000. [Online]. Available: <https://tminka.github.io/papers/>
- [44] J. Jayasinghe, J. Anguera, and D. Uduwawala, "On the behavior of several fitness functions for genetically optimized microstrip antennas," *Int. J. Sci. World*, vol. 3, no. 1, pp. 53–58, 2015.

- [45] B. A. Zeb, R. M. Hashmi, K. P. Esselle, and Y. Ge, "The use of reflection and transmission models to design wideband and dual-band Fabry-Pérot cavity antennas (invited paper)," in *Proc. URSI Int. Symp. Electromagn. Theory (EMTS)*, May 2013, pp. 1084–1087.
- [46] A. Hosseini, F. Capolino, F. D. Flaviis, P. Burghignoli, G. Lovat, and D. R. Jackson, "Improved bandwidth formulas for Fabry-Pérot cavity antennas formed by using a thin partially-reflective surface," *IEEE Trans. Antennas Propag.*, vol. 62, no. 5, pp. 2361–2367, May 2014.



Maria Kovaleva (S'14) received the B.S. degree (Hons.) in electrical engineering from the Moscow Technical University of Communications and Informatics, Moscow, Russia, in 2011. She is currently pursuing the Ph.D. degree with the Centre for Collaboration in electromagnetic and Antenna Engineering (C4CELANE), Macquarie University, Sydney, NSW, Australia.

She was an Antenna Design Engineer with the JSC NIKP (Russian Space Systems), Moscow, Russia, for three years. Her research interests include microwave and millimeter-wave antennas, resonant cavity antennas, evolutionary optimization methods in electromagnetics and surface electromagnetics.



David Bulger received the B.Sc. degree (Hons.) in mathematics from the University of Canterbury, Christchurch, New Zealand, in 1992, the M.Sc. degree (Hons.) in mathematics from Massey University, Palmerston North, New Zealand, in 1993, and the Ph.D. degree in mathematics from Central Queensland University, North Rockhampton, QLD, Australia, in 1997.

From 1997 to 1999, he was a Computer Programmer with the private sector, first with Kiwiplan NZ, Manukau, New Zealand, and then with Computer Systems Implementation Ltd., Christchurch, New Zealand. He returned to Academia with a Post-Doctoral Fellowship from 2000 to 2003 with Massey University, continuing the study of optimization algorithms he began in his Ph.D. Since 2004, he has been with the Statistics Department, Macquarie University, Sydney, NSW, Australia. He was a Principal Supervisor of three completed Ph.D. students and a completed M.Res. student, and an Associate Supervisor for two more completed Ph.D. students. His research interests include statistical applications in public health management, business planning and animal behavior, optimization theory and practice, quantum algorithms, and mathematical and statistical musicology.

Dr. Bulger received the New Zealand Mathematical Society's Annual National Predoctoral Thesis Competition in 1993.



Basit Ali Zeb (S'11–M'14) received the B.S. degree in electrical engineering degree (Hons.) from the University of Engineering and Technology, Lahore, Pakistan, in 2001, the M.S. degree in telecommunications from the Technical University of Denmark, Lyngby, Denmark, in 2005, and the Ph.D. degree in electronic engineering from Macquarie University, Sydney, NSW, Australia, in 2014.

He is currently a Research Associate with the Centre for Collaboration in Electromagnetic and Antenna Engineering (C4CELANE), Macquarie University. He has authored or co-authored over 35 journal and conference papers. His research interests include high-performance resonant cavity antennas, reconfigurable antennas, RF circuits and frequency selective surfaces for microwave, and millimetrewave applications.

Dr. Zeb received the Best Student Paper Award at the 2011 Twelfth Australian Symposium on Antennas, the Imperial College Press Prize, the Metamaterials 2011 Congress, and the Honorable Mention, Student Paper Competition in the 2012 IEEE APS/URSI International Symposium. He is a Reviewer of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, the IEEE ANTENNAS AND WIRELESS PROPAGATION LETTERS, the *IET Electronic Letters*, the *IET Microwaves, Antennas and Propagation*, and the European Conference on Antennas and Propagation.



Karu P. Esselle (M'92–SM'96–F'16) received the B.Sc. degree (Hons.) in electronic and telecommunication engineering from the University of Moratuwa, Moratuwa, Sri Lanka, and the M.A.Sc. and Ph.D. degrees in electrical engineering from the University of Ottawa, Ottawa, ON, Canada.

He has provided expert assistance to more than a dozen companies, including Intel, USA, Hewlett Packard Laboratory, Palo Alto, CA, USA; Cisco Systems, Richardson, TX, USA; Cochlear, Australia; Optus, Australia; ResMed, Australia; and Katherine-

Werke, Germany. He has served as the Division Executive and the Head of the Department several times. He is currently a Professor of electronic engineering with Macquarie University, Sydney, NSW, Australia, the Director of the WiMed Research Centre (one of the two), and the past Associate Dean Higher Degree Research of the Division of Information and Communication Sciences. He is also the Director of the Centre for Collaboration in electromagnetic and Antenna Engineering. He has authored over 500 research publications.

Dr. Esselle has served as a member of the Deans Advisory Council. He was elevated to prestigious IEEE Fellow grade for his contributions to resonance-based antennas. He is also a fellow of Engineers Australia. After two stages in the selection process, he has been selected by the IEEE Antennas and Propagation (AP) Society as one of two candidates in the ballot for 2019 President of the Society, which has over 8000 members worldwide. Only three people from Asia or Pacific have received this honor (one from Australia and two from Japan) in the 68-year history of this Society the premier global organization dedicated for antennas and propagation. He has also been selected as one of the three new distinguished lecturers of the IEEE AP Society from 2017 to 2019. He is the only Australian AP Distinguished Lecturer (DL) in almost two decades, and the second Australian AP DL ever. When he was elected to the IEEE AP Society Administrative Committee for a three-year term in 2014, he became the only person residing in the Asia-Pacific Region (IEEE Region 10) to be elected to this highly competitive position over a period of at least six years from 2010 to 2015. His other awards include the 2017 Engineering Excellence Award for Best Innovation and the 2016 and 2012 Engineering Excellence Awards for the Best Published Paper from the IESL NSW Chapter, the 2011 Outstanding Branch Counselor Award from the IEEE headquarters, USA, the 2009 Vice Chancellors Award for Excellence in Higher Degree Research Supervision, and the 2004 Innovation Award for best invention disclosure. His mentees have been received many fellowships, awards, and prizes for their research achievements. Thirty-six international experts who examined the theses of his recent Ph.D. graduates ranked them in the top 5% or 10%. He is the Chair of the Board of Management of Australian Antenna Measurement Facility and the Elected Chair of both the IEEE New South Wales (NSW) Section and the IEEE NSW AP/MTT Chapter, in 2016 and 2017, respectively. He is the General Co-Chair of TENSYP 2018, the Technical Program Committee Co-Chair of ISAP 2015, APMC 2011, and TENCON 2013, and the Publicity Chair of ICEAA 2016, IWAT 2014, and APMC 2000. He is the Foundation Counselor of the IEEE Student Branch with Macquarie University, and the Foundation Advisor of the IEEE MTT Chapter in Macquarie University. His papers have been cited over 4,800 times. He is the first Australian Antenna Researcher ever to reach Google Scholar h-index of 30 and his current h-index is the highest among Australian antenna researchers. Since 2002, his research team has been involved with research grants, contracts, and Ph.D. scholarships worth over 16 million dollars. His research has been funded by many national and international organizations, including Australian Research Council, Australia, Intel, U.S. Air Force, USA, Cisco Systems, Hewlett-Packard, and Australian and Indian Governments. He has been invited to serve as an International Expert/Research Grant Assessor by several nationwide research funding bodies overseas, including The Netherlands, Canada, Finland, Hong Kong, Georgia, South Africa, and Chile. He has been invited by Vice-Chancellors of Australian and overseas universities to assess applications for promotion to professorial levels. He has also been invited to assess grant applications submitted to Australia's most prestigious schemes, such as Australian Federation Fellowships and Australian Laureate Fellowships. He leads the Implantable Wireless Program of the WiMed Research Centre. In addition to the large number of invited conference speeches he has given, he has been an invited keynote speaker of the IEEE workshops and conferences. He is an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION and the IEEE ACCESS.