ADJOINT VARIABLE METHODS FOR DESIGN SENSITIVITY ANALYSIS WITH THE METHOD OF MOMENTS

N.K. Georgieva, S. Glavic, M.H. Bakr and J.W. Bandler

McMaster University, CRL223, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada

e-mail: talia@mcmaster.ca

tel: (905) 525 9140 ext. 27141

fax: (905) 523 4407

Outline

Introduction and Objectives

- design sensitivity analysis

The Adjoint Variable Method

- the direct differentiation method
- the adjoint variable method
- computational efficiency, feasibility, accuracy

Applications with Frequency-Domain Solvers

Conclusions

Introduction and Objectives

Design sensitivity analysis

- sensitivity of the state variables
- sensitivity of the response (or objective) function

Objectives

- obtain the response and its gradient in the design variable space through a single full-wave analysis
- applications with frequency-domain solvers
- feasibility of the approach

The Adjoint Variable Method

the linear EM problem

$$\boldsymbol{Z}(\boldsymbol{x})\boldsymbol{I} = \boldsymbol{V}$$

$\boldsymbol{x} = [x_1 \cdots x_n]^T$ - design parameters

$\boldsymbol{I} = [I_1 \cdots I_m]^T$ - state variables

define a scalar function (response function, objective function)

$$f\left(\boldsymbol{x}, \bar{\boldsymbol{I}}(\boldsymbol{x})\right)$$

objective

$$\nabla_{\boldsymbol{x}} f \ \text{ subject to } \ \boldsymbol{Z}(\boldsymbol{x})\boldsymbol{I} = \boldsymbol{V}, \qquad \nabla_{\boldsymbol{x}} f = \left[\frac{\partial f}{\partial x_1} \ \frac{\partial f}{\partial x_2} \ \cdots \ \frac{\partial f}{\partial x_n} \right]$$

The Adjoint Variable Method

state variable sensitivity

(direct differentiation method, DDM)

E.J. Haug, K.K. Choi and V. Komkov, *Design Sensitivity Analysis of Structural Systems*, 1986

J.W. Bandler, *Optimization, vol. 1, Lecture Notes*, 1994

$$\nabla_x \boldsymbol{I} = \boldsymbol{Z}^{-1}\left[\nabla_x \boldsymbol{V} - \nabla_x(\boldsymbol{Z}\bar{\boldsymbol{I}})\right]$$

$$\frac{\partial \boldsymbol{I}}{\partial x_i} = \boldsymbol{Z}^{-1}\left[\frac{\partial \boldsymbol{V}}{\partial x_i} - \frac{\partial \boldsymbol{Z}}{\partial x_i}\bar{\boldsymbol{I}}\right], \quad i = 1, \ldots, n$$

$$\nabla_x f = \nabla_x^e f + \nabla_{\boldsymbol{I}} f \cdot \nabla_x \boldsymbol{I}$$

$$\nabla_{\boldsymbol{I}} f = \begin{bmatrix} \frac{\partial f}{\partial I_1} & \cdots & \frac{\partial f}{\partial I_m} \end{bmatrix}; \quad \nabla_x \boldsymbol{I} = \begin{bmatrix} \frac{\partial I_1}{\partial x_1} & \cdots & \frac{\partial I_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial I_m}{\partial x_1} & \cdots & \frac{\partial I_m}{\partial x_n} \end{bmatrix}$$

The Adjoint Variable Method

response function sensitivity
(adjoint variable method, AVM)

$$\nabla_{\boldsymbol{x}} f = \nabla^{e}_{\boldsymbol{x}} f + \nabla_{\boldsymbol{I}} f \cdot \boldsymbol{Z}^{-1} \left[\nabla_{\boldsymbol{x}} \boldsymbol{V} - \nabla_{\boldsymbol{x}} (\boldsymbol{Z} \bar{\boldsymbol{I}}) \right]$$

$$\hat{\boldsymbol{I}} = \left[\nabla_{\boldsymbol{I}} f \cdot \boldsymbol{Z}^{-1} \right]^{T} = \left[\boldsymbol{Z}^{T} \right]^{-1} \left[\nabla_{\boldsymbol{I}} f \right]^{T}$$

$$\boldsymbol{Z}^{T} \hat{\boldsymbol{I}} = \left[\nabla_{\boldsymbol{I}} f \right]^{T}$$

$$\nabla_{\boldsymbol{x}} f = \nabla^{e}_{\boldsymbol{x}} f + \hat{\boldsymbol{I}}^{T} \left[\nabla_{\boldsymbol{x}} \boldsymbol{V} - \nabla_{\boldsymbol{x}} (\boldsymbol{Z} \bar{\boldsymbol{I}}) \right]$$

$$\frac{\partial f}{\partial x_i} = \frac{\partial_e f}{\partial x_i} + \hat{\boldsymbol{I}}^{T} \left[\frac{\partial \boldsymbol{V}}{\partial x_i} - \frac{\partial \boldsymbol{Z}}{\partial x_i} \bar{\boldsymbol{I}} \right], \quad i = 1, 2, \ldots, n$$

The Adjoint Variable Method

computational efficiency (single excitation mode)

	LU-decompositions	Back-substitutions
FDA	$n+1$	$n+1$
DDM	1	n
AVM	1	1

The Adjoint Variable Method

feasibility and accuracy of the AVM

finite-difference approximations within the AVM

the matrix sensitivity

$$\frac{\partial f}{\partial x_i} = \frac{\partial_e f}{\partial x_i} + \hat{\boldsymbol{I}}^T \left[\frac{\partial \boldsymbol{V}}{\partial x_i} - \frac{\partial \boldsymbol{Z}}{\partial x_i} \bar{\boldsymbol{I}} \right], \quad i = 1, \ldots, n$$

$$\frac{\partial \boldsymbol{Z}}{\partial x_i} \simeq \frac{\Delta \boldsymbol{Z}}{\Delta x_i}, \quad i = 1, \ldots, n$$

the adjoint excitation

$$\boldsymbol{Z}^T \hat{\boldsymbol{I}} = [\nabla_{\boldsymbol{I}} f]^T \qquad \hat{\boldsymbol{V}} = [\nabla_{\boldsymbol{I}} f]^T$$

$$\hat{\boldsymbol{V}} \simeq \left[\frac{\Delta f}{\Delta I_1}, \ldots, \frac{\Delta f}{\Delta I_m} \right]$$

Applications

1. Input impedance of a dipole (Pocklington's eqn., complex code)

sensitivity with respect to the normalized length

$L_n = L/\lambda$

$$\partial R_{in} / \partial L_n \quad \partial X_{in} / \partial L_n \quad \text{subject to } \boldsymbol{ZI} = \boldsymbol{V}$$

(1) finite-difference approach (FDA):

$$\frac{\partial Z_{in}(L_n^{(k)})}{\partial L_n} \simeq \frac{Z_{in}(L_n^{(k)} + \Delta L_n^{(k)}) - Z_{in}(L_n^{(k)})}{\Delta L_n^{(k)}} \qquad \Delta L_n^{(k)} = 0.01 L_n^{(k)}$$

Applications

adjoint variable method (AVM):

the matrix sensitivity

$$\frac{\Delta Z_{ij}}{\Delta L_n^{(k)}} \simeq \frac{Z_{ij}(L_n^{(k)} + \Delta L_n^{(k)}) - Z_{ij}(L_n^{(k)})}{\Delta L_n^{(k)}} \qquad \Delta L_n^{(k)} = 0.01 L_n^{(k)}$$

(2) complete re-meshing: full ΔZ matrix

(3,4) boundary layer: sparse ΔZ matrix

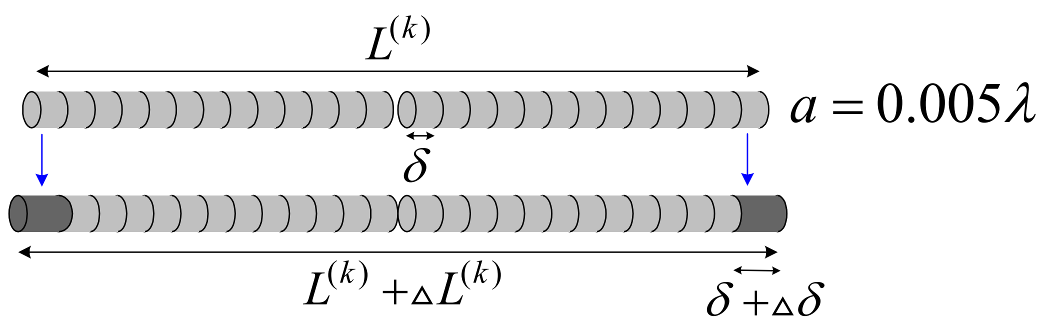


Fig. 1. The dipole and the boundary layer concept (*S. Amari*, 2001).

Applications

the adjoint excitation (analytical)

$$\hat{V}_b = \frac{\partial Z_{in}}{\partial I_b} = \frac{\partial (1/I_b)}{\partial I_b} = -\frac{1}{I_b^2},\ \hat{V}_j = 0 \text{ for } j \neq b$$

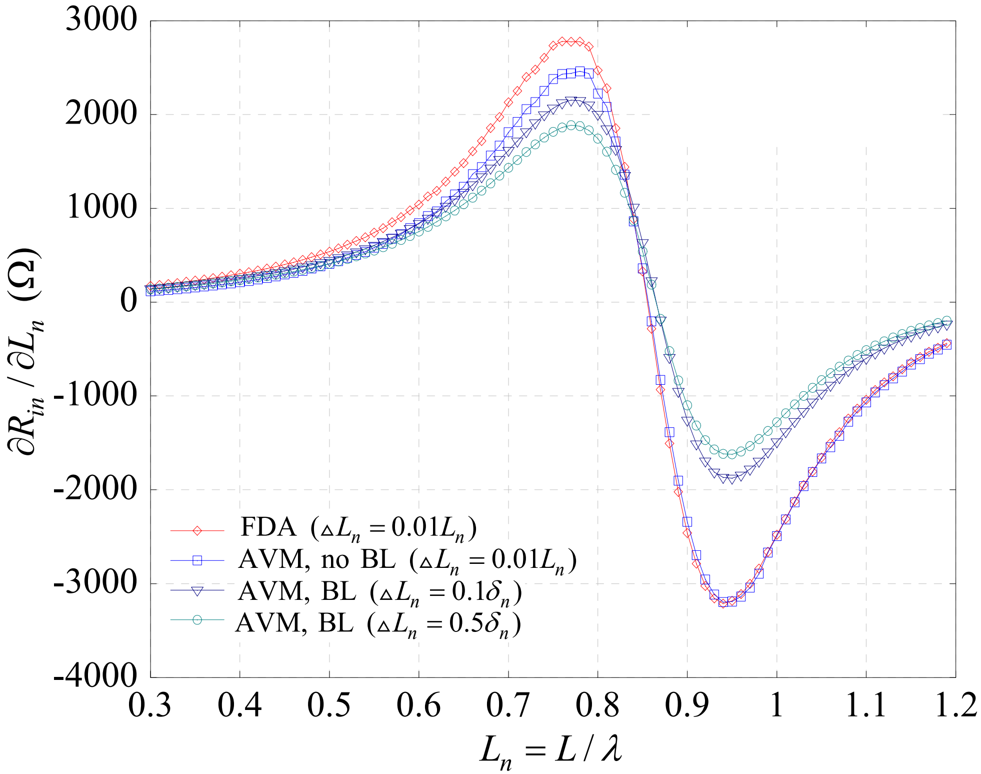


Fig. 2. Derivative of the input resistance of the dipole with respect to L_n.

Applications

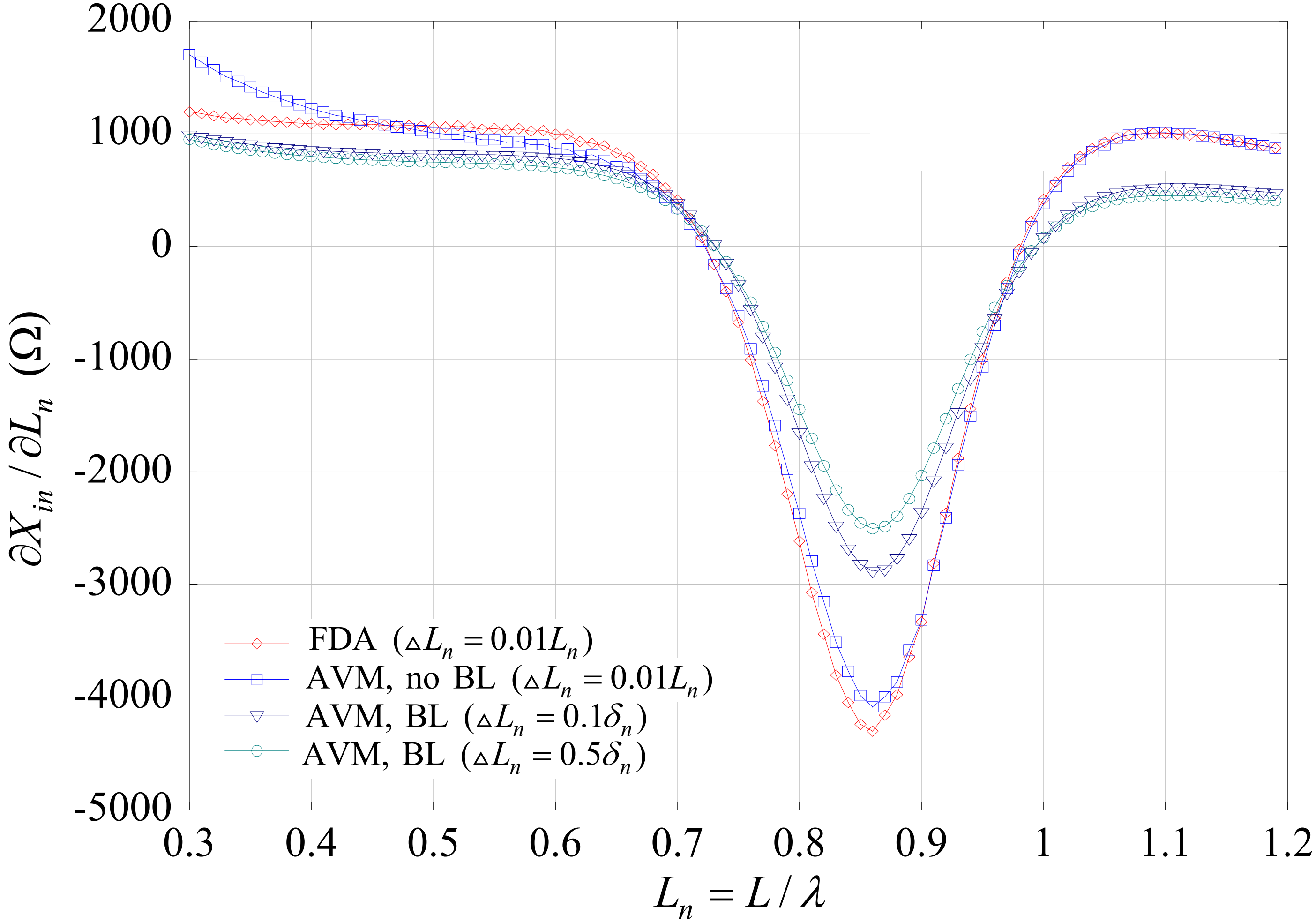
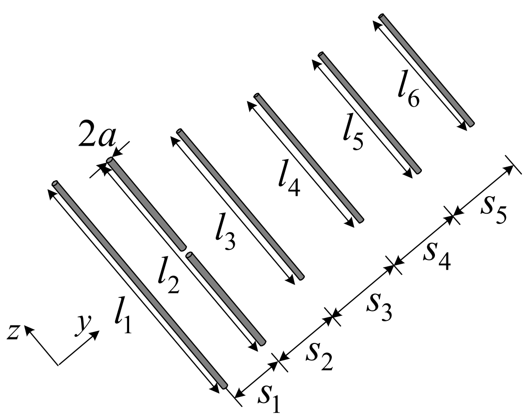


Fig. 3. Derivative of the input reactance of the dipole with respect to L_n.

2. Input impedance of a Yagi-Uda array

sensitivity with respect to the normalized separation distance driver-reflector and the normalized reflector length

$$\boldsymbol{x} = [l_{1n} \quad s_{1n}]^T$$



l_1/λ	l_2/λ	l_d/λ	s_1/λ	s_d/λ	a/λ
0.5243	0.45	0.406	0.2607	0.34	0.003
$l_3 = l_4 = l_5 = l_6 = l_d$; $s_2 = s_3 = s_4 = s_5 = s_d$					

Fig. 4. The geometry of the Yagi-Uda array (initial design).

Applications

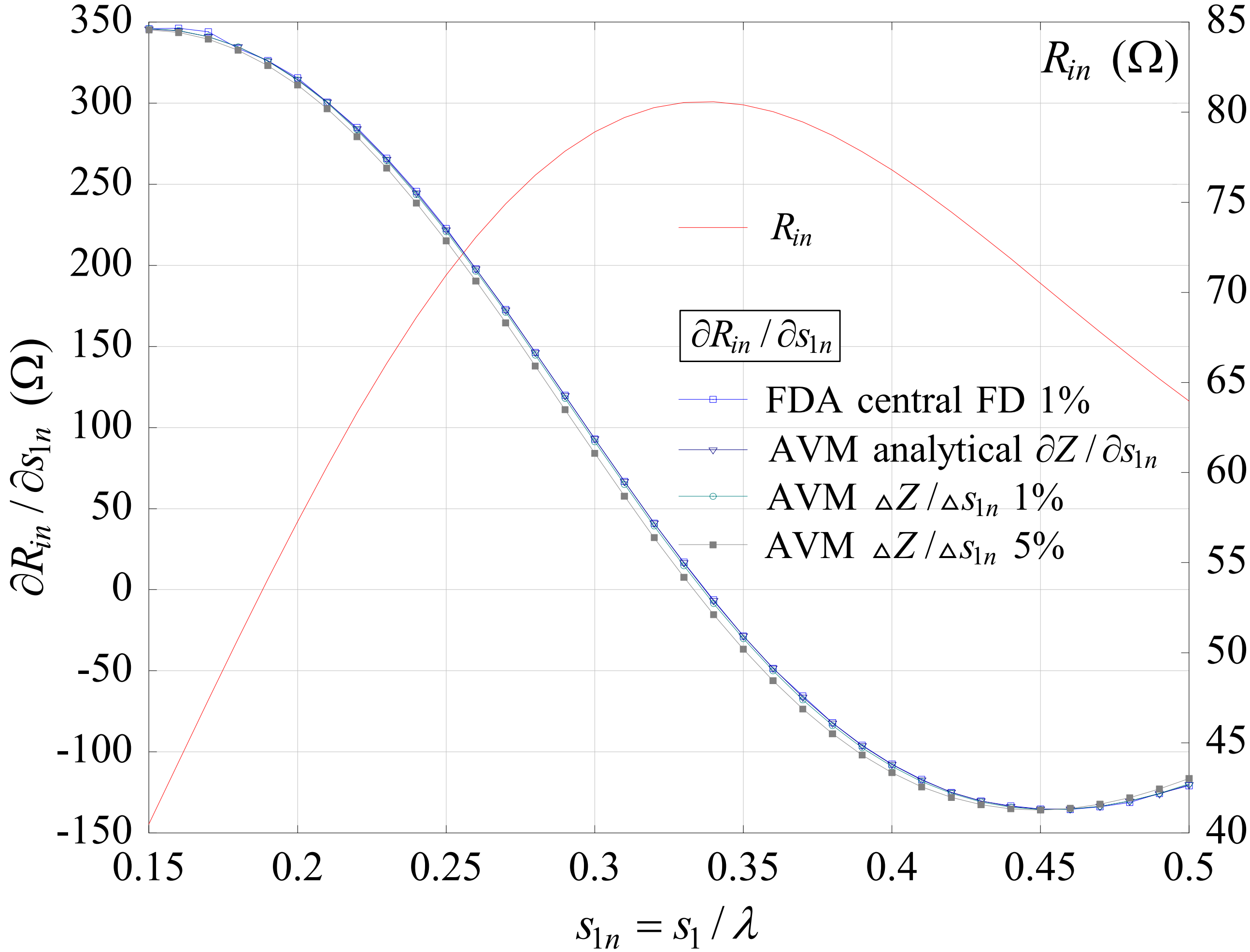


Fig. 5. Input resistance sensitivity with respect to s_{1n}.

Applications

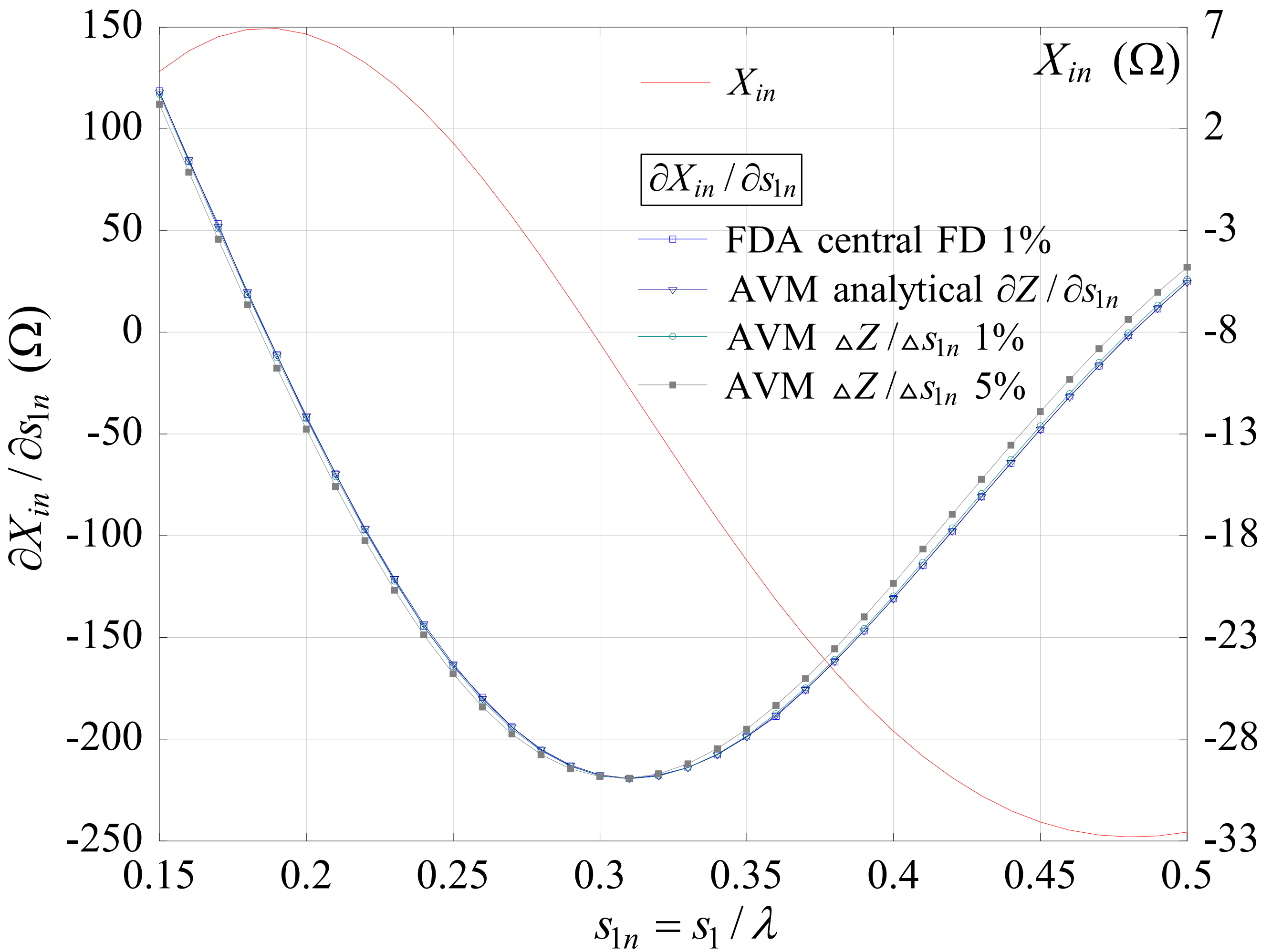


Fig. 6. Input reactance sensitivity with respect to s_{1n}.

Applications

3. Gain of a Yagi-Uda array (Pocklington's eqn., real code)

sensitivity with respect to the normalized separation distances

$s_k = s/\lambda,\ \ k = 1,\ldots,5$

$$\partial G / \partial s_k \quad \text{subject to} \quad \hat{\boldsymbol{Z}}\hat{\boldsymbol{I}} = \hat{\boldsymbol{V}}, \quad \hat{\boldsymbol{I}} = \begin{bmatrix} \mathrm{Re}\{\boldsymbol{I}\} \\ \mathrm{Im}\{\boldsymbol{I}\} \end{bmatrix}$$

the gain sensitivity depends on s_{n_i} explicitly

$$\frac{\partial G}{\partial s_{n_i}} = \frac{\partial_e G}{\partial s_{n_i}} - \hat{\boldsymbol{I}}^T \left(\frac{\partial \boldsymbol{Z}}{\partial s_{n_i}} \bar{\boldsymbol{I}} \right), \ \ i = 1,\ldots,5$$

Applications

the adjoint excitation $\hat{V}$ is a full vector

analytical

$$\hat{V}_k = \frac{\partial G}{\partial \operatorname{Re}(I_k)}$$

$$\hat{V}_{k+m} = \frac{\partial G}{\partial \operatorname{Im}(I_k)}$$

$$k = 1, \ldots, m$$

finite differences

$$\hat{V}_k \simeq \frac{\vartriangle G}{\vartriangle \operatorname{Re}(I_k)}$$

$$\hat{V}_{k+m} \simeq \frac{\vartriangle G}{\vartriangle \operatorname{Im}(I_k)}$$

Applications

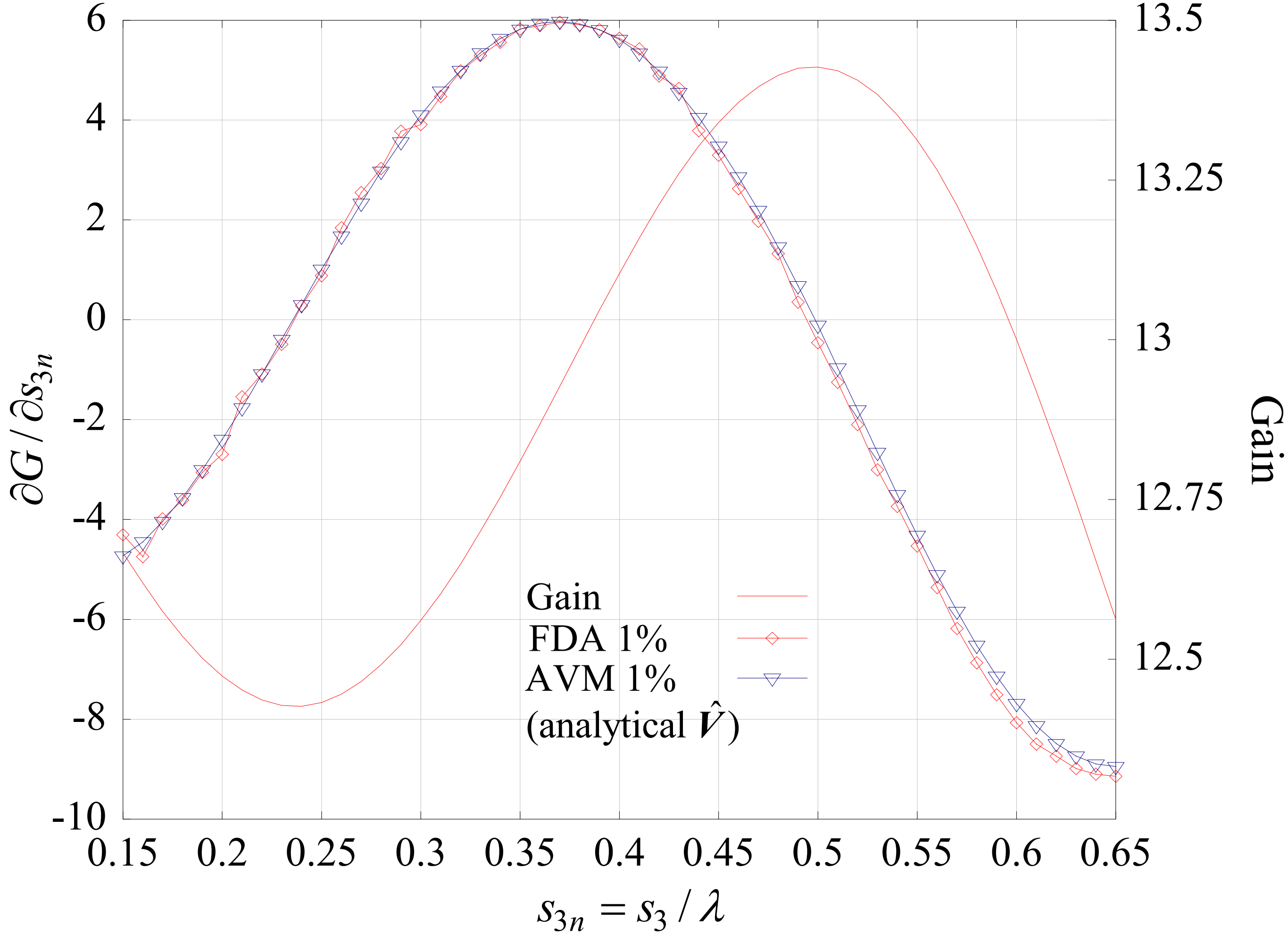


Fig. 7. Gain and gain sensitivity of the Yagi-Uda array with respect to s_{4n}.

Applications

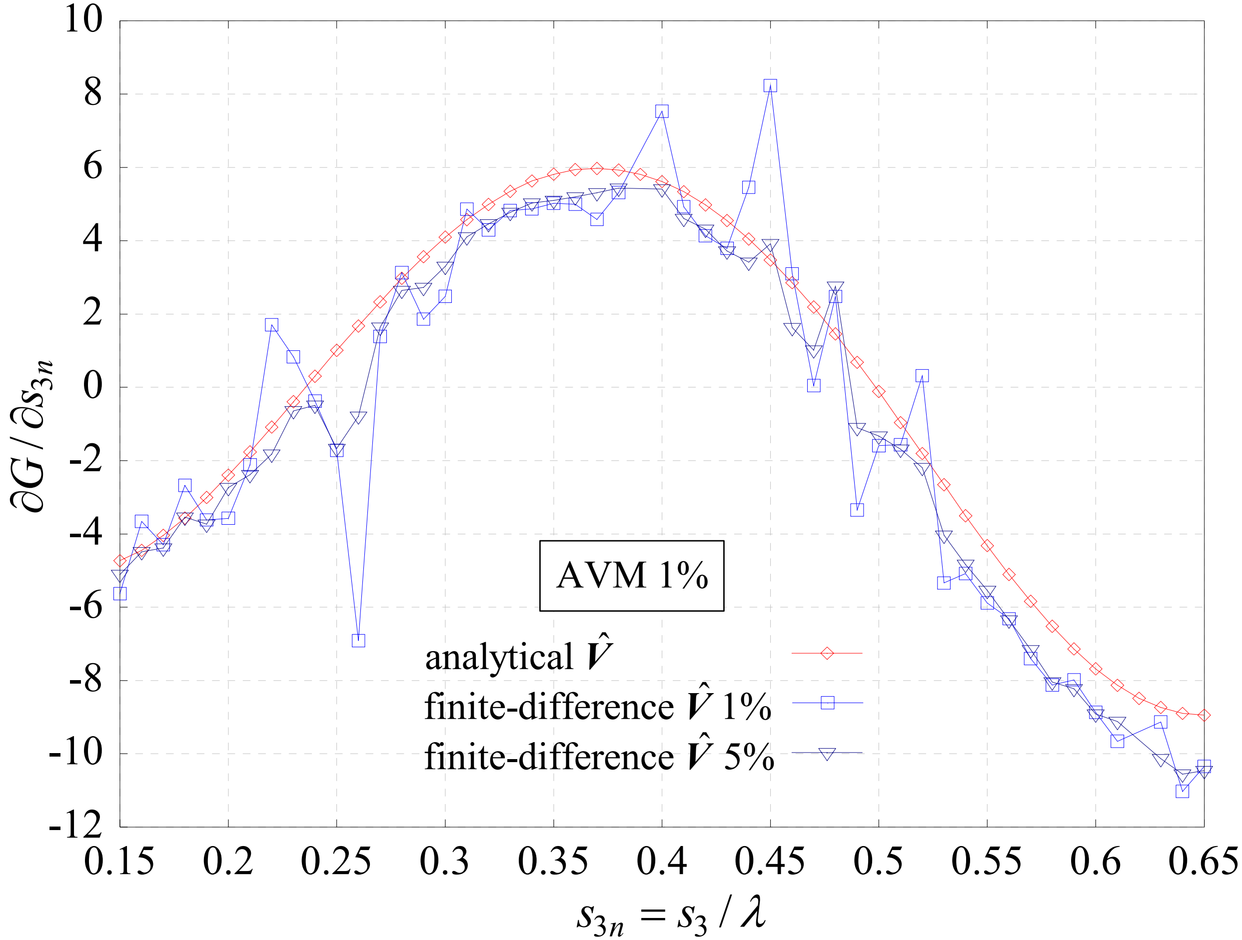


Fig. 8. Gain sensitivity of the Yagi-Uda array with respect to s_{4n}; finite-difference approximation of $\hat{V}$.

Applications

4. Optimization of the Yagi-Uda array for maximum gain and an input impedance of 73 Ω

design parameters

$$\boldsymbol{x} = [s_{1_n} \;\; s_{2_n} \;\; s_{3_n} \;\; s_{4_n} \;\; s_{5_n}]^T$$

objective function

$$f(\boldsymbol{x}) = 0.5\left[(\mathrm{Re}\{Z_{in}\} - 73)^2 + (\mathrm{Im}\{Z_{in}\})^2\right] - 0.5G^2$$

we start from a design already optimized for gain only

Applications

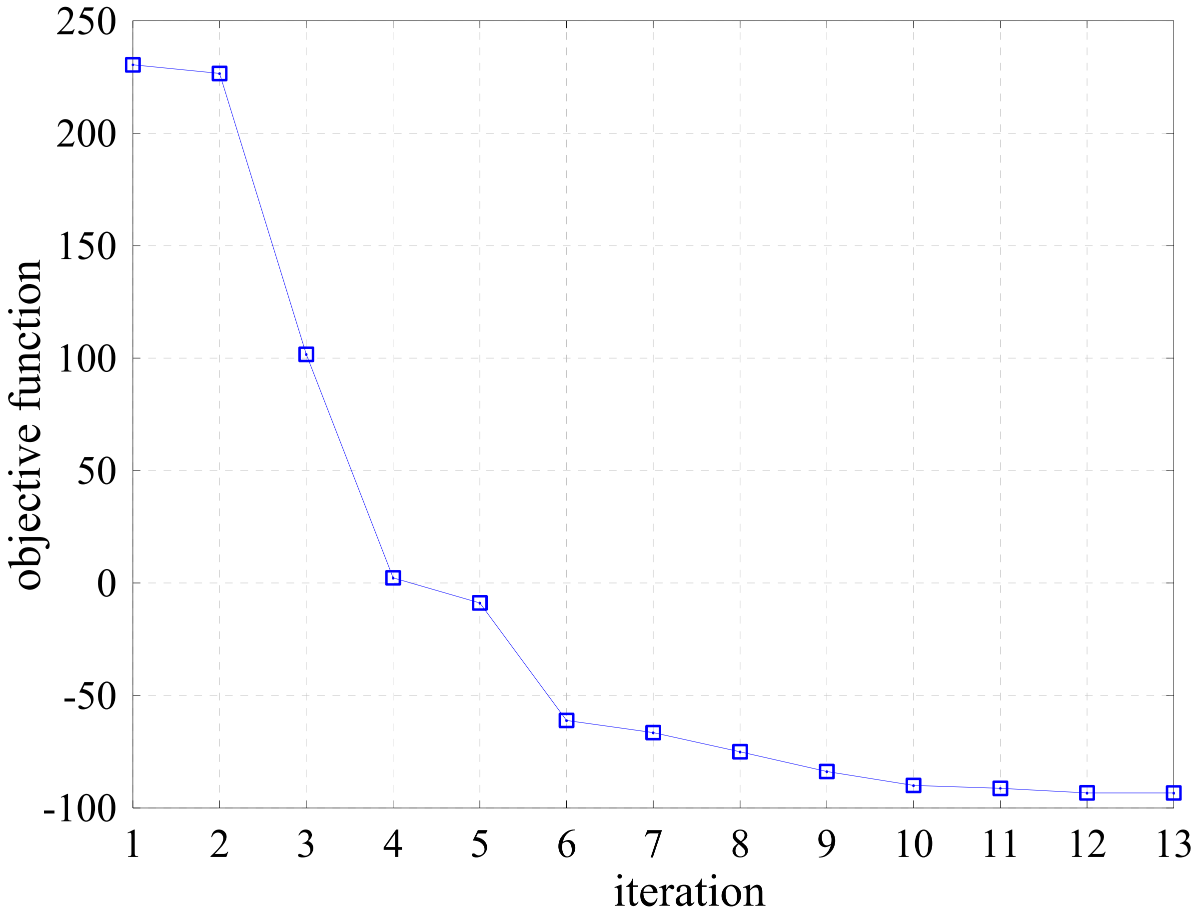


Fig. 9. The progress of the objective function during the optimization of the input impedance and the gain of the Yagi-Uda array.

Applications

TABLE I

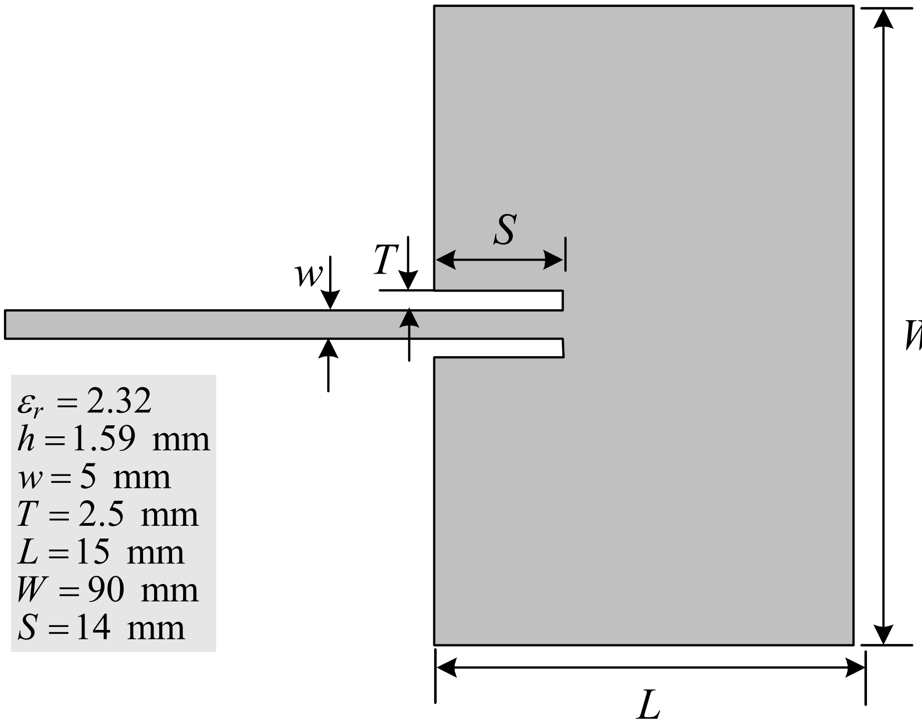
DESIGN PARAMETERS, INPUT IMPEDANCE AND GAIN OF THE YAGI-UDA ARRAY DESIGN

	s_{1n}	s_{2n}	s_{3n}	s_{4n}	s_{5n}	R_{in}	X_{in}	G
1	**0.2607**	**0.3400**	**0.3735**	**0.4471**	**0.4353**	**47.10**	**-4.15**	**15.08**
2	0.3455	0.4050	0.3301	0.3853	0.3765	77.77	-23.52	11.08
3	0.3544	0.4294	0.3639	0.4122	0.3544	81.02	-16.26	11.19
4	0.3158	0.3720	0.4229	0.4591	0.4158	73.33	13.25	13.08
5	0.3086	0.3613	0.4232	0.4519	0.4023	65.85	11.18	13.92
6	0.3450	0.3744	0.3953	0.4204	0.3909	70.23	-5.99	12.87
7	0.3214	0.3986	0.3535	0.4653	0.3432	72.36	-5.77	12.91
8	0.3062	0.3923	0.3844	0.4822	0.3362	72.85	5.46	13.41
9	0.2531	0.4357	0.3794	0.3607	0.3645	75.63	0.97	13.25
10	0.2999	0.4061	0.3777	0.4205	0.3627	72.99	-1.00	13.45
11	0.2874	0.4193	0.3685	0.4057	0.3825	72.26	-1.27	13.59
12	0.2884	0.4175	0.3749	0.4064	0.3937	71.51	0.92	13.77
13	**0.2906**	**0.4168**	**0.3771**	**0.4046**	**0.3966**	**71.80**	**0.38**	**13.75**

5. Optimization of a patch antenna for an input impedance of 50 Ω

design parameters

$$\boldsymbol{x} = [L\ W\ S]^T$$



$\varepsilon_r = 2.32$
$h = 1.59$ mm
$w = 5$ mm
$T = 2.5$ mm
$L = 15$ mm
$W = 90$ mm
$S = 14$ mm

objective function

$$f(\boldsymbol{x}) = \left(\mathrm{Re}\{Z_{in}\} - 50\right)^2 + \left(\mathrm{Im}\{Z_{in}\}\right)^2$$

Applications

$\boldsymbol{x}^{(0)} = [50 \;\; 90 \;\; 14]^T$ (mm) ⇨ $\boldsymbol{x}^{(4)} = [51.51 \;\; 96.39 \;\; 15.004]^T$ (mm)

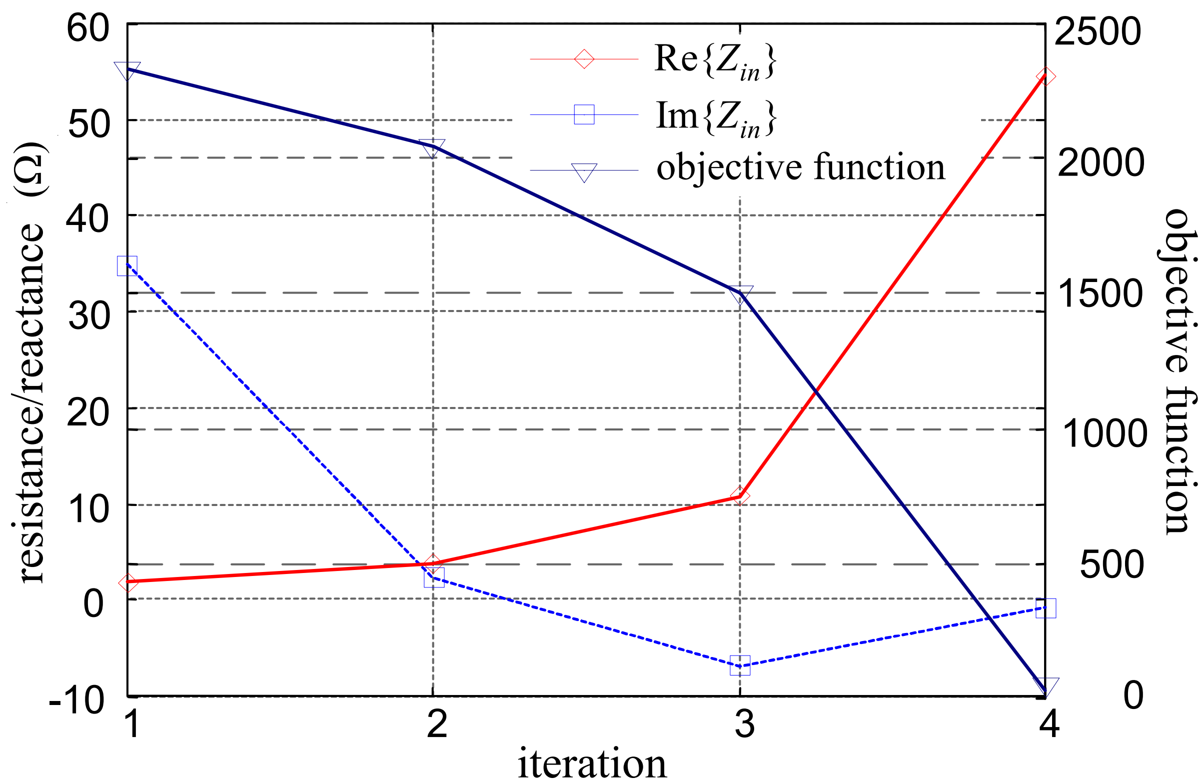


Fig. 10. The progress of the objective function during the optimization of the input impedance of the patch antenna.

Conclusions

The AVM is implemented into a feasible technique for the frequency-domain DSA of HF structures

- reduction of the CPU time requirements for the DSA by a factor of n *to* $(n+1)$
- improved accuracy and convergence
- feasibility: does not require significant modification of existing codes

Factors affecting the accuracy

finite differences with the $\partial_{x_i}\boldsymbol{Z}$ matrix: insignificant

finite differences with $\partial_{I_k} f$: significant

Conclusions

Applications of the DSA based on the AVM

- optimization
- modeling
- statistical and yield analysis

Limitations

- linear frequency-domain analysis
- extension to nonlinear frequency-domain analysis is straightforward