

CONTEMPORARY METHODS OF DATA PROCESSING IN EXPERIMENTAL PHYSICS

*G. Ososkov*¹

Joint Institute for Nuclear Research, Dubna

Three basic methods that are extensively applied at JINR to process recent experimental data are reviewed, namely, robust methods of mathematical statistics, artificial neural networks and wavelet analysis. This review primarily covers studies in which scientists from the Laboratory of Information Technologies participated, in particular, in collaborations with the leading centers of physics, such as CERN, DESY, BNL, GSI, etc. The main principles of the reviewed methods and the most useful and promising examples of their applications are discussed.

Представлен обзор трех основных методов, широко используемых в ОИЯИ для обработки экспериментальных данных, а именно: робастных методов математической статистики, искусственных нейронных сетей и вейвлет-анализа. Этот обзор охватывает прежде всего исследования, в которых принимают участие сотрудники Лаборатории информационных технологий, особенно в сотрудничестве с ведущими физическими центрами, такими как ЦЕРН, DESY, BNL, GSI и др. Обсуждаются основные принципы рассматриваемых методов и наиболее плодотворные и многообещающие примеры их применения.

PACS: 02.60.-X; 02.60.Gf; 02.90.+P

INTRODUCTION

Advances in theoretical physics in the last decades gave rise to revolution in experimental high energy physics (HEP). Now HEP began to use colliders operating in TeV-energy range, such as RHIC at BNL, and LHC, which is to be commissioned at CERN. Therefore, the facilities and the detection technique drastically changed due to recent achievements in the electronics, distributed computing systems, and object-oriented programming and the explosive expansion of Internet. The Joint Institute for Nuclear Research participates in a number of large HEP experiments, being carried out and forthcoming, in leading world nuclear research centers. Advanced electronic detectors — proportional, time projection, and silicon chambers, drift tubes, RICH detectors of Cherenkov radiation rings, and various calorimeters are in use in these experiments.

The particular features of the data from these detectors are as follows: recognized patterns are discrete and have complex texture, data arrive with extremely high rate; the number of background events, which are similar to good events, is larger than the number of the latter

¹E-mail: ososkov@jinr.ru

events by several orders of magnitude; noise counts are numerous and correlated; very high multiplicity of objects (tracks, Cherenkov radiation rings, showers) to be recognized in each event.

The basic requirements on data processing in current experiments are maximum speed of computing in combination with the highest attainable accuracy and high efficiency of methods of estimating physical parameters interesting for experimentalists. The implementation of these requirements under the above conditions demanded a new HEP data analysis where we are faced not with usual applied statistics, but with special *mass-production statistics*. Keywords are:

- data processing in *automatic mode* with minimum of human intervention;
- need in *very fast algorithms of*
 - parameter estimating and
 - hypothesis testing;
- data *contamination* due to
 - noise measurements;
 - measurements from neighboring objects.

In order to achieve that, new mathematical means are needed, in particular:

- *robust approach* based fast algorithms;
- *neural networks*;
- *wavelet analysis* for data filtering and jet search.

In this paper the main principles of the reviewed methods and the most useful and promising examples of their applications are expounded. It is based on the earlier survey [1] and more recent publications [2–7].

1. ROBUST ESTIMATES FOR HEAVY CONTAMINATED SAMPLES

As an example of the robust approach the *elastic template method* (ETM) can be taken. It has been successfully applied for CERES RICH ring recognition [8] and STAR TPC tracking in 1998 [9]. Its idea was: knowing a parametric form of curves to be recognized, accomplish both stages, *a curve recognition and fitting*, simultaneously. Gyulassy and Harlander (G&H) realized this idea (see refs. in [9]) by varying track equation parameters so that being bent with their variations, a template, i.e., a curve described by this equation, passed as close as possible to the points measured in a track. G&H proposed as a physical interpretation of their idea the interaction between

- the positively charged template distributed with the density $\rho_T(r)$,
- negatively charged spatial points measured in the track having charge density $\rho(r')$.

The better the elastic template fits points, the lower the energy of their interaction. The energy E of interaction between these two charges is

$$E = - \int dr' dr \rho_T(r) V(r - r') \rho(r') \rightarrow \min .$$

Here V is the *Lorentz potential* $V(x, t) = \frac{w^2(t)}{x^2 + w^2(t)}$ with a temperature-dependent width $w(t) = b - (a - b)\exp(-t/T)$, where T is a constant, a is the maximal distance, at which points are still accredited to this template, and $b \sim \sigma_{\text{res}}$ is spatial resolution of a detector, $b \ll a$. Taking measurement discreteness into account, we obtain the track interaction energy as the sum:

$$E(\pi, t) = -\frac{1}{N} \sum_i^N \frac{w^2(t)}{\varepsilon^2(\pi, x_i) + w^2(t)}.$$

Here, N is the number of points in the track, $\varepsilon(\pi, x_i)$ is the distance of the i th point measured in the space from the template and π is the set of the track parameters. The functional $E(\pi, t)$ depends on the points of only one track (although simultaneous fitting of all tracks is also possible, but not recommended). To avoid $E(\pi, t)$ local spurious minima the *simulated annealing procedure* is applied. On the first iteration $w(t)$ is taken for the highest temperature, when $E(\pi, t)$ has the only one minimum. Then $w(t)$ is narrowed gradually allowing more and more accurate search of the global minimum.

As was shown in [2], the elastic tracking algorithm is a *particular case of the general robust approach*: data contamination means the violation of the Least Squares (LS) crucial assumption of residual normality and, therefore, one has to consider gross-error model of a contaminated distribution of measurement errors as

$$f(\varepsilon) = (1 - c)\phi(\varepsilon) + c h(\varepsilon), \quad (1)$$

where c is a parameter of contamination, $\phi(\varepsilon)$ is the Gauss distribution $N(0, \sigma)$ and $h(\varepsilon)$ is some long-tailed noise distribution density. Following P. Huber's M -estimate theory, we replace the LS functional $\sum_i \varepsilon_i^2$ by a functional

$$L(p, \sigma) = \sum_i \rho(\varepsilon_i), \quad (2)$$

where ε_i are residuals and $\rho(\varepsilon)$ is a contribution function of Lorentz or other similar type. It's easy to show that for minimizing of the functional (2) by parameters, its derivative

$$\frac{\partial L(p)}{\partial p} = \sum_{i=1}^n \frac{\partial \rho(\varepsilon_i)}{\partial \varepsilon_i} \frac{\partial \varepsilon_i}{\partial p} = 0$$

by denoting $w(\varepsilon) = \frac{1}{\varepsilon} \frac{\partial \rho(\varepsilon)}{\partial \varepsilon}$ can be modified to the form

$$\frac{\partial L(p)}{\partial p} = \sum_{i=1}^n w(\varepsilon_i) \frac{\partial \varepsilon_i}{\partial p} \varepsilon_i = 0,$$

which is similar to the normal LS equations, but with replacement of the numerical weight coefficients to the *weight function* $w(\varepsilon)$ to be recalculated on each step of an iterative procedure.

There are at least two difficulties in (2) minimizing and choosing the weight function $w(\varepsilon)$: (i) lack of theoretical foundation for the use of; (ii) computational problems due to unknown initial values of parameters to start iterations and (iii) non-uniqueness of the solution.

Our novel approach for the first problem was proposed for the particular, but often case of the uniform contamination, when $h(\varepsilon) = h_0 = \text{const}$. It gives the optimal weight function

$$w_{\text{opt}}(\varepsilon) = \frac{1 + c_{\text{opt}}}{1 + c_{\text{opt}} \exp(\varepsilon^2/2)}.$$

The only parameter c_{opt} is determined by the contamination of data not in the whole range of the sample but within its essential part where all useful observations are concentrated.

A polynomial expansion of $w_{\text{opt}}(\varepsilon)$ up to the fourth order leads to the approximation

$$w_T(\varepsilon) = \begin{cases} (1 - \varepsilon/(c_T\sigma)^2)^2, & \text{if } |\varepsilon| < c_T\sigma, \\ 0, & \text{otherwise.} \end{cases}$$

It is the famous Tukey's bi-weights, which are easier to calculate than optimal ones.

The second problem of obtaining parameter initial values to start iteration procedure is usually solved by applying to measurements the Hough transform (HT) (see, for instance, [8]). In case of circles the coordinates of the parameter space are the ring centers and their radii. Through three arbitrary signal points a unique circle can be drawn. Histogramming of all obtained ring centers gives peaks for each circle, but due to high combinatorics direct applying of HT procedure is very time consuming. A quite fast improvement of HT was proposed [6] for circle recognizing. It is based on preliminary coarse histogramming of center data and then clustering non-zero bins to obtain areas of interest where more detailed HT search is fulfilled. The next 1D HT radius search is easy to realize.

Computational problems arise because the function $L(p, \sigma)$ with bounded $\rho(\varepsilon)$ often has several minima. Such a nonuniqueness of $L(p, \sigma) \rightarrow \min$ can be overcome by the joint estimate of location and scale parameters [2]. To clarify that let us consider $L(p, \sigma)$ as the function of two parameters. On the corresponding surface in Fig. 1 one can see «ravines», whose bottoms are formed by curved lines. Let us denote by N_1 the line going along the deepest ravine.

An obvious and easy computational algorithm can be proposed from this consideration: since σ is much less than the range of the sample, move gradually along N_1 until reach a small threshold $\sigma = \sigma_{\text{min}}$. One can take σ_0 equal to the range of the sample, $\sigma_{\text{min}} = 0.001$, $\sigma^{(k+1)} = 0.95 \sigma^{(k)}$.

Does anyone see the full similarity with the simulated annealing procedure?

Numerous revealing examples of the robust approach applications are given in [1, 2], including solution of left-right ambiguity problems for drift chamber tracking, Cherenkov ring recognition on the quite noisy background, solution of calibration and alignment problems for various detectors and some others.

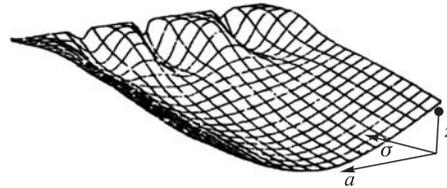


Fig. 1. Example of the surface $z = L(p, \sigma)$ for the sample containing 20 points grouped into three clusters. On these surfaces one can see «ravines», the number of which increases as $\sigma \rightarrow 0$

2. ARTIFICIAL NEURAL NETWORKS (ANN)

Let us remind briefly the basic concepts of the ANNs. Artificial neurons are simple logical devices specified by (i) activation level; (ii) topology of connections between neurons; (iii) the measure of interaction with other neurons, which is referred to as synaptic coupling power or weight; (iv) output level, which is related to the activation level by a certain, usually sigmoid, function. The weights of these connections are different and can be defined in dependence on the problem under consideration. The entire system consists of a vast number of identical neurons and the result of the operation of an ANN is almost not sensitive to the characteristics of a specific neuron. The system provides the possibility of the parallel processing of information.

The general input signal arriving at the j th neuron is

$$h_j = \sum_k w_{jk} x_k, \quad (3)$$

where x_k is the signal from the k th neuron of the network and w_{jk} is the synaptic-connection weight. The output signal of the j th neuron is the result of applying the activation function to this total signal, i.e., $y_j = g(h_j)$, where $g(u)$ is either a threshold function or a contraction sigmoid function like $g(u) = 1/(1 + \exp(-\lambda u))$.

The key characteristics of a network are the type of connections between neurons and network evolution dynamics determined by the activation function for neurons and the rule of varying weights upon this evolution.

ANNs that are extensively used in physics are determined by connections of two types: *feed-forward networks without feedback*, e.g., multilayer perceptrons (MLPs) or *recurrent networks*, where neurons are all connected with each other as in the Hopfield neural network. One can also consider *cellular automata* [1] as a special type of ANN with local connections.

Neural networks are successfully applied to the problems of classification, forecasting, and recognition. The simplicity of MLP structure stimulates many researchers to develop universal software packages that generate MLP on the basis of a specified number of layers and neurons in them and realize one of the network training methods. It is remarkable that one of the first such neural packages, JETNET, was developed by physicists at Lund University. For the same reasons, already in the early 1990s, several firms developing electronics succeeded in the hardware realization of wide-applicable ANNs in the form of integrated chips, which operate in parallel and allow the training of a network to an a priori simulated configuration (see references in [1]). Numerous papers on the application of various neural chips can be found in the special NIM issue A 389, 1997.

Concluding this section, we point out that MLP applications, most numerous in HEP, are, as a rule, based on a traditional scheme: first, ANN training by the backpropagation algorithm on a great training sample of simulated events, then embodying of the obtained ANN structure and weights in software or, more often, hardware implementation.

A substantial example of joint applications of robust methods, cellular automata and neural networks for data analysis in the OPERA experiment is given in [7].

2.1. Applications of Recurrent ANNs. As mentioned above, the general scheme of ANNs is valid also for the Hopfield neural network, which is a fully interconnected network reducing to the system of simple binary neurons. The evolution of the Hopfield neural network leads to a certain stable equilibrium state. Treating an ANN as a dynamical system, Hopfield used

the bilinear Lyapunov function as the network energy functional

$$E(s) = -1/2 \sum_{ij} s_i w_{ij} s_j \quad (4)$$

and demonstrated that, for the symmetric weight matrix $w_{ij} = w_{ji}$ with zero diagonal $w_{ii} = 0$ and asynchronous network dynamics, evolution leads to a decrease in the network energy to local minima corresponding to the points of network stability. Therefore, to find the stationary state of a network, it is necessary to find a minimum of the energy functional on the basis of the values of neuron states. For this aim, one applies the gradient descent method to energy functional (4) and arrives at the following set of equations controlling network dynamics:

$$s_i = \frac{1}{2} \left(1 + \text{sign} \left(-\frac{\partial E}{\partial s_i} \right) \right). \quad (5)$$

However, the procedure of the iterative solution of set (5) for binary neurons often leads to a certain local minimum of the energy functional. Moreover, binary neurons are unreal idealization in many practical applications. For this reason, Hopfield proposed the generalization of his ANN model to the case of neurons with a continuous set of states. The standard way of going over to neurons with continuous states is the introduction of statistical noise to a system with further application of the mean-field theory (see, e.g., [9]), which leads to averaging the values of the neuron states and to the replacement of step activation function to the following sigmoidal function:

$$v_i = \frac{1}{2} \left(1 + \tanh \left(-\frac{\partial E}{\partial v_i} \right) \frac{1}{T} \right) = \frac{1}{2} \left(1 + \tanh \left(-\frac{H_i}{T} \right) \right), \quad (6)$$

where the temperature T corresponds to the statistical noise level; $H_i = \langle \sum_j w_{ij} s_j \rangle_T$ is the local mean field of the neuron states. Values of neurons v_i are no more integers; i.e., a neuron is considered as active, if $v_i > v_{\min}$. The temperature T is decreasing according to the *simulated annealing* scheme. The values of ANN neurons are iteratively updated by equation (6) until reaching a stable point. Since v_i are no longer binary, it is possible to monitor neurons that join the points of tracks and, stimulated by the weights, increase continuously their activity as the network evolves. The level $v_{\min} = 0.5$ is usually taken as a threshold activity level.

Hopfield's studies stimulated great attention to these networks, in particular, because their evolution processes turned out to be connected with numerous optimization problems, which are usually formalized as the problems of seeking a functional extremum under constraints on the parameters of the functional. Among these problems, there is a problem of identifying tracks on the basis of data obtained in coordinate measurements by current electronic detectors such as proportional and drift chambers, or microstrip detectors. The determination of connections between network neurons from measured coordinate values leads to a very large number of neurons and even larger number of interneuron connections, the majority of which are not associated with tracks. To overcome this «size curse», the weight functions were introduced in such a way as to stimulate connections between neurons belonging to the same track and the energy functional was supplemented by restricting terms forbidding intertrack connections and an excessive increase in the number of tracks.

The well-known *method of segments* [9] introduces neurons s_{ij} at the experimental point set on the plane to determine whether the points i and j are connected or not; i.e., whether

the given directed segment s_{ij} belongs to the track or not. In further calculations according to scheme (4), the states of neurons cease to be integer-valued. Their values $v_{ij} = \langle s_{ij} \rangle_T$ determine the activity level of neurons. The energy functional was defined as two-term expression: $E = E_{\text{cost}} + E_{\text{constraint}}$. For recognition of smooth straight-line tracks, the first, cost, term was taken so that it favors the short adjacent segments with a small angle between them and the second, constraint, term was taken either to forbid bifurcation (i.e., the cases where more than one neural segment connect to one point) or to permit them (in cases of particle decays) and to balance between the number of active neurons and the number of experimental points. An example of successful applications of this method can be found in [11].

The idea of *rotor ANNs* (see refs. in [10]) determines a neuron as a unit vector emanating from the i th point measured on the track. The energy function of the rotor network is defined as

$$E = -\frac{1}{2} \sum_{ij} \frac{1}{|r_{ij}|^m} v_i v_j - \frac{1}{2} \alpha \sum_{ij} \frac{1}{|r_{ij}|^m} (v_i r_{ij})^2.$$

Here, the first term forces neighbouring rotors to be close to each other. The second term is in charge of the same between rotors and track-segments. In their initial state rotor's directions were taken randomly, but after several tens of iterations of the ANN evolution they usually aligned along the track. However, in its original view this idea did not find many applications because of too high number of iterations and also because of its tendency to straight tracks. However, after significant modifications this approach was successfully applied for the track recognition by data obtained on the magnetic spectrometer ARES [10], as well as for processing data of the vertical sounding of the ionosphere [12]. We used such a remarkable feature of any circle as the parallelism of its two tangents in points i and j connected by a chord R_{ij} after reflecting one of these tangents respectively to this chord. That gives the very simple view of the energy function without any constraints:

$$E = -\frac{1}{2} \sum_{ij} v_i \cdot v'_j,$$

even if they have large curvatures [3, 10].

In the *radius-based function (RBF) neural networks* the basic equation (3) for the neuron response, which is scalar product of the input vector \mathbf{X} and the weight-coefficient vector \mathbf{W}_j , is replaced by the distance between them. A geometrical interpretation of these two approaches is shown in Fig. 2.

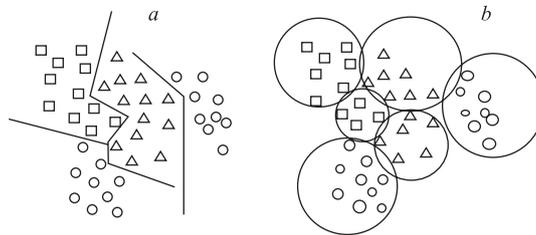


Fig. 2. The discrimination of three strongly overlapping classes by a multilayer perceptron (a) and a RBF network (b)

The metric $L_2 = \sum_j (x_j - w_{ij})^2$ is often used as a metric, although other metrics such as $\sum_j |x_j - w_{ij}|$, $\max_i |x_j - w_{ij}|$ or the so-called Makhalanobis distance are also in use depending on the problem to be solved. The activation of the RBF-neuron is then determined by its basis function, typically a Gaussian or sometime step function. Thus an RBF network has a hidden layer of radial elements, each reproducing the Gaussian response surface. The network output consists of their linear combination which can approximate an arbitrary function. For this reason, the training of RBF networks is very fast, though their applications are seriously complicated due to necessity to use a trial-and-error approach to determine Gaussian centers and widths.

Therefore, self-configuring radial basis (SCRB) neural networks were proposed in [4] to increase the speed of their functioning and to reduce essentially the number of neurons in the hidden layer. The SCRB network is constructing itself in the process of training, because in this process not only values of synaptic weights are calculated, but also the number of the neurons in hidden layers and even the number of layers are also determined. Details of the SCRB-network training procedure and examples of their applications to image recognition are given in [1, 3, 4]

3. WAVELET ANALYSIS AND ITS APPLICATIONS IN HEP

One-dimensional wavelet transform (WT) of the signal $f(x)$ has 2D form

$$W_{\Psi}(a, b)f = \frac{1}{\sqrt{C_{\Psi}}} \int \frac{1}{\sqrt{|a|}} \Psi\left(\frac{x-b}{a}\right) f(x) dx,$$

where Ψ is called a wavelet; b is a translation parameter; a is a dilation parameter or a scale, and $C_{\Psi} < \infty$ is a normalizing constant. As an example of a continuous wavelet transform, one can take the derivatives of the Gaussian function

$$\Psi(x) \equiv g_n(x) = (-1)^{n+1} \frac{d^n}{dx^n} e^{-x^2/2}, \quad n > 0,$$

with $C_{g_n} = 2\pi(n-1)!$. Known as *vanishing momentum* (or Gaussian) wavelets, $g_n(x)$ are especially suitable for handling bell-shaped signals, which can be well approximated by a Gaussian.

Such a remarkable fact that the wavelet transformation of a Gaussian $g(x; A, x_0) = A \exp(-(x-x_0)^2/2\sigma^2)$ looks as the corresponding wavelet was used in [13] for estimating the Gaussian parameters and for the very high resolution of bell-shaped overlapping signals.

Gaussian wavelets g_2 and g_4 were successfully used in [5] for analyzing angular distributions of secondary particles in high-energy nucleus-nucleus interactions. Presenting the distribution of the secondary particles on the pseudorapidity $\eta = -\ln(\tan(\theta/2))$ in an event as

$$f(\eta) = \frac{dn}{d\eta} = \frac{1}{N} \sum_{i=1}^N \partial(\eta - \eta_i),$$

where N is multiplicity of single charged particles in the event, θ is polar angle and η_i is pseudorapidity of i th particle, one obtains wavelet transform of $f(\eta)$ as

$$W_{\Psi}(a, b) = \frac{1}{N} \sum_{i=1}^N a^{-1/2} \Psi \left(\frac{x - b}{a} \right).$$

Figure 3 shows the g_2 spectrum of an event involving six particles with the different pseudorapidities. The values of the wavelet coefficients are presented by different grey levels. The whiter levels correspond to larger coefficient values. Thus, it demonstrates that the wavelet transform g_2 enables us to classify particles for scales less than 0.3. In [5], this approach was applied to analyze 884 and 504 events where sulfur and oxygen nuclei with momenta of 200 GeV/c per nucleon from the SPS accelerator (CERN) interacted with the NIKFI BR-2 nuclear photoemulsion piles. It was observed that the distributions of the groups on pseudorapidities have irregularities showing preferences of emission angles of the groups. The pairs of particles with close pseudorapidities, $\Delta < 10^{-5}$, are found for the first time. Gaussian wavelets were also applied in small-angle neutron scattering experiment for spectrum smoothing [14].

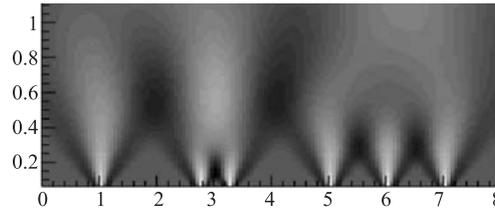


Fig. 3. The g_2 spectrum of an event involving six particles

The discrete character of experimental data measured by various counters and electronic detectors made more applicable the use of discrete wavelet family. It is especially advantageous for quite urgent problems of peak detection on experimental spectra. Special program packages were elaborated at the JINR Laboratory of Information Technologies, implementing fast algorithms for both discrete and continuous wavelets. It is intensively used for data smoothing and filtering.

There are a wide group of important problems of processing such 2D objects as images and, in particular, a new phenomenon observed in high-energy hadron collisions as jets, which are groups of particles strongly correlated in the space and time.

One step of a wavelet transform of a 2D signal is performed by transforming each dimension of the signal independently. Then two-dimensional sub-band s_1 that contains the low pass part in both dimensions is transformed further, as shown in Fig. 4, where d_k is high pass of the k th step of this decomposition, h_k and v_k are 1D wavelet transforms for horizontal and vertical image scans, correspondingly.

2D histogram of pseudorapidity versus φ using as weights transverse momenta with two jets of different width contaminated by 800 tracks uniformly distributed in space and then filtered is depicted in Fig. 5.

This algorithm is more effective than standard procedure LUCCELL and is quite resistant to background noise.

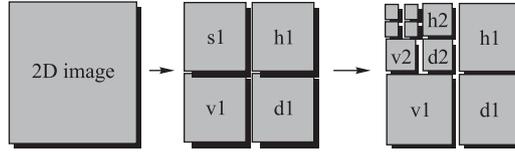


Fig. 4. Scheme of 2D wavelet transform

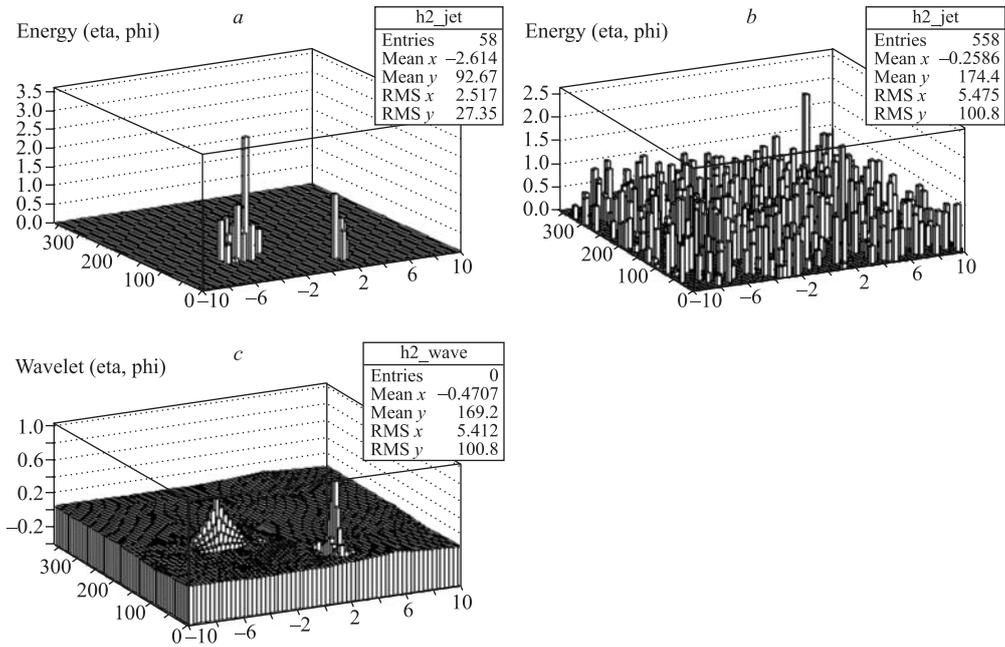


Fig. 5. a) Two simulated jets before adding background; b) uniform noise added; c) two peaks with different width after filtering by 2D wavelets

CONCLUSION

Three basic methods — robust methods of mathematical statistics, artificial neural networks and wavelet analysis — extensively applied at JINR to process recent experimental data have been expounded. This paper primarily covers studies in which scientists of the Laboratory of Information Technologies participated, in particular, in the collaborations with the leading physical centers such as CERN, DESY, BNL, GSI, etc. The basic principles of the reviewed methods and the most useful and promising examples of their applications have been discussed. In particular, examples of the successful applications of robust approaches under very different experimental conditions as well as the remarkable coincidence (see Sec. 1) of the robust approach with powerful elastic-tracking methods indicate that the robust approach is very promising. In the next section we paid more attention to the fully interconnected and radial basis neural networks than to multilayer feed-forward ones more popular in HEP because we wanted to stress the great potential of the first two types of ANN.

The wavelet analysis can provide large capabilities. Qualitatively new possibilities for analyzing complex implicit dependences can be provided by two-dimensional wavelets, which are little used yet.

REFERENCES

1. *Ososkov G. et al.* // Part. Nucl. 2002. V. 33, No. 3. P. 676–745.
2. *Ososkov G.* // Tatra Mountains. Math. Publ. 2003. V. 26. P. 1–16.
3. *Ososkov G., Stadnik A.* // Optical Memory & Neural Networks. 2003. V. 12, No. 3. P. 199–218.
4. *Ososkov G., Stadnik A.* // Nucl. Instr. Meth. A. 2003. V. 502/2-3. P. 529–531.
5. *Uzhinskii V. et al.* // Yad. Phys. 2004. V. 67. P. 156–162;
see http://arxiv.org/PS_cache/hep-ex/pdf/0206/0206003.pdf
6. *Akishin P. G. et al.* JINR Commun. E10-2006-48. Dubna, 2006.
7. *Dmitrievsky S., Gornushkin Yu., Ososkov G.* JINR Preprint E10-2005-216. Dubna, 2005.
8. *Muresan L. et al.* // JINR Rapid Commun. 1997. No. 1[81]. P. 27–44.
9. *Lasiuk B. et al.* // Proc. CHEP'98, Chicago, 1998.
10. *Kisel I., Neskoromnyi V., Ososkov G.* // Part. Nucl. 1993. V. 24, No. 6. P. 1551–1595.
11. *Ososkov G. et al.* // Math. Modelling. 1999. V. 11, No. 10. P. 116–126.
12. *Galkin I. et al.* // Radio Sci. 1996. V. 31, No. 5. P. 1119–1128.
13. *Ososkov G., Shitov A.* // Comp. Phys. Commun. 2000. V. 126, No. 1–2. P. 149–157.
14. *Islamov A. et al.* // Nucl. Instr. Meth. A. 2003. V. 502, No. 2–3. P. 498–500.