USE OF THE ANALOG NEURAL NETWORKS IN THE ADAPTIVE ANTENNA CONTROL SYSTEMS

Zdeněk TOBEŠ, Zbyněk RAIDA
Dept. of Radio Electronics
Brno University of Technology
Purkyňova 118, 612 00 BRNO
Czech Republic

Abstract

In the paper, original control system of adaptive antennas, which is based on Kalman filter, is presented and compared with earlier approaches to the problem. The designed control circuit eliminates some disadvantages of the control circuits based on the classical Kalman neural network and the Wang one, and enables a real time processing of quickly changing signals processed by adaptive antennas. Especially, the dependence of the convergence rate on ratio of eigenvalues and the risk of instability are significantly reduced.

Keywords

Analog recurrent neural networks, Kalman filter, adaptive antenna, convergence rate

1. Introduction

An adaptive antenna array is an antenna system, that automatically sets minims of its directivity pattern to directions from which the most powerful interference signals come. Other words, the directivity pattern of the adaptive antenna is synthesized to optimize the signal to interference ratio at the antenna output. Or - an adaptive antenna can be understood as a spatial filter which passes signals coming from the main lobe direction (from which the desired signal comes) and which adaptively suppresses signals arriving from other directions.

At present, most adaptive antennas are based on the pilot signal method and the steering vector one.

In the case of the pilot signal, a transmitter transmits a signal, which is known at the receiving side during a learning period. Hence, an error signal, which equals to the difference between the desired pilot and the actual signal at the antenna output, can be defined and the mean squared error can be minimized to obtain an optimal signal to interference ratio at the antenna output.

In the case of the steering vector, the mean power of signal at the antenna output is minimized. If the minimization is constrained in order not to influence parameters of the antenna system in the main lobe direction (from which a desired signal comes), then the minimization reduces interference signals only and the signal to interference ratio at the antenna output is optimized again.

From the optimization point of view, adaptive antennas are parallel systems, which are asked to be very fast in most applications. Since digital processors work in a sequential way and since parallel multiprocessor systems are expensive, analog parallel processors seem to be a suitable alternative for the control of an optimization process (since analog circuits can work quickly and they are relatively cheap). One approach consists in using analog neural networks. These networks are analog circuits based on certain adaptive algorithm (Least Mean Squares [3], Kalman filter [12]) and consisting of certain number of same cooperating units. Since the neurons work in a parallel way, these networks are suitable for solving sets of simultaneous equations. Since steering algorithms for the control of adaptive antennas can be converted to the form of a set of simultaneous linear equations, the analog neural networks are applicable also in adaptive antennas.

In the paper, we compare adaptive antenna control system based on the Wang network (WN) and originally developed adaptive antenna control system based on the Simplified Kalman network (SKN) and its improvement, which exhibits better convergence properties.

2. Pilot Signal Systems

The pilot signal method is based on transmitting a pilot signal by transmitter during the adaptation period, that is known at receiving side and which serves for adaptation of the antenna to the interference environment. The mean squared difference between the signal on the antenna output and the known "pilot" signal (mean square error) is proportional to the power of received interference signals. If mean squared error is minimized, then the power of received interference signals is minimized too. The pilot signal has to be of similar spectral and directional characteristics as the incoming signal (desired signal), which is transmitted after the end of adaptation.

The removal of interferences is provided by changes of directivity pattern, which leads to the minimization of the interference power. Setting the main lobe and nulls of directivity pattern to desired directions is reached by complex weighting of signals at outputs of antenna elements. Optimal complex weights (producing optimal signal to interference ratio in Wiener sense) can be sought by gradient algorithms, which iteratively change setting of weights in contra-direction of the gradient of the mean squared error.

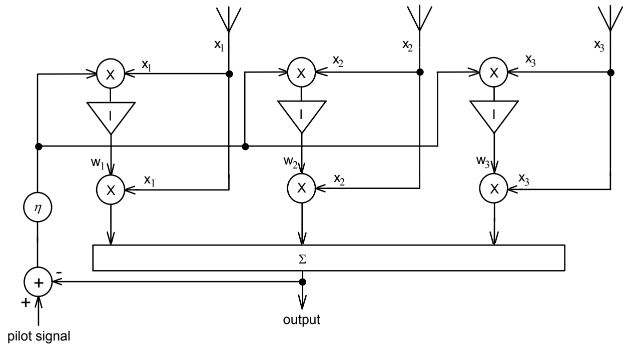

Fig. 1 The pilot-signal-based adaptive antenna controlled by WN.

As already mentioned, the difference between the pilot signal and the signal at the antenna output (the error signal) is expressed according to [8] as

$$e(t) = d(t) - \mathbf{X}^{\mathrm{H}}(t)\cdot\mathbf{W}(t). \qquad (1)$$

In (1), $\mathbf{W}(t)$ is a vector of complex weights at the outputs of antenna elements, $d(t)$ is a pilot signal and $\mathbf{X}(t)$ denotes a vector of signals at the outputs of antenna elements (Fig. 1). The mean squared error is minimal if [8]

$$\nabla_{\mathbf{W}} E\left[e(t)e^{*}(t)\right] = 0 \qquad (2)$$

where E denotes the mean value and * complex conjugate value. Substituting (1) into (2) and performing a few mathematical manipulations [8], Wiener-Hoff equation for computing optimal weights is obtained

$$\mathbf{R}\mathbf{W}_{opt} = \mathbf{P}\,. \qquad (3)$$

Here, $\mathbf{R} = E\,\{\mathbf{X}(t)\mathbf{X}^{\mathrm{H}}(t)\}$ is the auto-correlation matrix of signals at the antenna elements outputs, $\mathbf{P} = E\,\{d^{*}(t)\mathbf{X}(t)\}$ is the cross-correlation vector of the pilot signal and signals at the outputs of antenna elements, and $^{\mathrm{H}}$ denotes Hermitian conjugate value.

The pilot signal method is very popular because of its simplicity. Unfortunately, there are problems with realization of the pilot signal. Estimation of its proper statistical parameters is the first hard nut, and problem of synchronizing the received pilot with generated one is second one.

2.1 Adaptive Antenna Controlled by Wang Network

Adaptive antenna controlled by WN is depicted in Fig. 1. Blocks containing "I" denote integrators; blocks containing "X" represent analog multipliers.

The task solved by ANN is the Wiener-Hoff equation, where the estimates of the left-side matrix elements are computed by couples of multipliers. Hence, the respective Wang ANN can be described as follows:

$$\frac{d\mathbf{W}(t)}{dt} = \frac{\eta}{C_{int}R_{int}}\left[\mathbf{X}(t)\,d^{*}(t) - \mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\,\mathbf{W}(t)\right]. \qquad (4)$$

Here, $C_{int}R_{int}$ is a time constant of the integrator and η is an adaptation constant. The weighting vector can be obtained when solving (4) by the variation of parameters. Handling with the equation

$$\frac{d\mathbf{W}(t)}{dt} + \frac{\eta}{C_{int}R_{int}}\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\,\mathbf{W}(t) = 0\,, \qquad (5)$$

the vector $\mathbf{W}(t)$ is obtained in the form

$$\mathbf{W}(t) = \left\{\exp\left[\int_0^t \frac{-\eta}{C_{int}R_{int}}\,\mathbf{X}(\tau)\;\mathbf{X}^{\mathrm{H}}(\tau)\;d\tau\right]\right\}\mathbf{C}\,, \qquad (6)$$

where $\mathbf{C}$ is a vector of constant elements. In order to obtain the solution of (4), we suppose the vector $\mathbf{C}$ being time-dependent. Then, eqn. (4) can be rewritten to

$$\left\{\frac{d\,\mathbf{C}(t)}{dt} + \frac{-\eta}{C_{int}R_{int}}\;\mathbf{X}(t)\;\mathbf{X}^{\mathrm{H}}(t)\;\mathbf{C}(t) + \right.$$
$$\left. + \;\frac{\eta}{C_{int}R_{int}}\;\mathbf{X}(t)\;\mathbf{X}^{\mathrm{H}}(t)\;\mathbf{C}(t)\right\}\cdot$$
$$\cdot\left\{\exp\left[\int_0^t \frac{-\eta}{C_{int}R_{int}}\;\mathbf{X}(\tau)\;\mathbf{X}^{\mathrm{H}}(\tau)\;d\tau\right]\right\}$$
$$= \frac{\eta}{C_{int}R_{int}}\;\mathbf{X}(t)\;d^{*}(t)\cdot \qquad (7)$$

From this, we can express $\mathbf{C}(t)$ as

$$\mathbf{C}(t) = \int_0^t \left\{\frac{\eta}{C_{int}R_{int}}\left[\exp\int_0^\tau \frac{\eta}{C_{int}R_{int}}\;\mathbf{X}(t)\;\mathbf{X}^{\mathrm{H}}(t)\;dt\right]\cdot\right.$$
$$\left.\cdot\;\mathbf{X}(\tau)\,d^{*}(\tau)\,d\tau\right\} + \mathbf{K} \qquad (8)$$

Then, the relation for $\mathbf{W}(t)$ can be rewritten to

$$\mathbf{W}(t) = \left\{\exp\int_0^t \frac{-\eta}{C_{int}R_{int}}\;\mathbf{X}(\tau)\;\mathbf{X}^{\mathrm{H}}(\tau)\,d\tau\right\}\cdot\{\;\mathbf{K} +$$
$$\left.\int_0^t \frac{\eta}{C_{int}R_{int}}\left[\exp\int_0^\tau \frac{\eta}{C_{int}R_{int}}\;\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\,dt\right]\mathbf{X}(\tau)\,d^{*}(\tau)\,d\tau\right\} \qquad (9)$$

where $\mathbf{K}$ is a vector of constant elements, which are evaluated considering initial conditions (substituting $t = 0$ to eqn. 9). First, time integrals in (9) are evaluated. As stated in [14], we replace the exponential function according to

$$\exp\int \frac{\eta}{C_{int}R_{int}}\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\,dt = \exp\left[\frac{\eta\,t}{C_{int}R_{int}}\;\mathbf{R}\right]. \qquad (10)$$

Hence, time integral in brackets of (9) can be rewritten to

$$\int_0^t \frac{\eta}{C_{int}R_{int}}\left[\exp\int_0^\tau \frac{\eta}{C_{int}R_{int}}\;\mathbf{X}(t)\;\mathbf{X}^{\mathrm{H}}(t)\,dt\right]\mathbf{X}(\tau)\,d^{*}(\tau)\,d\tau =$$
$$= \int_0^t \frac{\eta}{C_{int}R_{int}}\;\exp\left[\frac{\eta\,\tau}{C_{int}R_{int}}\,\mathbf{R}\right]\mathbf{X}(\tau)\,d^{*}(\tau)\,d\tau \qquad (11)$$

and evaluated as

$$\int_0^t \frac{\eta}{C_{int}R_{int}} \exp\left[\frac{\eta\tau}{C_{int}R_{int}}\mathbf{R}\right]\mathbf{X}(\tau)\, d^*(\tau)\, d\tau =$$

$$= \exp\left[\frac{\eta t}{C_{int}R_{int}}\mathbf{R}\right]\mathbf{R}^{-1}\, E\{\mathbf{X}(t)\, d^*(t)\} \tag{12}$$

where the product $\mathbf{X}(t)d^*(t)$ is replaced by its averaged value [14]. Hence, (9) be can rewritten to

$$\mathbf{W}(t) = \exp\left[\frac{-\eta t}{C_{int}R_{int}}\mathbf{R}\right]\left\{\exp\left[\frac{\eta t}{C_{int}R_{int}}\mathbf{R}\right]\cdot\mathbf{R}^{-1}\cdot E\{\mathbf{X}(t)\, d^*(t)\} + \mathbf{K}\right\}. \tag{13}$$

For $t = 0$, a constant vector $\mathbf{K}$ can be expressed as

$$\mathbf{K} = -\mathbf{R}^{-1}\, E\{\mathbf{X}(t)\, d^*(t)\} + \mathbf{W}(0)\,. \tag{14}$$

Then, the average weighting vector converges to

$$\overline{\mathbf{W}}(t) = \left\{\mathbf{R}^{-1}\, E\{\mathbf{X}(t)\, d^*(t)\} + \exp\left[\frac{-\eta t}{C_{int}\, R_{int}}\mathbf{R}\right]\mathbf{W}(0)\right\} + - \exp\left[\frac{-\eta t}{C_{int}R_{int}}\mathbf{R}\right]\mathbf{R}^{-1}\, E\{\mathbf{X}(t)\, d^*(t)\}\,. \tag{15}$$

Using transform [14], we can obtain

$$T_j = [C_{int}R_{int}]/[\eta\lambda_j], \qquad j = 1, 2, \ldots, N, \tag{16}$$

where λ_j represents j-th eigenvalue of the matrix $\mathbf{R}$ and N is a number of eigenvalues. This is the same result as in the case of the Wang network for solution of set of simultaneous linear equations [17]. It can be concluded, that convergence properties are not very good. From (16), it can be seen that the bigger ratio of eigenvalues, the more slowly the network converges. Another problem is the use of optimal setting of adaptive parameters for another ratio of eigenvalues, which can result in creating of an unstable state of the circuit, especially when the used ratio of eigenvalues is higher than the ratio of eigenvalues, for which are the circuit parameters set as optimal. Therefore, the improvements of the Wang network were made – see, for example, [3], [12], [14] and [15].

2.2 Adaptive Antenna Controlled by Simplified Kalman Network

SKN can control pilot-signal based adaptive antennas as well. On the basis of properties of SKN, discussed in [18], we can say that the control circuit based on SKN should exhibit better convergence properties than such circuit based on Wang network. An adaptive antenna controlled by SKN is depicted in Fig. 2. Here, blocks P and I denote predictor and integrator respectively, which are depicted in Fig. 4. Mathematical description of the circuitry is based on the equations of non-simplified Kalman filter [12]. These equations can be rewritten to

$$\frac{d\mathbf{W}(t)}{dt} = \frac{\eta}{C_{int}R_{int}}\,\mathbf{K}(t)\left[\mathbf{X}(t)d^*(t) - \mathbf{X}(t)\mathbf{X}^{\mathrm{H}}(t)\mathbf{W}(t)\right], \tag{17}$$

$$\mathbf{K}(t) = \mathbf{P}(t)\left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]^{\mathrm{T}}\rho\,, \tag{18}$$

$$\frac{d\,\mathbf{P}(t)}{dt} = -\frac{1}{C_{int}R_{int}}\,\mathbf{K}(t)\,\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\,\mathbf{P}(t) \tag{19}$$

by considering the Wiener-Hoff equation as the solved task. Here, $C_{int}R_{int}$ denotes a time constant of Kalman predictor and a time constant of the integrator (see Fig. 4), which are considered to be equal, and η is an adaptation coefficient (see Fig. 2), which will be for simplicity in following derivation considered to be equal to one. Next, $\mathbf{P}(t)$ is a predicted state-error, ρ is a coefficient representing a residual error, $\mathbf{K}(t)$ denotes a matrix of a Kalman gain and $\mathbf{X}(t)$ is the vector of input signal. The predicted state error can be expressed as

$$d\mathbf{P}/dt = -\mathbf{P}(t)\left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]^{T}\left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]\cdot \mathbf{P}(t)\ \rho/[C_{int}\, R_{int}]. \tag{20}$$

This can be arranged to

$$\int[\mathbf{P}(t)\,\mathbf{P}(t)]^{-1}\, d\mathbf{P}(t) + \mathbf{I}\,C =$$

$$= -\int\left\{\left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]^{\mathrm{T}}\left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]\rho\, dt/[C_{int}\, R_{int}]\right\}, \tag{21}$$

where C denotes a constant. Considering

$$-\int\left\{\left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]^{\mathrm{T}}\left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]\rho\, dt/[C_{int}\, R_{int}]\right\} =$$

$$= -\mathbf{R}^{\mathrm{T}}\,\mathbf{R}\,[t\rho]/[C_{int}\, R_{int}], \tag{22}$$

we can find that

$$\mathbf{P}(t) = \left(\frac{\rho t}{C_{\mathrm{int}}R_{\mathrm{int}}}\mathbf{R}^{T}\mathbf{R} + \mathbf{I}C\right)^{-1}. \tag{23}$$

Here, $\mathbf{R}$ is auto-correlation matrix of signals at outputs of antenna elements. The constant C can be determined from the initial condition $\mathbf{P}(t)$=$\mathbf{P}_0$. Then, we can find

$$C = 1/P_0 \tag{24}$$

and therefore

$$\mathbf{P}(t) = \left[\mathbf{R}^{\mathrm{T}}\mathbf{R}\,(\rho t)/(C_{int}\, R_{int}) + \mathbf{I}/P_0\right]^{-1}. \tag{25}$$

Now, Kalman gain can be expressed as

$$\mathbf{K}(t) = \left[\mathbf{R}^{\mathrm{T}}\mathbf{R}\,(\rho t)/(C_{int}R_{int}) + \mathbf{I}/P_0\right]^{-1}\cdot \left[\mathbf{X}(t)\mathbf{X}^{\mathrm{H}}(t)\right]^{\mathrm{T}}\rho. \tag{26}$$

We substitute (26) to (17) and express the weighting vector. Then, (17) can be rewritten to

$$\frac{d\mathbf{W}(t)}{dt} = \frac{1}{C_{int}\,R_{int}} \left[\frac{\rho\, t}{C_{int}\,R_{int}}\,\mathbf{R}^{\mathrm{T}}\mathbf{R} + \frac{\mathbf{I}}{P_0} \right]^{-1} \cdot \tag{27}$$

$$\cdot \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]^{\mathrm{T}} \rho \left[\mathbf{X}(t)\,d^*(t) - \mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\mathbf{W}(t)\right].$$

In order to obtain the solution of (27), we solve first

$$\frac{d\mathbf{W}(t)}{dt} + \frac{1}{C_{int}\,R_{int}} \left[\frac{\rho\, t}{C_{int}\,R_{int}}\,\mathbf{R}^{\mathrm{T}}\mathbf{R} + \frac{\mathbf{I}}{P_0} \right]^{-1} \cdot$$

$$\rho \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]^{\mathrm{T}} \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\mathbf{W}(t)\right] = 0. \tag{28}$$

Applying a similar procedure as used in Section 2.1, we can obtain

$$\mathbf{W}(t) = \mathbf{C}' \exp\int\left\{ \frac{-\rho}{C_{int}\,R_{int}} \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]^{\mathrm{T}} \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right] \cdot \right.$$

$$\left. \cdot \left[\frac{\rho\, t}{C_{int}\,R_{int}}\,\mathbf{R}^{\mathrm{T}}\mathbf{R} + \frac{\mathbf{I}}{P_0} \right]^{-1} \right\} dt \cdot \tag{29}$$

Here **C'** is a vector of constants. In order to obtain the solution of (27), we assume the vector of constants **C'** being time-dependent. Then, we can rewrite (27) to

$$\frac{d\mathbf{C}'(t)}{dt} \exp\int\left\{ \frac{-\rho}{C_{int}\,R_{int}} \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]^{\mathrm{T}} \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right] \cdot \right.$$

$$\left. \left[\frac{\rho\, t}{C_{int}\,R_{int}}\mathbf{R}^{\mathrm{T}}\mathbf{R} + \frac{\mathbf{I}}{P_0} \right]^{-1} \right\} dt = \frac{1}{C_{int}\,R_{int}} \left[\frac{\rho\, t}{C_{int}\,R_{int}}\mathbf{R}^{\mathrm{T}}\mathbf{R} + \frac{\mathbf{I}}{P_0} \right]^{-1} \cdot$$

$$\cdot \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]^{\mathrm{T}} \left[\mathbf{X}(t)\,\rho\, d^*(t)\right]. \tag{30}$$

Hence, the constant **C'**(t) can be expressed as

$$\mathbf{C}'(t) = \int\left\{ \frac{1}{C_{int}\,R_{int}} \left[\frac{\rho\, t}{C_{int}\,R_{int}}\,\mathbf{R}^{\mathrm{T}}\mathbf{R} + \frac{\mathbf{I}}{P_0} \right]^{-1} \cdot \right.$$

$$\left. \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]^{\mathrm{T}} \left[\mathbf{X}(t)\,d^*(t)\ \rho\ \gamma\right] \right\} dt + \mathbf{K}', \tag{31}$$

where

$$\gamma = \exp\int\left\{ \frac{1}{C_{int}\,R_{int}} \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]^{\mathrm{T}} \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right] \right.$$

$$\left. \left[\frac{\rho\, t}{C_{int}\,R_{int}}\,\mathbf{R}^{\mathrm{T}}\mathbf{R} + \frac{\mathbf{I}}{P_0} \right]^{-1} \right\} dt \tag{32}$$

and **K'** is a vector of constants. Now, we need to evaluate time integrals in (31). We note, that (31) is rewritten to

$$\mathbf{C}'(t) = \int\left\{ \frac{d\gamma}{dt} \left[\mathbf{X}(t)\mathbf{X}^{\mathrm{H}}(t)\right]^{-1} \left[\mathbf{X}(t)\,d^*(t)\right] \right\} dt + \mathbf{K}' \cdot \tag{33}$$

Finally, (33) can be rearranged (considering similar assumptions as in the previous chapter) to

$$\mathbf{C}'(t) = \gamma\ \mathbf{R}^{-1}\ E\left\{\mathbf{X}(t)\,d^*(t)\right\} + \mathbf{K}' . \tag{34}$$

Now, we can express the weighting vector as

$$\mathbf{W}(t) = \mathbf{R}^{-1}\ E\left\{\mathbf{X}(t)\,d^*(t)\right\} +$$

$$+ \exp\int\left\{ \frac{-\rho}{C_{int}\,R_{int}} \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]^{\mathrm{T}} \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right] \cdot \right.$$

$$\left. \cdot \left[\frac{\rho\, t}{C_{int}\,R_{int}}\,\mathbf{R}^{\mathrm{T}}\mathbf{R} + \frac{\mathbf{I}}{P_0} \right]^{-1} \right\} dt\ \mathbf{K}' \cdot \tag{35}$$

As already stated, the exponential function can be replaced according to

$$\exp\int\left\{ \frac{-\rho}{C_{int}\,R_{int}} \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right]^{\mathrm{T}} \left[\mathbf{X}(t)\,\mathbf{X}^{\mathrm{H}}(t)\right] \cdot \right.$$

$$\left. \left[\frac{\rho\, t}{C_{int}\,R_{int}}\,\mathbf{R}^{\mathrm{T}}\mathbf{R} + \frac{\mathbf{I}}{P_0} \right]^{-1} \right\} dt =$$

$$= \exp\left\{ -\ln\left[\rho\, t\ P_0\ \mathbf{R}^{\mathrm{T}}\mathbf{R} + \mathbf{I}\ C_{int}\ R_{int}\right]\right\}, \tag{36}$$

and (35) can be rewritten to

$$\mathbf{W}(t) = \mathbf{R}^{-1}\ E\left\{\mathbf{X}(t)d^*(t)\right\} + \left[\rho t P_0\ \mathbf{R}^{\mathrm{T}}\mathbf{R} + \mathbf{I} C_{int} R_{int}\right]\mathbf{K}' . \tag{37}$$

In order to find an expression for constant vector **K'**, $t = 0$ is substituted to (37). Then

$$\mathbf{K}' = -\,C_{int}\,R_{int}\ \mathbf{R}^{-1}\ E\left\{\mathbf{X}(t)\,d^*(t)\right\} + \mathbf{W}(0)\,C_{int}\,R_{int} . \tag{38}$$

Finally, we can express time course of the average weighting vector

$$\overline{\mathbf{W}}(t) = \left[\mathbf{I} - C_{int}\,R_{int}\left(\rho\, t\, P_0\ \mathbf{R}^{\mathrm{T}}\mathbf{R} + \mathbf{I}\,C_{int}\,R_{int}\right)^{-1}\right] \cdot$$

$$\cdot\, \mathbf{R}^{-1}\ E\left\{\mathbf{X}(t)\,d^*(t)\right\} +$$

$$+ C_{int}\,R_{int}\ \left[\rho\, t\, P_0\ \mathbf{R}^{\mathrm{T}}\mathbf{R} + \mathbf{I}\,C_{int}\,R_{int}\right]^{-1}\ \mathbf{W}(0). \tag{39}$$

Convergence properties of the described adaptive antenna system (including influence of eigenvalue ratio) are given by the term $[\rho P_0 t\, \mathbf{R}^{\mathrm{T}}\mathbf{R} + \mathbf{I}\ C_{int} R_{int}]^{-1}$, which are very similar to convergence properties of Kalman neural network, described in [18]. We can use the matrix transform described in [14] in order to obtain the dependence on eigenvalues: following [14], we define $\mathbf{W}'(t) = \mathbf{E}^{-1}\mathbf{W}(t)$, $\mathbf{X}'(t) = \mathbf{E}^{-1}\mathbf{X}(t)$, where **E** denotes a transform matrix. We multiply (39) by $\mathbf{E}^{-1}$ and rewrite (39) to

$$\mathbf{W}'(t) = \left[\mathbf{I} - \left(\Lambda \frac{P_0 \rho t}{C_{int} R_{int}} + \mathbf{I}\right)^{-1}\right] \mathbf{R}^{-1}\left\{\mathbf{X}'(t) d^*(t)\right\} \cdot \tag{40}$$

Here, Λ represents the matrix of eigenvalues. Comparing (40) to (16), the form of the convergence process of Wang ANN is given by the exponential function of a negative argument. If argument is a function of time, its decrease is quicker than inverted value of time. Computer simulations show that circuit from Fig. 2 converges more quickly for the same input signal, which can be explained by the influence of real circuit elements (in ideal case, both networks can converge to zero within very short time, because we are not limited in setting of adaptive parameters).

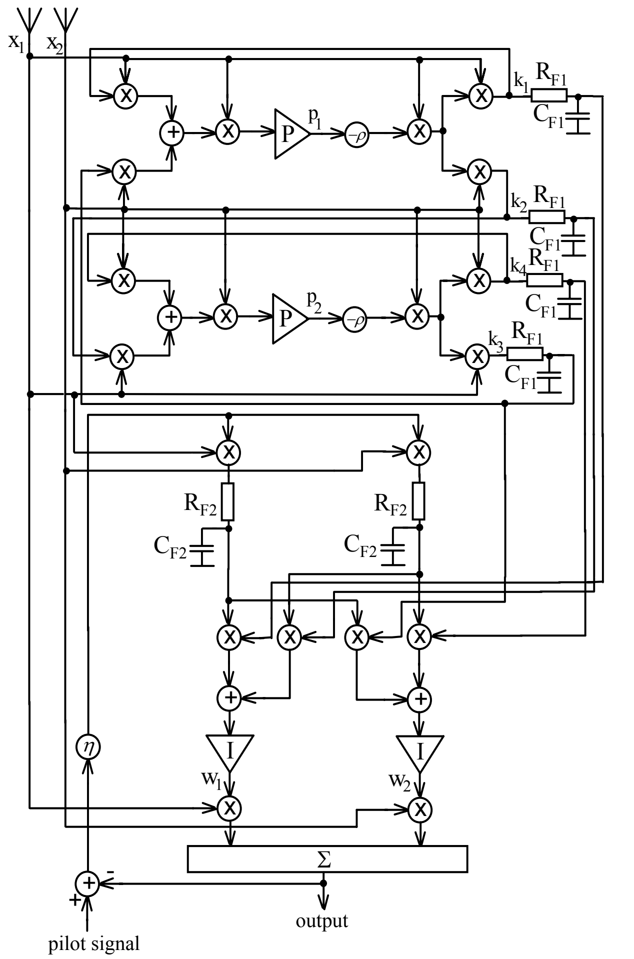

Fig. 2 Pilot-signal-based adaptive antenna controlled by classical simplified Kalman ANN.

In real case, convergence rate is limited by properties of real circuit elements (especially opamps), which can cause an unstable state of ANN, if there is a high gain of the corresponding closed loops in the circuit (caused by high or low value of an adaptation parameter, e.g.). Higher convergence rate of SKN is caused by multiplying the error signal by Kalman gain, which is decreasing with time very quickly. Therefore, gains of the closed loops in the lower part of SKN are not time-independent - they are decreasing in time due described multiplying. Hence, creating an unstable state requires higher values of gains of closed loops at the beginning of the convergence process comparing to Wang ANN, which leads to better stability and convergence properties. In case of SKN, there are also more possibilities of influencing closed-loops gains because Kalman gain depends on the input signal and on the upper-circuit adaptation parameters, which allows finer tuning of the network.

Properties of the SKN-based adaptive antenna can be improved applying improved versions of the classical SKN [18]. An analog realization of adaptive antenna controlled by improved SKN is depicted in Fig. 3. Here, the block P denotes a predictor, and the block I' is a modified integrator from Fig. 5. Such network for solving of simultaneous linear equations was at first presented in [18], where it was

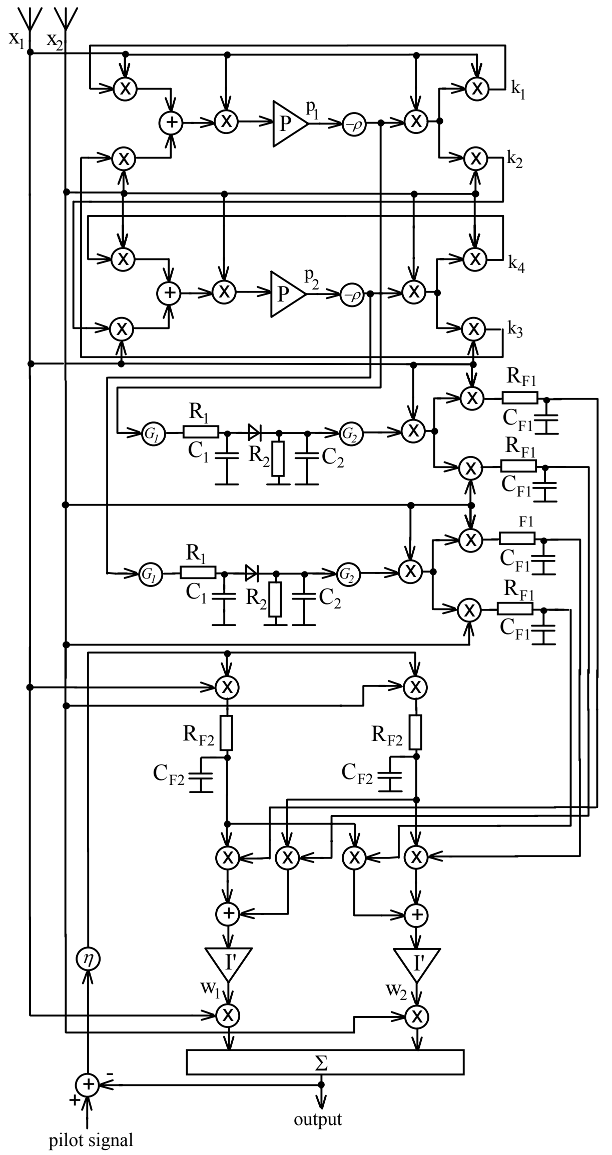

Fig. 3 The pilot-signal-based adaptive antenna controlled by improved SKN.

created by combining of separately developed improvements, which led to the shortest convergence time of all developed improved networks. Unfortunately, time course of combined SKN cannot be expressed analytically due to the enormous complexity; only the SKN with only one improvement used in circuit from Fig. 3 can be analyzed, which was performed in [18]. However, the computer simulations show that these improvements separately used do not lead to strong decreasing of the convergence rate and dependence of the convergence rate on the eigenvalue ratio of the input signal matrix, but its combination exhibits such properties (see part 4 of this work). Considering use of adaptive antennas in such areas as mobile communications are [10], [11], such properties of the control circuit are the significant contribution to its errorless function.

3. Steering Vector Systems

The steering vector method is based on such minimization of mean output power, which does not influence properties of the antenna in the main lobe direction, from

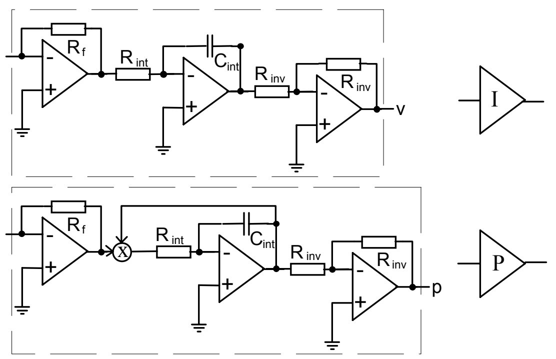

Fig. 4 Blocks used in the control circuit based on SKN: a) integrator, b) predictor

which the desired signal comes, and reduces mean power of the interference signals at the output of the antenna, since interferences are assumed to come from other directions. The mean output power [8] can be expressed as

$$\min_{\mathbf{W}} \left\{ E\left[y(t) y^*(t) \right] \right\}, \tag{41}$$

and its minimization is constrained by the requirement of fixed properties of directivity pattern in the main lobe direction [8]

$$\mathbf{S}^{\mathrm{T}} \mathbf{W}(t) = 1 . \tag{42}$$

Here, $y(t)$ is the output signal of the adaptive antenna and $\mathbf{S}$ is column vector, which is generally of the form

$$\mathbf{S} = \left[1, \exp\left(j \frac{2\pi d}{\lambda_0} \cos\theta_0 \right), \ldots, \exp\left(j \frac{2\pi d}{\lambda_0} (M-1) \cos\theta_0 \right) \right]^{\mathrm{T}} . \tag{43}$$

Here, M is a number of elements of the antenna array, d is the element spacing, λ_0 is the wavelength of the plane wave in free space and θ_0 is the look direction angle (the angle between the axis of the linear antenna array and the direction of the arrival of the desired signal).

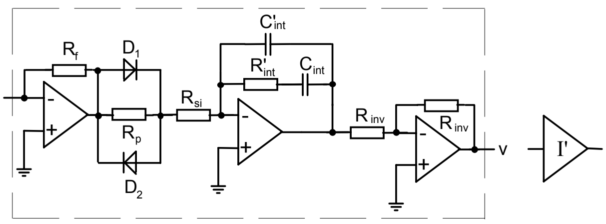

Fig. 5 Modified integrator

The constrained minimization can be solved by the method of Lagrange multipliers [5] producing

$$L = E\left[y(t)\, y^*(t) \right] + \lambda \left[\mathbf{C}^{\mathrm{T}} \mathbf{W}(t) - 1 \right], \tag{44}$$

which is to be minimized. In (44), λ denotes the Lagrange multiplier. The necessary condition for the minimum of (44) can be expressed as

$$\nabla_{\mathbf{W}} L = 0 . \tag{45}$$

Substituting (44) into (45), considering (for simplicity) $\mathbf{S}$ to be the column vector of ones (the desired signal has the same phase at all elements) and evaluating the resultant equation yields the following relation

$$\begin{bmatrix} E\{x_1^*(t)x_1(t)\} & E\{x_1^*(t)x_2(t)\} & \ldots & E\{x_1^*(t)x_N(t)\} & 0.5 \\ E\{x_2^*(t)x_1(t)\} & E\{x_2^*(t)x_2(t)\} & \ldots & E\{x_2^*(t)x_N(t)\} & 0.5 \\ \ldots & \ldots & & \ldots & \ldots \\ E\{x_N^*(t)x_1(t)\} & E\{x_N^*(t)x_2(t)\} & \ldots & E\{x_N^*(t)x_N(t)\} & 0.5 \\ 1 & 1 & \ldots & 1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \ldots \\ w_N \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \ldots \\ 0 \\ 1 \end{bmatrix}, \tag{46}$$

which provides (N+1) relations for N unknown complex weights w and an unknown Lagrange multiplier λ.

The adaptive antenna based on this control system was so far designed by 2 ways: The first one was published in [5], where solving of the quadratic programming problem was applied to the Wang network, and the second one was published in [4] and [6] and deals with the Hopfield network. Because of relative bad properties of the Hopfield network, the first approach only will be mentioned bellow.

3.1 Adaptive Antenna Controlled by Wang Network

Since the above-described steering vector system is the complex-value constrained quadratic programming problem, which cannot be solved by the neural network directly, it is necessary to separate the real and the imaginary parts of elements of the weighting vector and the matrix of the input signal. [4] shows that the product of complex values $\mathbf{W}^{\mathrm{H}}\mathbf{RW}$**,** which represents the mean output power of the adaptive antenna and which is to be minimized, can be converted to the form

$$\mathbf{W}^{\mathrm{H}} \mathbf{R}\, \mathbf{W} = \mathbf{v}^{\mathrm{T}} \begin{bmatrix} \mathbf{R}_r & -\mathbf{R}_i \\ \mathbf{R}_i & \mathbf{R}_r \end{bmatrix} \mathbf{v}, \tag{47}$$

where

$$\mathbf{v} = \left[\mathbf{W}_r \quad \mathbf{W}_i \right]^{\mathrm{T}} . \tag{48}$$

Here, $\mathbf{W}_r$ is real part of $\mathbf{W}$, $\mathbf{W}_i$ is imaginary part of $\mathbf{W}$, $\mathbf{R}_r$ and $\mathbf{R}_i$ denote real and imaginary part of $\mathbf{R}$, respectively.

Discussing the way of the design of ANN for solving of this task, problem is described as a minimization of

$$\Phi(\mathbf{v}) = \tfrac{1}{2} \mathbf{v}^{\mathrm{T}} \mathbf{G}\, \mathbf{v} \tag{49}$$

subject to

$$\mathbf{C}\,\mathbf{v} = \mathbf{b}\,, \tag{50}$$

where

$$\mathbf{G} = \begin{pmatrix} \mathbf{R}_r & -\mathbf{R}_i \\ \mathbf{R}_i & \mathbf{R}_r \end{pmatrix}, \tag{51}$$

$$\mathbf{b} = \begin{pmatrix} 1 & 0 \end{pmatrix}^{\mathrm{T}}, \tag{52}$$

$$\mathbf{C} = \begin{pmatrix} \mathbf{S}_r^{\mathrm{T}} & \mathbf{S}_i^{\mathrm{T}} \\ -\mathbf{S}_i^{\mathrm{T}} & \mathbf{S}_r^{\mathrm{T}} \end{pmatrix}. \tag{53}$$

Here, **C** is created same way as **G** (for more information see [4]), $\mathbf{S}_r$ and $\mathbf{S}_i$ denote here real and imaginary part of **S**.

The Langrangian of the quadratic programming problem with equality constraints is defined according to [5] as

$$L(\mathbf{G},\lambda) = \tfrac{1}{2}\mathbf{v}^{\mathrm{T}}\mathbf{G}\,\mathbf{v} + \lambda^{\mathrm{T}}(\mathbf{C}\mathbf{v} - \mathbf{b}) \tag{54}$$

where λ is two-dimensional column vector of Langrangian multipliers (the length of this vector is determined by the vector **b**). From the condition (see [2])

$$\nabla_{\mathbf{W}} L = 0\,, \tag{55}$$

we can obtain matrix-form algebraic equation

$$\begin{pmatrix} \mathbf{G} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{C} & \mathbf{N}_1 \end{pmatrix} \begin{pmatrix} \mathbf{v} \\ \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{N}_2 \\ \mathbf{b} \end{pmatrix}, \tag{56}$$

where $\mathbf{N}_1$ is a null matrix of size 2×2, and $\mathbf{N}_2$ is n-dimensional null column vector. According to [5], $(n+2)$-dimensional linear system of algebraic equations in (55) has a unique solution if **G** is positive definite and **C** has full rank (n is the dimension of **G**). The dynamics of Wang ANN for the solution of this problem can be described by

$$\frac{d\mathbf{x}(t)}{dt} = -\eta\,\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\,\mathbf{x}(t) + \eta\,\mathbf{Z}^{\mathrm{T}}\mathbf{y} \tag{57}$$

where

$$\mathbf{Z} = \begin{pmatrix} \mathbf{G} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{C} & \mathbf{N}_1 \end{pmatrix}, \tag{58}$$

$$\mathbf{x}(t) = \begin{pmatrix} \mathbf{v}(t) & \lambda(t) \end{pmatrix}^{\mathrm{T}}, \tag{59}$$

$$\mathbf{y} = \begin{pmatrix} \mathbf{N}_2 & \mathbf{b} \end{pmatrix}^{\mathrm{T}}, \tag{60}$$

and η is a learning constant (as mentioned above). The relation (57) can be rewritten to

$$\begin{pmatrix} d\mathbf{v}(t)/dt \\ d\lambda(t)/dt \end{pmatrix} = -\eta \begin{pmatrix} \mathbf{G}^2 + \mathbf{C}\mathbf{C}^{\mathrm{T}} & \mathbf{G}\mathbf{C} \\ \mathbf{C}^T\mathbf{G} & \mathbf{C}^{\mathrm{T}}\mathbf{C} \end{pmatrix} \begin{pmatrix} \mathbf{v}(t) \\ \lambda(t) \end{pmatrix} + \eta \begin{pmatrix} \mathbf{C}\mathbf{b} \\ \mathbf{N}_2 \end{pmatrix}. \tag{61}$$

Wang ANN for solving this problem consists of $(m+2)$ neurons representing variables $v_i(t)$ and $\lambda_i(t)$. The first matrix is denoted as

$$\mathbf{M} = -\eta \begin{pmatrix} \mathbf{G}^2 + \mathbf{C}\mathbf{C}^{\mathrm{T}} & \mathbf{G}\mathbf{C} \\ \mathbf{C}^{\mathrm{T}}\mathbf{G} & \mathbf{C}^{\mathrm{T}}\mathbf{C} \end{pmatrix}, \tag{62}$$

and the biasing threshold vector is denoted as

$$\theta = \eta\begin{pmatrix} \mathbf{C}\mathbf{b} & \mathbf{N}_2 \end{pmatrix}^{\mathrm{T}}. \tag{63}$$

Then, we can rewrite (61) to

$$\begin{pmatrix} d\mathbf{v}(t)\,/\,dt \\ d\lambda(t)\,/\,dt \end{pmatrix} = \mathbf{M}\begin{pmatrix} \mathbf{v}(t) \\ \lambda(t) \end{pmatrix} + \theta. \tag{64}$$

This is the classical task to be solved by Wang ANN of $(m+2)$ neurons. Convergence properties are given by eigenvalues of **M**, as seen from (64). The dependence of the convergence rate upon ratio of eigenvalues of this matrix is similar as this dependence in the case of the classical Wang network, described in [17]. Computer simulations are difficult to perform because of complexity of the circuitry, which causes very high computational requirements – from reason, real opamps can not be used, as well as multipliers providing correlation coefficients. Therefore, it is not necessary to observe the dependence of the convergence rate on the circuit parameters and on the eigenvalue ratio of the input signal matrix, because by setting of circuit parameters an arbitrary low convergence rate can be reached.

3.2 Adaptive Antenna Controlled by the Simplified Kalman Network

Kalman ANN can control the steering-vector antenna too [9]. The basis for deriving control relations is

$$\frac{d\mathbf{x}(t)}{dt} = -\eta\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\mathbf{x}(t) + \eta\mathbf{Z}^{\mathrm{T}}\mathbf{y}\,, \tag{65}$$

which was presented in previous chapter as an equation describing convergence process of Wang ANN for solving of the quadratic programming problem. In order to obtain desired equations for steering vector system implemented by Kalman ANN, relation (65) has to be applied to the basic equations of Kalman filter. Then, (65) can be rewritten to

$$\begin{pmatrix} d\mathbf{v}(t)/dt \\ d\lambda(t)/dt \end{pmatrix} = -\mathbf{K}(t)\begin{pmatrix} \mathbf{G} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{C} & \mathbf{N}_1 \end{pmatrix}\begin{pmatrix} \mathbf{v}(t) \\ \lambda(t) \end{pmatrix} + \mathbf{K}(t)\begin{pmatrix} \mathbf{N}_2 \\ \mathbf{b} \end{pmatrix} \tag{66}$$

where

$$\mathbf{K}(t) = \rho\,\mathbf{P}(t)\begin{pmatrix} \mathbf{G} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{C} & \mathbf{N}_1 \end{pmatrix}^{\mathrm{T}} \tag{67}$$

$$\frac{d\,\mathbf{P}(t)}{dt} = -\mathbf{K}(t)\begin{pmatrix} \mathbf{G} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{C} & \mathbf{N}_1 \end{pmatrix}\mathbf{P}(t)\,. \tag{68}$$

$\mathbf{K}(t)$ is a Kalman gain vector, $\mathbf{P}(t)$ is a vector of predicted state error and ρ is an adaptive parameter. Other symbols denote the same quantities as in the previous chapter. On the basis of presented equations, we can express an equation of the convergence process of weighting vector as

$$\mathbf{v}(t) = \begin{pmatrix} \mathbf{G} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{C} & \mathbf{N}_1 \end{pmatrix}^{-1}\begin{pmatrix} \mathbf{N}_2 \\ \mathbf{b} \end{pmatrix} + \tag{69}$$

$$+\left[\mathbf{v}(0)-\begin{pmatrix}\mathbf{G} & \mathbf{C}^{\mathrm{T}}\\ \mathbf{C} & \mathbf{N}_1\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{N}_2\\ \mathbf{b}\end{pmatrix}\right]\left[\rho t P_0\begin{pmatrix}\mathbf{G} & \mathbf{C}^{\mathrm{T}}\\ \mathbf{C} & \mathbf{N}_1\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}\mathbf{G} & \mathbf{C}^{\mathrm{T}}\\ \mathbf{C} & \mathbf{N}_1\end{pmatrix}+\mathbf{I}\right]^{-1}.$$

The last brackets in (69), i.e. ratio of eigenvalues of the matrix product in these brackets determine convergence properties. Influence of eigenvalue ratio is similar to the classic Kalman ANN. Realization of Kalman ANN for this task is quite easy - it can be solved by the classic Kalman ANN, which is extended to have m+2 neurons. Because the complexity of the steering vector based adaptive antenna controlled by SKN is much more complicated than such adaptive antenna controlled by Wang ANN, the computer simulations were not performed from previous reasons.

4. Computer Simulations

In Tab.2, computer simulations related to the discussed circuits are presented. Computer simulations are performed by using PSPICE for a two-element adaptive antenna based on a pilot signal. Values of circuit parameters were set for the case of optimal convergence time for input signal with frequency 100 kHz (frequency shift between antenna array and ANN is supposed) and other parameters following the third row of the Tab. 1 (all the simulations were performed with this setting of circuit parameters).

	Phase of interference		Phase of signal		Weights	λ_1/λ_2
	1st el.	2nd el.	1st el.	2nd el.		
1	-60°	+60°	-40°	-40°	$\mathbf{v}=\begin{pmatrix}0.449\\0.156\end{pmatrix}$	4
2	-60°	+135°	-40°	-40°	$\mathbf{v}=\begin{pmatrix}0.511\\1.998\end{pmatrix}$	526
3	-60°	+139°	-40°	-40°	$\mathbf{v}=\begin{pmatrix}0.515\\10.088\end{pmatrix}$	13092

Tab. 1 Observed input signals (in all cases, the same input signal frequency is considered, amplitude of all the signals 1V).

	Wang's net	classical SKN	modified SKN
λ_1 / λ_2	t [ms]	t [μs]	t [μs]
4	6.31	divergence	13.32
526	6.55	divergence	16.79
13092	4.26	227.76	20.87

Tab. 2 Dependency of the convergence time on the eigenvalue ratio of the input signal matrix.

5. Conclusion

The paper presents an original control system of the adaptive antenna, based on improved simplified Kalman filter. In section 2, control systems of the pilot signal based adaptive antenna, based on classical Wang and simplified Kalman network, were presented and analyzed. In chapter 3, adaptive antenna based on the steering vector method is discussed in a similar way. By using results of computer simulations, the mentioned control circuits were compared with circuit based on modified simplified Kalman network, which was found to exhibit very high convergence rate and very low dependence on the eigenvalue ratio.

References

[1] ČERNOHORSKÝ, D., NOVÁČEK, Z. Antennas and Propagation of EM Waves. Brno: Brno University of Technology, 1992.

[2] LUENBERGER, D. E. Linear and Nonlinear Programming. Addison Wesley Reading, MA, 1984.

[3] COMPTON, R. T., Jr. Improved Feedback Loop for Adaptive Arrays. IEEE Transactions on Aerospace Electronic Systems. 1980, vol. 16, no. 2, p. 128 – 136.

[4] CHANG, P. R., YANG, W. H., CHAN, K. K. A Neural Network Approach to MVDR Beamforming Problem. IEEE Trans. on Antennas and Propagation. 1992, vol. 40, no. 3, p. 313 – 322.

[5] WANG, J. Recurrent Neural Network for Solving Quadratic Programming Problems with Equality Constraints. Electronics Letters. 1992, vol. 28, no. 14, p. 1345 – 1347.

[6] KENNEDY, M. P., CHUA, L. O. Neural Networks for Nonlinear Programming. IEEE Transactions on Circuits and Systems. 1988, vol. 35, no. 5, p. 554 – 562.

[7] CHEN, Y. H., CHIANG, C. T. Adaptive Beam-Forming Using the Constrained Kalman Filter. IEEE Trans. on Antennas and Propagation. 1993, vol. 41, no. 11, p. 1576 – 1580.

[8] WIDROW, B., MANTEY, P. E., GRIFFITHS, L. J., GOODE, B. B. Adaptive Antenna Systems. Proc. IEEE. 1967, vol. 55, no. 12, p. 2143 – 2159.

[9] TOBEŠ, Z., RAIDA, Z. Analog Neural Networks for the Control of Adaptive Antennas. In Proceedings of the International Symposium on Antennas JINA'96. 1996, Nice (France): France Telecom, p. 601 – 604.

[10] GODARA, L. C. Applications of Antenna Arrays to Mobile Communications, Part I: Performance Improvement, Feasibility, and System Considerations. Proceedings of the IEEE. 1997, vol. 85, no. 7, p. 1031 – 1060.

[11] GODARA, L. C. Applications of Antenna Arrays to Mobile Communications, Part II: Beam-Forming and Direction-of-Arrival Considerations. Proceedings of the IEEE. 1997, vol. 85, no. 8, p. 1195 – 1245.

[12] RAIDA, Z. Improvement of Convergence Properties of Wang's Neural Network. Electronics Letters. 1994, vol. 30, no. 22, p. 1864 - 1866.

[13] RAIDA, Z. Stability of digital adaptive antennas. Ph.D. thesis. Brno: Brno University of Technology, 1994 (in Czech).

[14] KLEMES, M. A. Practical Method of Obtaining Constant Convergence Rates in LMS Adaptive Arrays. IEEE Trans. on Antennas and Propagation. 1986, vol. 34, no. 3, p. 440 - 446.

[15] TOBES, Z., RAIDA, Z. Stability Problems of Wang's Neural Networks. In Proc. of the Conference Radioelektronika '96. Brno (Czech Republic), 1996, p. 366 - 369.

[16] BARTSCH, H.J. Mathematical formulas. Praha: SNTL Praha, 1987 (in Czech).

[17] WANG,J. Electronic Realization of Recurrent Neural Network for Solving Simultaneous Linear Equations. Electronics Letters. 1992, vol. 28, no. 5, p. 493 - 495.

[18] TOBEŠ, Z., RAIDA, Z. Improvements of Analog Neural Networks Based on Kalman Filter. Radioengineering. 2002, vol. 11, no. 1, p. 6 – 13.