

ADAPTIVE ALGORITHMS IN RADIO DIRECTION FINDING

Rudolf SELLER Péter OLASZ, and Gábor GYIMESI

Department of Microwave Telecommunications
Technical University of Budapest
H-1521 Budapest, Hungary
Phone: +36 1 463-3687, Fax: +36 1 463-3289
e-mail: t-seller@nov.mht.bme.hu

Received: Jan. 20, 1998

Abstract

Conventional radio direction finding methods suffer a lack in performance under certain configurations of the radio frequency environment. A typical example is the case of two or more transmitters spaced closely in terms of azimuth angle. Several adaptive algorithms have been introduced to enhance the angular resolution and accuracy of radio measurement. Compared to the traditional methods these algorithms provide considerably higher accuracy in determining the direction of arrival and higher grade of radiating source separation can be achieved. In this paper a brief overview of conventional and two adaptive estimation methods is provided as a literature summary, which is followed by a qualitative analysis and comparison of these three methods in terms of dynamic range and resolution as new results. Finally software simulation results are presented to demonstrate the advantages of adaptive methods as well as their sensitivity to versatile performance degrading conditions.

Keywords: adaptive signal processing, spectral estimation, antennas, radar.

1. Introduction

To distinguish the different adaptive approaches it is necessary to understand the common principle of the radio direction measurement. We assume that a linear antenna array is located in the electromagnetic environment to be measured. It is also assumed that this system is operating under aperture far field conditions, which means that the receiver array is spaced distant enough from all the transmitters so that the incident field can be estimated as a superposition of plane waves (see *Fig. 1*).

Under the above conditions there is a strong parallelism between the well known time \leftrightarrow frequency domain and spatial frequency \leftrightarrow angular domain, that is widely exploited in antenna theory and design. The most spectacular example for this relationship is linear antenna array design, where the design of the array, a spatial filter, is derived from conventional filter design methods. In this approach the transfer function of a frequency domain filter corresponds to the antenna characteristics in the angular domain.

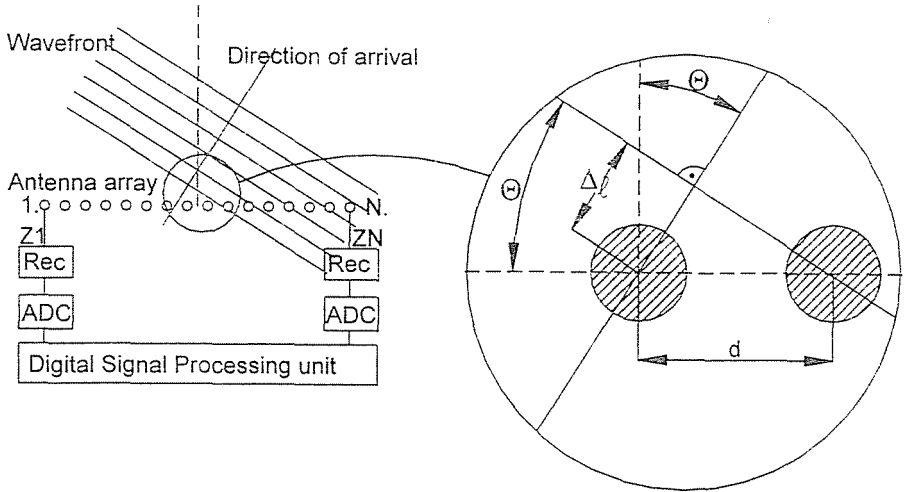
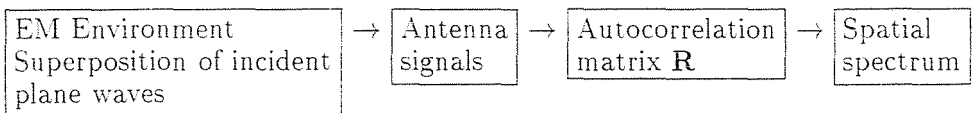


Fig. 1.

There is a clear evidence that radio direction finding raises basically the same questions as spectrum analysis. That is why the techniques discussed below are commonly referred to as adaptive spectral estimation methods. The dualism of time and spatial domain is summarized below.

Analogies	
Time domain	Spatial domain
time	distance, displacement
frequency	spatial frequency
correlation	spatial correlation
spectrum	spatial spectrum
frequency domain filter	spatial filter

The main point of radio direction finding is determining the direction of arrival (DOA) of the radiating sources. This can be done by determining the spatial spectrum of the input process, i.e. the incoming signal vector of the individual array elements. The power spectrum of any stochastic processes can be derived from its auto-correlation function, as they are Fourier transform pairs. Finally the correlation can be calculated by the convolution of the time domain signal vector coming from all antenna elements. The summary of this processing flow is shown here:



Similarly to any real measurement situation the available data covers

a finite time and spatial span rather than the infinite time period and/or spatial range required by definition of any correlation function. In this case the spatial domain is sampled at the antenna element positions, and the time domain is limited to discrete samples of a finite interval, too, which explains the word 'estimation' in both spectral and spatial sense.

After having drawn the rough skeleton of the successive steps in the data flow we give a more detailed description of each phase.

1.1. Mathematical Model [1] [3] [5] [6]

In the mathematical model we assume a linear antenna array of N elements, M different interfering sources and thermal noise. The antenna elements are equally spaced and the distance is not greater than half of the wavelength of the incident field (Shannon's sampling law). The antenna elements are isotropic or omnidirectional. The interfering sources are sinusoidal. The effects of non-zero bandwidth will be taken into consideration later in this paper. The thermal noise is Gaussian white noise with zero mean value, σ^2 variance and is uncorrelated with all the interfering sources. The signal of a single antenna element can be written in the following form:

$$z_k(t) = n_k(t) + \sum_{m=1}^M p_m(t) g_k(Q_m) e^{j\omega_m x_k}, \quad k = 1, \dots, N. \quad (1)$$

Where n_k is the thermal noise component; p_m is the power density of the m -th source at the array's position; g_k is the gain in the direction of the source - this value is actually independent of Θ , as we assumed omnidirectional or isotropic antenna elements. The exponential component needs further explanation: ω_m is the spatial frequency, and x_k is the distance of a given antenna element, measured in wavelength units from the end of the array ($\omega = 2\pi \sin \Theta$, $x_k = k \frac{d}{\lambda}$). One sample of all antenna elements at a given instant can be expressed as follows:

$$\begin{bmatrix} z_1(t) \\ z_2(t) \\ \vdots \\ z_N(t) \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1M} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ v_{N1} & v_{N2} & \cdots & v_{NM} \end{bmatrix} \cdot \begin{bmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_M(t) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \\ \vdots \\ n_N(t) \end{bmatrix}. \quad (2)$$

The \mathbf{z} is the input signal vector, each vector element corresponds to an antenna element. The \mathbf{V} matrix contains the gains multiplied by the phase difference of the single elements, thus this depends on both the antenna array configuration and the electromagnetic environment, each column corresponds to an interfering source, each row to an antenna element. The

incident power at the array is represented by the \mathbf{p} vector and \mathbf{n} is the noise vector.

Until now we set up the model for the electromagnetic environment. Now the first step of processing has to be done with this data. Determining the autocorrelation matrix of the incoming signals can be done as follows:

$$\mathbf{R} = E \left\{ \mathbf{z}(t, \xi) \mathbf{z}^H(t, \xi) \right\}, \quad (3)$$

where $E\{\}$ is the ensemble averaging operator which – for ergodical processes – is equal to the time average.

Using the \mathbf{V} matrix shown above and exploiting the fact that noise and interference are uncorrelated, this equation can be rewritten:

$$\mathbf{R} = \mathbf{V} E \left\{ \mathbf{p}(t, \xi) \mathbf{p}^H(t, \xi) \right\} \mathbf{V}^H + E \left\{ \mathbf{n}(t, \xi) \mathbf{n}^H(t, \xi) \right\}. \quad (4)$$

We introduce the following notation

$$\mathbf{P} = E \left\{ \mathbf{p}(t, \xi) \mathbf{p}^H(t, \xi) \right\}, \quad (5-a)$$

$$E \left\{ \mathbf{n}(t, \xi) \mathbf{n}^H(t, \xi) \right\} = \sigma^2 \mathbf{I}. \quad (5-b)$$

Now we can rewrite (4) in a shorter form, that will be used in this paper:

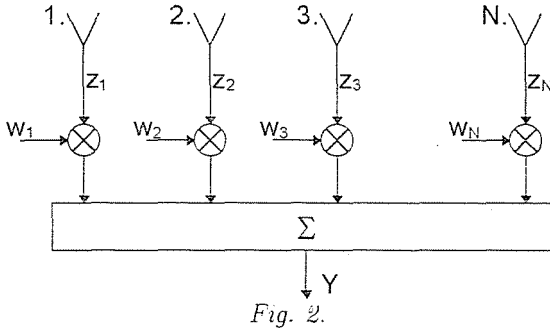
$$\mathbf{R} = E \left\{ \mathbf{z}(t, \xi) \mathbf{z}^H(t, \xi) \right\} = \mathbf{V} \mathbf{P} \mathbf{V}^H + \sigma^2 \mathbf{I}. \quad (6)$$

The mean value should be calculated on an infinite interval of time. In real systems this is impossible, only a certain number of samples can be used, thus the autocorrelation matrix can only be estimated. During practical tests it turned out that about 100 samples result in sufficient accuracy. There is, however, another limiting factor in real EM environments for the number of samples: the stationarity of the observed process is not always satisfied, as the autocorrelation function is a slowly varying function of the time. This means that the time interval required to take the samples for the measurement has to be within the time constant of the quasistationarity.

2. Principles of Different Spectrum Estimating Methods

2.1. Conventional Beamforming [1] [2] [3] [5]

Conventional beamforming estimates the spatial spectrum by making a Fourier transform on the incoming $\mathbf{z}(t)$ signal's correlation matrix after applying a triangle window. This is what we call Bartlett estimation. The 'triangle' is because the weighting coefficients are $|w_k| = 1$, $k = 1, \dots, N$



and therefore the incoming correlation series (estimating the correlation function) will be multiplied by a triangle window. In the spatial frequency domain our antenna array will have a 'sinc' pattern. The problem is that this pattern has a high sidelobe level. We can suppress these sidelobes by selecting another windowing function, but then the mainlobe's beamwidth increases and the angular resolution decreases. Note that low sidelobe level and narrow main lobe contradict each other.

If we vary the weighting coefficients' phase linearly, we get to the electronically scanned beam antenna. We sum the incoming signals from the individual elements (make the $Y = \mathbf{w}^T \mathbf{z}$ product) with a phase delay reflecting the current direction of arrival (DOA). This complex output voltage is a maximum, if the mainlobe and DOA match. Such a system can be seen on Fig. 2.

The $\mathbf{w}(\Theta)$ weighting (here: scanning) vector has the following form:

$$w_k = e^{-j\omega_m x_k}, \quad k = 1, \dots, N. \quad (7)$$

According to our analogy the scanning main lobe equals to a filter having a 'sinc' transmission function shifting in spatial frequency domain. If the beam points to Θ direction, the system's output power is:

$$P(\Theta) = \overline{|Y|^2}. \quad (8)$$

As we already know,

$$Y = \mathbf{w}^T(\Theta) \mathbf{z}, \quad (9)$$

so the power spectrum is the following:

$$P_{\text{CBF}}(\Theta) = \frac{1}{N^2} \mathbf{w}^H(\Theta) \mathbf{R} \mathbf{w}(\Theta), \quad (10)$$

where the $\{\}^H$ operator is the transponate-conjugate operator; \mathbf{R} is the incoming signal series' correlation matrix.

The Rayleigh limit is valid for the systems resolution, which means that two targets can be separated only if they have a distance which is equal to the distance of the first null from the main beam's peak. The angular resolution's value is the half of the main beam's width:

$$\beta_R = \frac{\lambda}{D}, \quad (11)$$

where D is the linear antenna array's length; λ is the incoming plane wave's wavelength. We can increase the resolution, if we increase the electrical length of the array. As this always results in increasing the N number of elements, we run into practical limits.

Advantages:

- + Low computational needs;
- + Area under the spectral curve corresponds to the power of the incoming signal.

Drawbacks:

- High sidelobe level, which means lower dynamic range;
- Flat main beam (hard to find DOA);

2.2. Maximum Signal-to-Noise-Interference Algorithm [1] [3] [7]

This method was introduced by CAPON in 1969. This adaptive algorithm could be described by a changing FIR filter in the spatial domain, which alters at every frequency to produce a maximum of the $\frac{\text{Signal}}{\text{Noise} + \text{Interference}}$ ratio. (That gives the name Maximum Signal-to-Noise-Interference Ratio - MSINR.)

We can also call this method 'Minimum Variance' as the aim is to hold the Y outgoing power originating from noise and interference at minimum level at every direction (spatial frequency), while the currently examined direction is put through with unity gain. This is like a fictive scanning main lobe antenna array, which has unity signal gain and is producing an antenna pattern which minimizes the effect of all other interference sources. If we scan in such a way through the entire angular interval, tuning our filter by the algorithm given above, the SINR will be the lowest at the places where the interference and the fictive source are matching. That is the way we can estimate the spectral distribution function from the incoming correlation matrix.

Let us see how we can get the power spectral density. As we know:

$$P(\Theta) = \mathbf{w}^H(\Theta) \mathbf{R} \mathbf{w}(\Theta). \quad (12)$$

We can compute the SNR of the fictive signal (\mathbf{s} vector) and the noise + interference sources (\mathbf{n} vector):

$$\frac{\text{Signal}}{\text{Noise} + \text{Interference}} = \frac{|\mathbf{w}^T \mathbf{s}|^2}{|\mathbf{w}^T \mathbf{z}|^2} = \frac{|\mathbf{s}^T \mathbf{w}|^2}{\mathbf{w}^H \mathbf{R} \mathbf{w}}. \quad (13)$$

We know that the optimum value for \mathbf{w} to achieve the highest possible SINR is:

$$\mathbf{w}(\Theta) = \mu \mathbf{R}^{-1} \mathbf{s}^*(\Theta), \quad (14)$$

where $\mathbf{s}(\Theta)$ is the weighting vector which would turn a classical beamformer's main lobe into Θ direction (that's why it is often called steering vector); μ is a complex constant – the main beam unity gain can be achieved by properly choosing this. The maximized SNR is:

$$\max \left\{ \frac{\text{Signal}}{\text{Noise} + \text{Interference}} \right\} = \frac{|\mu|^2 |\mathbf{s}^T \mathbf{R}^{-1} \mathbf{s}^*|^2}{(\mu^* \mathbf{s}^T \mathbf{R}^{-1}) \mathbf{R} (\mu \mathbf{R}^{-1} \mathbf{s}^*)} = \mathbf{s}^T(\Theta) \mathbf{R}^{-1} \mathbf{s}^*(\Theta), \quad (15)$$

$$\text{MSINR} = \mathbf{s}^T(\Theta) \mathbf{R}^{-1} \mathbf{s}^*(\Theta) = \mathbf{s}^H(\Theta) \mathbf{R}^{-1} \mathbf{s}(\Theta).$$

The gain is unity if $\mathbf{s}^T \mathbf{w} = 1$, so μ is:

$$\mu = [\mathbf{s}^T(\Theta) \mathbf{R}^{-1} \mathbf{s}^*(\Theta)]^{-1}. \quad (16)$$

We substitute the constant into (12), then we have the spatial spectrum, which gives the reciprocal value of the SINR (since we scaled μ for unity gain):

$$P_{\text{MSINR}}(\Theta) = \frac{1}{\mathbf{s}^H(\Theta) \mathbf{R}^{-1} \mathbf{s}(\Theta)}. \quad (17)$$

Advantages:

- + High angular resolution;
- + Wide dynamic range;
- + The peak's maximal value corresponds to the incoming power from Θ direction;
- + Low sidelobes.

Drawbacks:

- High computational performance required;
- Bandwidth-sensitive;
- Correlation-sensitive.

2.3. Maximum Entropy Method [1] [3] [4] [5]

This method is also called Howells-Applebaum (HA). The method is strongly coupled to the linear prediction algorithm, and – in case of a one-dimensional antenna array – the two methods give the same Power Spectral Density (PSD). The linear prediction method operates on a FIR filter's coefficients to minimize the error signal at the output. This can be done by removing all the deterministic components (increase the entropy) from the output signal so it looks like a Gaussian white noise – that is what we call whitening. The ADPCM coding mechanism used in speech encoding works just the same way – only the output (white noise-like) and the filter coefficients are transmitted.

The HA method's approach is to use only really measured data, but this data must be utilized fully – in contrast to the Fourier methods, where data is being lost (windowing function) and violated by using false data (estimating 0-s at the unknown places). The HA method wants to estimate the unknown points in the less determinant (maximal entropy) way, or with other words, to continue the function in the most probable way.

The MEM methods spectrum can be described by the following formula in a vectorial form (derivation omitted for shortness):

$$P_{\text{MEM}}(\Theta) = \frac{1}{|\mathbf{s}^H(\Theta)\mathbf{R}^{-1}\delta|^2 |\mathbf{s}(\Theta)|^2}, \quad (18)$$

where:

- \mathbf{R} is the autocorrelation matrix;
- δ is a steering vector, usually defined for a linear array in the following way: $\delta^T = [1 \ 0 \ \dots \ 0]$.

Advantages:

- + Provides higher angular resolution than MSINR;
- + Great dynamic range;
- + Low sidelobe ripple.

Drawbacks:

- High computational performance required;
- Bandwidth-sensitive;
- Correlation-sensitive.

3. Qualitative Analysis and Comparison of the Described Methods [5]

In this section the dynamic range and resolution of the previously described methods will be calculated in the simplest possible case of radio direction finding environment. The applied model is a special, simple case of the one described in section 1.1:

- linear antenna array consisting of N isotropic elements spaced $\lambda/2$ from each other
- one signal source with power P_{sig} at Θ_0 azimuth angle (distant enough to apply the plane wave model) and Gaussian white noise with 0 mean value and σ_0 .

In order to calculate the minimal necessary signal-to-noise ratio we exploit the fact that the minimal dynamic range must be at least 3 dB in order to fulfil the Rayleigh resolution limit.

The autocorrelation matrix is as follows under the above conditions:

$$\mathbf{R} = \sigma_0^2 \mathbf{I} + P_{\text{sig}} \mathbf{s}(\Theta_0) \mathbf{s}^H(\Theta_0), \quad (19)$$

where

$$\mathbf{s}^T(\Theta) = \left[1 \quad e^{-j2\pi f_d \frac{\lambda}{2}} \quad \dots \quad e^{-j2\pi f_d \frac{\lambda}{2}(N-1)} \right]; \quad f_d = \frac{1}{\lambda_0} \sin \Theta. \quad (20)$$

3.1. The Conventional Method (Bartlett Estimation)

The power spectrum of the Bartlett estimation is

$$P_{\text{CBF}}(\Theta) = \frac{1}{N^2} \mathbf{w}^H(\Theta) \mathbf{R} \mathbf{w}(\Theta). \quad (21)$$

Substituting the current autocorrelation matrix (19) into this equation we obtain

$$\begin{aligned} P_{\text{CBF}}(\Theta) &= \frac{1}{N^2} \left[\sigma_0^2 \mathbf{w}^H(\Theta) \mathbf{I} \mathbf{w}(\Theta) + P_{\text{sig}} \mathbf{w}^H(\Theta) \mathbf{s}(\Theta_0) \mathbf{s}^H(\Theta_0) \mathbf{w}(\Theta) \right] = \\ &= \frac{1}{N^2} \left[\sigma_0^2 N + P_{\text{sig}} \mathbf{w}^H(\Theta) \mathbf{s}(\Theta_0) \mathbf{s}^H(\Theta_0) \mathbf{w}(\Theta) \right], \end{aligned} \quad (22)$$

this expression has its maximum at $\Theta = \Theta_0$:

$$P_{\text{CBF}}(\Theta_0) = \frac{1}{N^2} \left[\sigma_0^2 N + P_{\text{sig}} N^2 \right]. \quad (23)$$

The estimated mean power in the direction of arrival of the signal:

$$P_{\text{CBF}}(\Theta_0) = \frac{\sigma_0^2}{N^2} + P_{\text{sig}}. \quad (24)$$

At angles different from Θ_0 the magnitude of $P_{\text{sig}} = \mathbf{w}^H(\Theta) \mathbf{s}(\Theta) \mathbf{s}^H(\Theta) \mathbf{w}(\Theta)$ rapidly decreases, whereas the noise power $\frac{\sigma_0^2}{N}$ remains constant. The minimal estimated value is $\frac{\sigma_0^2}{N}$. Now the minimal SNR necessary to obtain a dynamic range of the required 3 dB will be determined.

$$\Delta_{\text{CBF}}^{3\text{dB}} = \frac{P_B(\Theta_0)}{P_{\text{CBFmin}}(\Theta)} = \frac{\frac{\sigma_0^2}{N} + P_{\text{sig}}}{\frac{\sigma_0^2}{N}}, \quad (25)$$

$$\left(\frac{P_{\text{sig}}}{\sigma_0^2}\right)_{\min} = \frac{1}{N}; \quad \left(\frac{P_{\text{sig}}}{\sigma_0^2}\right)_{\min}^{\text{dB}} = -10 \lg N. \quad (26)$$

While determining the dynamic range of the conventional or Bartlett estimation the sinc-like shape of the Bartlett window's Fourier transform has to be taken into consideration. We chose the level of the first sidelobe's peak of the *sinc* envelope as the lower reference of the estimation's dynamic range. Its relative level is:

$$\text{sinc}\left(\frac{3\pi}{2}\right) = -\frac{2}{3\pi}, \quad (27)$$

thus the dynamic range:

$$\Delta_{\text{CBF}} = \frac{P_{\text{CBF}}(\Theta_0)}{P_{\text{CBF}}(\Theta_1)} = \frac{\frac{\sigma_0^2}{N} + P_{\text{sig}}}{\frac{\sigma_0^2}{N} + \left(\frac{2}{3\pi}\right)^2 P_{\text{sig}}} = \frac{\frac{1}{N} + \frac{P_{\text{sig}}}{\sigma_0^2}}{\frac{1}{N} + \left(\frac{2}{3\pi}\right)^2 \frac{P_{\text{sig}}}{\sigma_0^2}},$$

$$\Delta_{\text{CBF}}^{\text{dB}} = 10 \lg \left(\frac{\frac{1}{N} + \frac{P_{\text{sig}}}{\sigma_0^2}}{\frac{1}{N} + \left(\frac{2}{3\pi}\right)^2 \frac{P_{\text{sig}}}{\sigma_0^2}} \right). \quad (28)$$

3.2. The MSINR (Capon) Method

The MSINR power spectrum was derived in the 2nd section (17)

$$P_{\text{MSINR}}(\Theta) = \frac{1}{\mathbf{s}^H(\Theta)\mathbf{R}^{-1}\mathbf{s}(\Theta)}, \quad (29)$$

where \mathbf{R} is the current autocorrelation matrix (19), now its inverse is needed:

$$\mathbf{R}^{-1} = \frac{1}{\sigma_0^2} \left[\mathbf{I} - P_{\text{sig}} \frac{\mathbf{s}(\Theta_0)\mathbf{s}^H(\Theta_0)}{\sigma_0^2 + NP_{\text{sig}}} \right]. \quad (30)$$

By substituting \mathbf{R}^{-1} into the denominator of the MSINR spectrum (30)

$$\begin{aligned} \frac{1}{P_{\text{MSINR}}(\Theta)} &= \mathbf{w}^H(\Theta)\mathbf{R}^{-1}\mathbf{w}(\Theta) = \\ &= \frac{1}{\sigma_0^2} \left[\mathbf{w}^H(\Theta)\mathbf{I}\mathbf{w}(\Theta) - P_{\text{sig}} \frac{\mathbf{w}^H(\Theta)\mathbf{s}(\Theta_0)\mathbf{s}^H(\Theta_0)\mathbf{w}(\Theta)}{\sigma_0^2 + NP_{\text{sig}}} \right] = \\ &= \frac{1}{\sigma_0^2} \left[N - P_{\text{sig}} \frac{\mathbf{w}^H(\Theta)\mathbf{s}(\Theta_0)\mathbf{s}^H(\Theta_0)\mathbf{w}(\Theta)}{\sigma_0^2 + NP_{\text{sig}}} \right]. \end{aligned} \quad (31)$$

At $\Theta = \Theta_0$ angle, thus in the direction of arrival the power density is:

$$\frac{1}{P_{\text{MSINR}}(\Theta)} = \frac{1}{\sigma_0^2} \left[N + P_{\text{sig}} \frac{N^2}{\sigma_0^2 + NP_{\text{sig}}} \right], \quad (32)$$

so the MSINR power estimation in the signal source's direction:

$$P_{\text{MSINR}}(\Theta_0) = \frac{\sigma_0^2}{N} + P_{\text{sig}}. \quad (33)$$

At azimuth angles different from (Θ_0) the value of $\mathbf{w}^H(\Theta)\mathbf{s}(\Theta_0)\mathbf{s}^H(\Theta_0)\mathbf{w}(\Theta)$ rapidly decreases, whereas the $\frac{\sigma_0^2}{N}$ noise power remains constant. Thus the estimated minimum value is $\frac{\sigma_0^2}{N}$. The dynamic range of the estimation can now be derived:

$$\begin{aligned} \Delta_{\text{MSINR}} &= \frac{P_{\text{MSINR}}(\Theta_0)}{P_{\text{MSINRmin}}(\Theta)} = \frac{\frac{\sigma_0^2}{N} + P_{\text{sig}}}{\frac{\sigma_0^2}{N}} = 1 + N \frac{P_{\text{sig}}}{\sigma_0^2}, \\ \Delta_{\text{MSINR}}^{\text{dB}} &= 10 \lg \left(1 + N \frac{P_{\text{sig}}}{\sigma_0^2} \right). \end{aligned} \quad (34)$$

The lowest possible signal-to-noise ratio required to achieve the required minimal dynamic range of 3 dB can be calculated.

$$\Delta_{\text{MSINR}}^{3\text{dB}} = 2 = \frac{P_{\text{MSINR}}(\Theta_0)}{P_{\text{MSINRmin}}(\Theta)} = \frac{\frac{\sigma_0^2}{N} + P_{\text{sig}}}{\frac{\sigma_0^2}{N}}, \quad (35)$$

$$\left(\frac{P_{\text{sig}}}{\sigma_0^2}\right)_{\text{min}} = \frac{1}{N}; \quad \left(\frac{P_{\text{sig}}}{\sigma_0^2}\right)_{\text{min}}^{\text{dB}} = -10 \lg N. \quad (36)$$

3.3. The MEM Method

We use the MEM estimation of the power spectrum described in 2.3

$$P_{\text{MEM}}(\Theta) = \frac{1}{N^2} \frac{1}{|\mathbf{w}^T(\Theta)\mathbf{R}^{-1}\delta|^2}, \quad (37)$$

where \mathbf{R} equals (19) and $\delta^T = [1 \ 0 \ \dots \ 0]$.

Substituting the same \mathbf{R}^{-1} matrix as in MSINR case (30) into the power spectrum estimation (37) we obtain

$$\mathbf{w}^T(\Theta)\mathbf{R}^{-1}\delta = \frac{1}{\sigma_0^2} \left[\mathbf{w}^T(\Theta)\mathbf{I}\delta - P_{\text{sig}} \frac{\mathbf{w}^T(\Theta)\mathbf{s}(\Theta_0)\mathbf{s}^H(\Theta_0)\delta}{\sigma_0^2 + NP_{\text{sig}}} \right], \quad (38)$$

which gives

$$\mathbf{s}^T(\Theta_0)\mathbf{R}^{-1}\delta = \frac{1}{\sigma_0^2} \left[N - P_{\text{sig}} \frac{N}{\sigma_0^2 + NP_{\text{sig}}} \right] = \frac{1}{\sigma_0^2 + NP_{\text{sig}}}, \quad (39)$$

at $\Theta = \Theta_0$, thus in the direction of the incident wave, which results in the MEM power estimation in the DOA in:

$$P_{\text{MEM}}(\Theta_0) = \frac{1}{N^2} \frac{1}{|\mathbf{s}^T(\Theta_0)\mathbf{R}^{-1}\delta|^2} = \frac{|\sigma_0^2 + NP_{\text{sig}}|}{N^2} = \left| \frac{\sigma_0^2}{N} + P_{\text{sig}} \right|^2. \quad (40)$$

At angles different from (Θ_0) $\mathbf{w}^T(\Theta)\mathbf{s}(\Theta_0)\mathbf{s}^H(\Theta_0)\delta$ tends to zero rapidly again, whereas the value of $\left| \frac{\sigma_0^2}{N} \right|^2$ related to the noise power remains constant.

Thus the estimated minimum of $P_{\text{MEM}}(\Theta)$ is $\left| \frac{\sigma_0^2}{N} \right|^2$.

Now the dynamic range of the estimation can be derived:

$$\Delta_{MEM} = \frac{S_{MEM}(\Theta_0)}{S_{MEM}(\Theta)} = \frac{\left[\frac{\sigma_0^2}{N} + P_{sig}\right]^2}{\left[\frac{\sigma_0^2}{N}\right]^2} = \left[1 + N \frac{P_{sig}}{\sigma_0^2}\right]^2,$$

$$\Delta_{MEM}^{dB} = 20 \lg \left[1 + N \frac{P_{sig}}{\sigma_0^2}\right]. \quad (41)$$

The lowest acceptable signal-to-noise ratio required to achieve the required minimal dynamic range of 3 dB can be calculated similarly to the previous cases.

$$\Delta_{MEM}^{3dB} = 2 = \frac{P_{MEM}(\Theta_0)}{P_{MEM}(\Theta)} = \left(\frac{\frac{\sigma_0^2}{N} + P_{sig}}{\frac{\sigma_0^2}{N}}\right)^2,$$

$$\left(\frac{P_{sig}}{\sigma_0^2}\right)_{MEMmin} = \frac{\sqrt{2} - 1}{N}; \quad \left(\frac{P_{sig}}{\sigma_0^2}\right)_{MEMmin}^{dB} = 10 \lg \frac{\sqrt{2} - 1}{N}. \quad (42)$$

3.4. Summary and Comparison

Table 1 summarizes the required minimal signal-to-noise ratio and the dynamic range as a function of the SNR and the number of antenna elements for the three discussed methods.

Table 1.

	CBF	MSINR	MEM
$\left(\frac{P_{sig}}{\sigma_0^2}\right)_{min}$	$-10 \lg N$	$-10 \lg N$	$10 \lg \frac{\sqrt{2} - 1}{N}$
Δ_{3dB}	$10 \lg \left(\frac{\frac{1}{N} + \frac{P_{sig}}{\sigma_0^2}}{\frac{1}{N} + \left(\frac{2}{3\pi}\right)^2 \frac{P_{sig}}{\sigma_0^2}}\right)$	$10 \lg \left(1 + N \frac{P_{sig}}{\sigma_0^2}\right)$	$20 \lg \left(1 + N \frac{P_{sig}}{\sigma_0^2}\right)$

Fig. 3 shows the dynamic range as a function of the SNR in the case of a 10 element antenna array, Fig. 4 shows the angular resolution versus the SNR with 2 elements. The relative angular resolution is defined as follows:

$$\frac{\Theta_{3dB}}{\Theta_{CBF_3dB}},$$

where the numerator is the 3 dB angular resolution of the given method, and the denominator is the angular resolution of the CBF method at infinite SNR.

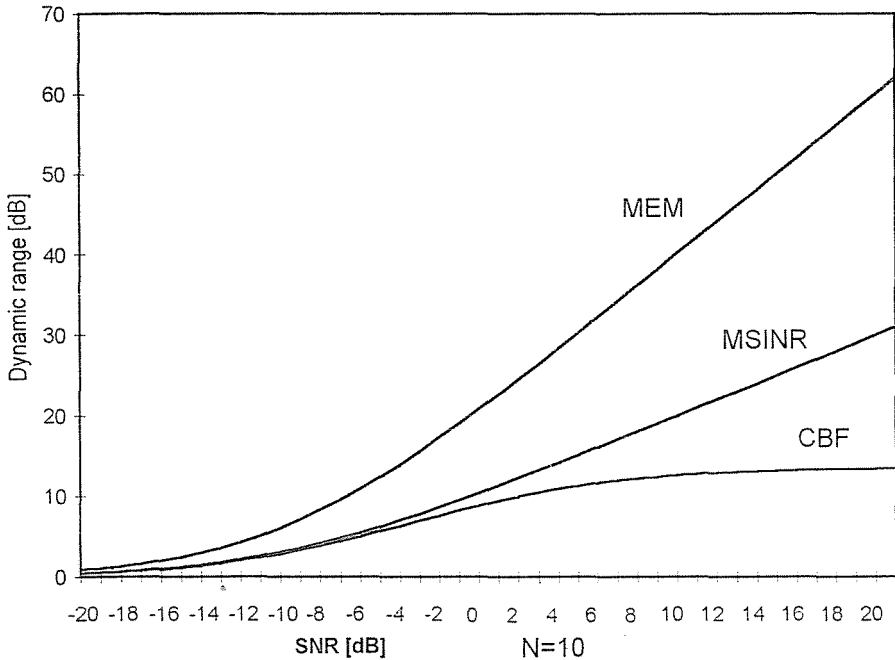


Fig. 3.

The following characteristics can be observed on *Fig. 3* and *Fig. 4*:

- the CBF and the MSINR methods have the same minimal SNR, whereas the MEM has significantly lower SNR requirement for the 3 dB dynamic range,
- the resolution of the CBF does not increase with the input SNR above a certain level, whereas the MEM and MSINR methods are linear functions of the SNR at higher values, with MEM having twice the slope of MSINR,
- the strong relationship of resolution of dynamic range.

4. Performance Reducing Effects

4.1. Bandwidth

In the mathematical model we assumed an unmodulated carrier. In practical applications this assumption is never met. Transmission of information requires a certain bandwidth, that is characteristic for the data transmitted

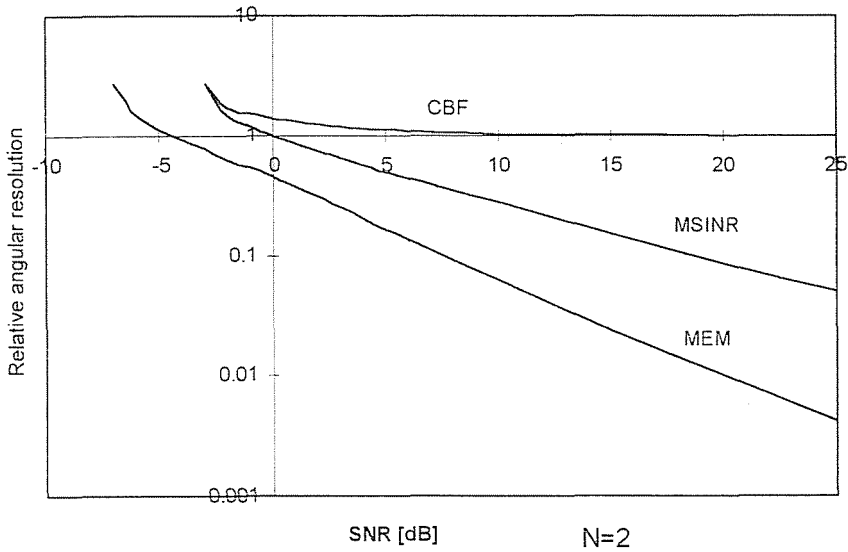


Fig. 4.

and the modulation process. In many cases, however, the relative bandwidth is fairly small, many telecommunication applications occupy a narrow band around the carrier. On the field of mobile communications for example the relative bandwidth of the NMT 450 system is roughly $25 \text{ kHz}/450 \text{ MHz} = 5.5 \cdot 10^{-5}$, in radar applications $2 \text{ MHz}/1.5 \text{ GHz} = 1.33 \cdot 10^{-3}$, for CB and military short wave radios it is $12.5 \text{ kHz}/30 \text{ MHz} = 4.1 \cdot 10^{-4}$. The effect of narrowband signals will be shown in the part describing computer simulation results.

4.2. Correlation

Correlation is one of the weaknesses of all adaptive methods described in this paper. Correlated signals occur very often in free space propagation environment as a consequence of multipath effects. The reflected and the direct wave have a correlation coefficient close to unity if the modulation bandwidth is small compared with the reciprocal value of the time delay caused by the reflection. Unfortunately, this practically important phenomenon has quite a dramatic effect on the performance of the adaptive methods to be described. A representative simulation result will also be shown to demonstrate the performance reduction. To overcome this problem correlation destruction methods can be applied, but this subject exceeds the coverage of this paper.

4.3. Number of Antenna Elements

The most important restriction for the number of elements in the array is that it must exceed the number of interference sources at least by one. Otherwise the system has not enough degree of freedom to be able to determine the spatial spectrum. Reflected signals of multipath propagation cause the number of interfering sources to increase, furthermore the effect mentioned above has an unpleasant consequence. Even in the case of relatively few interfering signals the increase of the antenna element number will culminate in better angular resolution as the correlation matrix will be greater and better conditioned. Based on better set of data, the estimation process will do a better job as well.

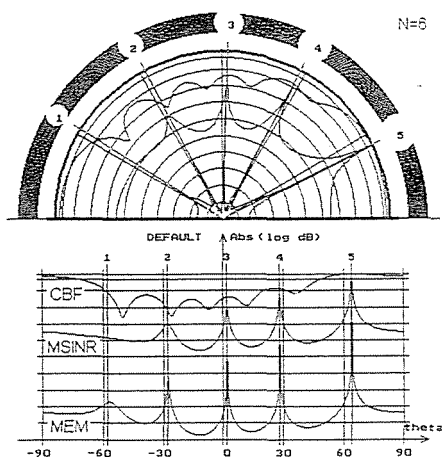


Fig. 5.

5. Computer Simulation Results

After the theory let us see some results which demonstrate the better performance of the adaptive algorithms. Fig. 5 shows 5 sources with increasing amplitude, and the result of the three (Direct, MSINR and MEM) algorithms trying to find all the sources.

The two adaptive algorithms were able to find the sources with a wide dynamic range, while the direct method failed. The first source's SNR was below the minimal limit of the MSINR method. The figure clearly shows that the MEM method has a very wide dynamic range, and is able to find even the smallest source.

Fig. 6 shows the case of closely spaced sources with increasing angular distance.

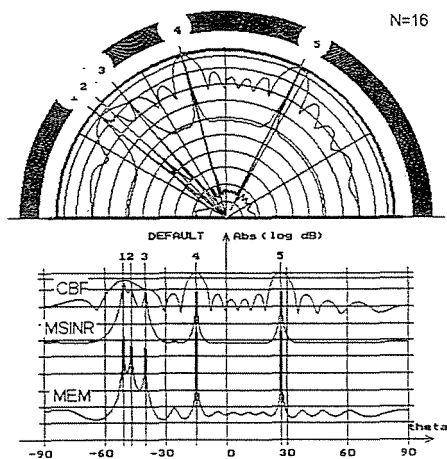


Fig. 6.

The result again is that the adaptive methods can achieve higher peaks and therefore separate the sources. The best is again the MEM method, but the MSINR is also supplying useful data.

Now let us see the drawbacks of the two adaptive methods. The first performance limiting effect was the bandwidth. *Fig. 7* shows the effect of the increasing relative channel bandwidth in two steps. The direct method is practically insensitive to the bandwidth, while the other methods suffer a decrease in performance. It can, however, be stated that the adaptive methods are not subject to such performance degradation even in this case, which would result in a resolution poorness comparable to the conventional method.

The next reducing effect was the correlation. This is one of the most important factors as will see from *Fig. 8*, where the correlation is switched off and totally on between the two sources. The picture shows the greatest drawback of the MEM method: the correlation sensitivity. If the sources are correlated, both methods produce weak results. Therefore, if using adaptive methods, we must usually use decorrelation or correlation-destroying algorithms like Spatial Smoothing Process (SSP) or Modified SSP. These algorithms are able to decorrelate the correlation matrix, but have the effect of decreased resolution, as they use some samples for the decorrelation.

The last problem is the effect of the Gaussian noise present in the environment. *Fig. 9* shows one case, where the noise level was set to -50 dB and then to 0 dB. The adaptive methods show again significant decrease in performance at $\text{SNR} = 0$ dB, the MEM shows even increasing sidelobes (note that these sidelobes are still lower than those of the direct method).

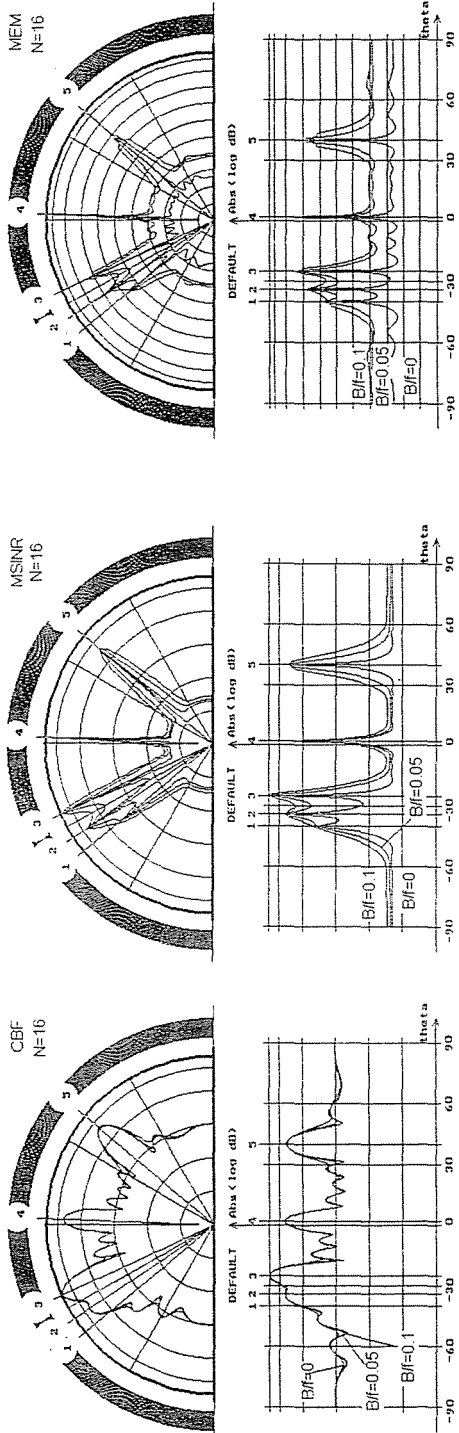


Fig. 7.

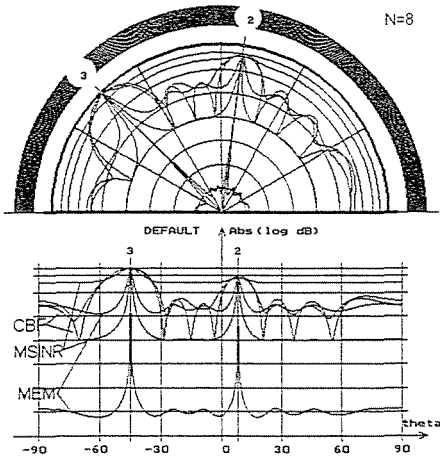


Fig. 8.

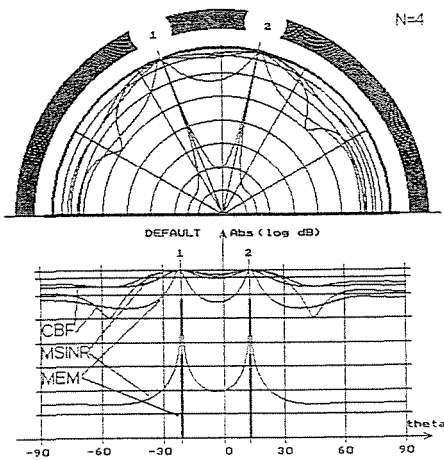


Fig. 9.

6. Concluding Remarks

The elaborated program gave us a powerful tool for further experimentation on this field, by creating various simulated electromagnetic situations and compare the results with our expectations. Similarly, it can be used in the education of subjects concerning wave propagation, antennas and radio measuring systems. We also got an idea of the computational performance and numerical stability required during our experience with the program.

The coverage of the program could be extended by implementing another very promising method, the so-called MUSIC algorithm. According

to the papers studied this method gives the highest accuracy among all mentioned here.

Practical experience can only be gathered by using hardware implementation. As the basic theory described in the first section cannot only be applied on radio frequency and electromagnetic waves, the easiest and cheapest test configuration can be constructed at the ultrasonic frequency region using acoustic waves. With acoustic models, however, one has to be very careful, as this range is heavily loaded with noise of different origin. A possible radio frequency application is surveying channel usage within a single cell of NMT 450 or GSM network, in order to determine optimal base station placement.

References

- [1] FARINA, A. (1992): Antenna-Based Signal Processing Techniques for Radar Systems, Artech House.
- [2] STEYSKAL, H. – ROSE, J. F. (1989): Digital Beamforming for Radar Systems, *Microwave Journal*, 01/1989.
- [3] DEATS, B. W. – FARINA, D. J. – BULL, J. F. (1991): Super-Resolution Signal Processing Aids RCS Testing, *Microwave & RF*, 03/1991.
- [4] ABLES, J. G. (1992): Maximum Entropy Spectral Analysis, *Modern Spectrum Analysis, IEEE*.
- [5] SELLER, R. (1996): Módszerek céltárgyparaméterek rádiólokációs mérési pontosságának növelésére; Műszaki doktori értekezés, BME VIK.
- [6] GABRIEL, W. F. (1980): Spectral Analysis and Adaptive Array Superresolution Techniques, *Proceedings of the IEEE*, Vol. 68, No. 6, June.
- [7] CAPON, J. (1978): High-Resolution Frequency-Wavenumber Spectrum Analysis, *Modern Spectrum Analysis, IEEE PRESS*.