One of the most popular experimental techniques for investigation of brain activity is the so-called
method of evoked potentials: the subject repeatedly makes some movements (by his/her finger),
whereas brain activity and some auxiliary signals are recorded for further analysis. The key problem
is the detection of points in the myogram that correspond to the beginning of the movements. The
more precisely the points are detected, the more successfully the magnetoencephalogram is processed
aiming at the identification of sensors that are closest to the activity areas.

This paper proposes a statistical approach to this problem based on mixtures models that uses
a specially modified method of moving separation of mixtures of probability distributions (MSM-
method) to detect the start points of the finger’s movements. We demonstrate the correctness of the
new procedure and its advantages as compared with the method based on the notion of the myogram
window variance.

Introduction

The human brain is the center of the nervous system; the cerebral cortex processes all information.
Therefore, research on brain activity is one of the most important problems of modern medicine. In
order to investigate human mental activity and obtain new information concerning the structure and
relationships between the functional areas in the brain, various statistical methods can be used.

A precise preoperative localization of irreplaceable areas in the brain is the basis of surgical planning
and the minimization of possible postoperative aftereffects for patients with brain disorders. In clinical
practice, localization of the primary motor cortex (M1) is one of the most important and difficult prob-
lems, especially the problem of detection of the hand activity’s area in M1. Due to various involvements
of the central nervous system (e.g., as a result of epilepsy), information concerning the configuration of
areas by anatomical data is deficient.

Magnetoencephalography (MEG) is a method of preoperative localization of various brain areas. It
is often used in combination with magnetic resonance imaging (MRI) to localize activity in the areas of
the human brain. The method of evoked potentials is one of the most popular techniques to determine
the exact location of the motor areas of the cortex [1, 2].

The external actions and some movements of the subject can be considered as the events for the
analysis by the method of evoked potentials. Such methodology leads to the increase in the signal-to-
noise ratio and allows revealing the brain activity corresponding to the events. The noise in this case is a
superposition of the physical noise (say, generated by the noises of sensors, amplifiers, analog-to-digital
conversion, external signals, network interferences, etc.) and the physiological one (which is just the background brain activity). The main problem of the method relates to the detection of the starting point of a movement.

In brief, the scheme of the experiment is as follows. The subject puts his hand on the table and taps his forefinger for a few minutes. His/her MEG signals and myogram are recorded during the experiment. Additionally, the information about contacts with the table surface can be registered (say, by the accelerometer, photocell button, etc.). All signals are taken synchronously and with a strictly fixed sampling frequency. MEG signals are the dataset of sensors located on the subject’s head; each time series is called a channel.

The main aim of the analysis of the experimental data is the detection of the area of the cerebral cortex that is responsible for the beginning of the movement. This is a particular case of the so-called inverse problem of finding the source of a signal by the characteristics of the field generated by the source. One of the simplest solutions is to find the channel with the best response by averaging parts of the MEG signals over the starts of the movements. Then the corresponding curve can be chosen to improve the signal-to-noise ratio. However, start points cannot be determined by the MEG signals due to the part of the noise in the channel (it equals 0.95 and more). But the beginning of the movement can be found from the myogram, and this is sufficient for the averaging of MEG. At the same time, the values of a button can be used to adjust the fact of movement in the neighborhood of some point.

In the paper [3] an algorithm was proposed for solving the problem under discussion. That algorithm was aimed at the proper identification of reference points. For this purpose a simple property of a myogram was used: its window variance associated with muscle movements, due to the physiological characteristics of the human muscles, exceeds significantly the one related to the rest period (the window width is 30–50 ms).

The simplest reasonable mathematical model for a myogram is a cyclic non-stationary random process that can be represented as

\[ \xi(t) = \sum_i ((s_i(t) + \varepsilon_i(t) + \theta_i(t))1\{t_i \leq t < t_{i+1}\}), \]

where the random process \( s_i(t) \) corresponds to the signal component related to the finger movement (\( s_i(t) = 0 \) out of the movement period), \( \varepsilon_i(t) \) is the rest noise (it equals zero during a movement), and \( \theta_i(t) \) is the movement noise (it equals zero during a rest period); in neurophysiology the half-interval \( [t_i, t_{i+1}] \) is called an epoch (each epoch includes the rest interval before the movement and the movement itself), and \( i \) is the epoch number.

The window variance obtained from the myogram is also a cyclic non-stationary process, but much less contaminated with noise. The transfer to the window variance allows one to eliminate trends and to emphasize rest-to-movement transition moments, which are then determined by the threshold processing. The method proposed in [3] demonstrated very high accuracy. However, due to the noticeable non-normality of the distribution of noise, within the method proposed in [3] the thresholds were determined rather artificially.

The present paper proposes some developments of the method of [3]. In the following sections we describe some statistical procedures based on mixture models designed for precise detection of the points corresponding to the beginning of movements.

**Smoothing the signal by moving separation of finite mixtures**

To reveal the changes in the structure of the stochastic processes in time, the so-called method of moving separation of mixtures (MSM method) is successfully used. This method was proposed in [4]. As examples of efficient performance of this method, the papers [5–7] should be mentioned containing results for the financial markets, for the traffic in information systems, and for plasma turbulence, respectively.
The key point in this method is that the volatility of the process is decomposed into two components, dynamical and diffusive.

Within the framework of this method, the one-dimensional distributions of the increments of the basic process are approximated by finite location-scale mixtures of normal distributions. The theoretical background of these models can be found in [4].

To analyze the dynamics of the changes in the stochastic process, the problem of statistical estimation of unknown parameters of distributions should be successively solved for a part of a sample that moves in the direction of astronomical time (i.e., the initial sample is divided into sliding or moving sub-samples often called windows). Typically, the window (sub-sample) size is fixed. Once the analyzed parameters are obtained for a current location, the window should be moved by one element of the initial sample (i.e., the method will analyze the next sub-sample). This allows one to detect all possible changes in the behavior of components.

We assume that the cumulative density function for the corresponding moment of time (location of a window) can be represented as

\[
F(x) = \sum_{i=1}^{k} \frac{p_i}{\sigma_i \sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{ -\frac{(t - a_i)^2}{2\sigma_i^2} \right\} dt,
\]

where

\[
\sum_{i=1}^{k} p_i = 1, \quad p_i \geq 0
\]

(for all \(x \in \mathbb{R}, a_i \in \mathbb{R}, \sigma_i > 0, i = 1, \ldots, k\)). The model (1) is called a finite location-scale normal mixture. The parameters \(p_1, \ldots, p_k\) are weights satisfying (2). The parameter \(k\) is the number of mixture components. The parameters \(a_1, \ldots, a_k\) are associated with the dynamic component of the volatility (variance) of the process, and the parameters \(\sigma_1, \ldots, \sigma_k\) are associated with the diffusion one; see [4]. Namely, if \(Z\) is a random variable with distribution function (1), then its variance can be represented as the sum of two components:

\[
\text{D}Z = \sum_{i=1}^{k} p_i (a_i - \overline{a})^2 + \sum_{i=1}^{k} p_i \sigma_i^2,
\]

where

\[
\overline{a} = \sum_{i=1}^{k} p_i a_i.
\]

The first term on the right-hand side of (3) depends only on the weights \(p_i\) and the expected values \(a_i\) of the components of mixture (1). Since \(Z\) is an increment of the basic process, then \(a_i\) is the expected value of the increment, i.e., the trend component. Hence, the first component is the part of the total variance (changeability) that is due to existing elementary trends. It is called a dynamic component of the variance. At the same time, the second term on the right-hand side of (3) depends only on the weights \(p_i\) and the variances \(\sigma_i^2\) of components and represents the purely stochastic diffusive component of the total variance.

**Detection of starting points from the myogram by the MSM method within the finite mixture model**

The main idea of detection of movements by the MSM method is as follows. First of all, the dynamic and diffusive components should be estimated. In most situations, various modifications of the usual
EM-algorithm are involved [4]. The results of such a decomposition are shown in Fig. 3 (the initial myogram time series) and Fig. 2 (the dynamic and diffusive volatility (variance) components).

It seems more promising to use the obtained dynamic component of the total variance as the initial data (time series) for further analysis; see below. It is very interesting that the distribution of the values of the dynamic component obtained over the rest period is very far from being normal; see Fig. 1.

One of the key problems in detection of movements by the MSM method is the possible time delay. The analysis is based on sub-samples, and the influence of a new moment of time (one element of the sample) could appear with a delay. But we are interested in the precise detection of movements. The solution is based on idea of double sample processing by the MSM method: in the forward and backward directions. Comparing the probable points for both directions, we can find the right location of movements.

To detect the points of movements, we use the myogram as the initial data (sample). To avoid trends in the myogram (and in the dynamic volatility component), the differences of successive elements of the sample should be found. We tested various sizes of window (e.g., 20, 30, 50) to compare the character of components. In the paper we present the results for sub-samples of size equal to 50 elements in each position of a window.

The algorithm described above was applied to the myogram. The results for both directions are represented on Fig. 2. The $x$-axis corresponds to the time of experiments (in milliseconds), and the $y$-axis demonstrates the corresponding values of components. The thin solid lines are components (above is the dynamic component and below is the diffusive one).

We can see the balance statement of the subject at the beginning of components approximately until the point 3000. Note that it is very important to use the diffusive component to determine the mode due to possible ambiguities in the dynamic component. We could exploit the balance statement of the subject to estimate the bounds to be used for movements detection. Figure 1 demonstrates the histogram for the sample corresponding to the balance statement of the subject. The distribution is multimodal; it can be approximated by the mixture model (1), too. So, the standard technique based on the quantiles of unimodal distribution (like the $w$-sigma rule for a normal distribution) cannot be used to fix the bounds. Surely, the bounds can be chosen empirically with the help of additional information. For example, the data of a photocell button can be used to evaluate the bounds for all signals in the experiment by a few movements in the beginning of a myogram. In the following sections we discuss a more valid statistical approach based on the chi-squared tests.

The probable points of movements are determined by the crossings of the bounds. Next we need
to group the probable points of both directions. It seems reasonable to classify groups with the help of metric related to the window size. For example, all of the probable points that are located inside the interval with length $j \times \text{(window size)}$ should be classified as a group. And the starting point of a movement is the average of their time locations.

Figure 2 demonstrates the points of movements obtained by the method based on myogram window variance and by the method based on moving separation of mixtures: the vertical red dashed lines and the green triangles, respectively. We can see a good compliance between the results of the two statistical methods. The only exception is the triangle near the mark 6200 ms (the second triangle from the left). There are two reasons for that point. First, at the beginning of the experiment, the subject adjusts to the conditions and can move not only his finger but his head or leg as well. But this influences the myogram. Second, for the method based on myogram window variance, the information from the photocell button was available. It allows one to omit some false moments. For the MSM method, we had no data but the myogram.

We compared the results of the method for the component graphs. But we analyzed the differences in the myogram. So we need to compare the starting points of movements with the window variance method for the initial myogram as is demonstrated in Fig. 3.

![Fig. 2. The forward (left) and backward (right) direction of analysis. The thin solid lines are components (the dynamic one is located on top of the figure, the diffusive component is arranged below). The points of movements by the method based on myogram window variance and by the method based on moving separation of mixtures: the vertical red dashed lines and the green triangles respectively.](image)

The $x$-axis corresponds to the time of experiments (in ms), and the $y$-axis to the corresponding values of the myogram. The thin solid line demonstrates the signals of the myogram, and the vertical red solid lines and green triangles are the movement points for the two methods respectively. There are no changes for the conclusions mentioned above.

It should be noted that there are a few differences between forward and backward results. The shape and values of backward diffusive components almost coincided with the forward one except the end of the graphs. For the backward mode it is an adjustment area for a smoothed EM algorithm. A similar explanation is correct for the dynamic components too. Also, a backward dynamic component is a reflection of a forward one. This can be explained by the reasoning that the dynamic component is an expectation.
Detection of starting points by moving grid method from the dynamic component

The main idea is to treat the dynamic component as a random process (this is actually so) and process its distributions assuming that they are normal mixtures themselves (see Fig. 1). To separate these mixtures we use the modified two-step grid-based method [8], using only its first step. As a result, we obtain the values of the weights $p_i$ of the $i$th node of the grid for every window. We set the window size equal to 100 and the window shift equal to 1.

In order to compare the vectors $p_i$ of probabilities $(p_{i,1}, \ldots, p_{i,K})$ obtained on different nonintersecting (for independence) windows, for each window with number $i > 100$ we calculate the value

$$z_i = \|p_i - p_{i-100}\| = \left[ \sum_{j=1}^{K} (p_{i,j} - p_{i-100,j})^2 \right]^{1/2}.$$

Let us set the threshold value for $z_i$ in scope as $\theta = 0.97$ to highlight only extreme changes of the vector $p_i = (p_{i,1}, \ldots, p_{i,K})$ ($K$ is the number of nodes in the grid). All $z_i > \theta$ are colored purple. Purple events are grouped in tight groups so that every group is continuous, meaning that it consists of windows following one another. So we do not have any issues with detecting groups instead of single events.

From each group we take the first value and shift it by $+150$ (window size used for this decomposition (100) + window size used to obtain source data, the dynamic component (50)). We also take into consideration that every group might have a reflection after the detection value. We consider the next group as a reflection if it is no further than 300 windows apart from the original group. Reflections are excluded from the detection process.

The simple algorithm above produces the following time points (marked green, see Fig. 4) (left): 4074, 5608, 6256, 8446, 11284, 12938, 15017, 17327, 19685, 21321, 23531. The actual events are (red lines): 4032, 8443, 11298, 12917, 14976, 17326, 19688, 21337, 23539.

As we can see, between the first real event and the second real event we detect a couple of auxiliary events. All other points are estimated with an average accuracy of 7 ms, which is a great result.

The previous experiment led us to the conclusion that the vector $p$ of weights is usually volatile. Let us focus on the events themselves. If we analyze the behavior of vector $p$ inside the detected events, we
notice that it is changing slower than it does outside of detected events. As a benefit of this fact we use the chi-square test to detect the periods of “stability,” which are much easier to determine.

For each window, we calculate the $P$-value of chi-square test with 5 bins. Once this is done, we will select the windows where the $p$-value is almost equal to 1, i.e., greater than 0.9999. These windows are marked purple; see Fig. 4 (right).

We are interested in long periods of stability, not random (noisy) events, so we filter the groups that are less than 50 ms. Once we apply this filtering, we take the first point in each group. In the previous experiment we added a fixed value of +150 to each event. In this case the stability period is detected only once the window is fully over the event, so we need to add only +50 ms (the window size used to obtain the source data, the dynamic component).

Using the method above, we obtain the following detected points (marked green, shifted above purple intervals to make them more visible): 4048, 6249, 8450, 11286, 12941, 15017, 17330, 19703, 21340, 23553. The actual events are (red lines): 4032, 8443, 11298, 12917, 14976, 17326, 19688, 21337, 23539.

As before, the method detects events between the first actual event and the second one, but in this case we have only one false detection, which means that this method is better. The average accuracy of real events estimation is 12 ms.

Detection of starting points by moving grid method from the myogram

We applied the two-step decomposition algorithm [8] to the source data, the myogram time series, directly. We used two different distribution families: Generalized Hyperbolic (GH) [9] and Generalized Variance Gamma (GVG) distributions [10] to fit the data. The cumulative distribution functions of both of these laws are special normal variance-mean mixtures of the form

$$F(x) = \int_0^\infty \Phi\left(\frac{x - \alpha u}{\sigma \sqrt{u}}\right) dG(u), \quad x \in \mathbb{R},$$

where $\Phi(x)$ is the standard normal distribution function,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt,$$

and if $F$ is the GH-distribution function, then $G$ is the generalized inverse Gaussian distribution function [11] whereas if $F$ is the GVG-distribution function, then $G$ is the generalized gamma distribution function [12].
Both distribution families demonstrated good fit, but the GVG-distributions have slightly better $P$-values when applying the chi-square test. The figures below demonstrate the comparison of fitting GVG-distributions and GH-distributions for two randomly chosen windows. Based on the above, for further analysis we used only the GVG-distributions.

![Figure 5](image-url) GH vs GVG goodness-of-fit comparison, window No. 19251 (left) and window No. 10951 (right).

We have a particular interest in the $\alpha$ parameter. It is volatile, but near the events it tends to rapidly decrease, and then to rapidly grow, consistently demonstrating large absolute values. This means that we are able to detect events by watching the average absolute values of $\alpha$.

We use threshold 1.0 to select only high values (highlighted in purple, see Fig. 6). Similar to the cases above, we will group the data and select only the first value in a group. To produce the final prediction, we need to subtract 200 from the values (100, the initial windows size + 100 for average $\alpha$ calculation).

The result is: 3111, 4051, 4654, 7550, 7792, 11312, 12945, 15044, 17352, 19684, 21367. The actual events are (red lines): 4032, 8443, 11298, 12917, 14976, 17326, 19688, 21337, 23539.

As we can see, the “$\alpha$-method” detected false events around the first and second actual events. Also, the detection of the last event is missing. Other events are estimated very accurately ($\sim$ 25 ms).

![Figure 6](image-url) Movement detection using the $\alpha$ parameter of the GVG-distribution.

Other metrics calculated for particular GVG-distributions can also be used. Any numerical characteristics (moments, quantiles, assymetry, curtosis, etc.) can be calculated and used for detection if they
demonstrate specific behavior before/after or within the analyzed events.

Conclusions

The paper demonstrates the efficiency of the proposed statistical methods for solving an important medical problem. This method is based on mixture models and implements numerical procedures of separation of mixtures. Different techniques realizing this method are discussed. For example, the method based on the MSM approach could involve the processing of additional data (from accelerometer and photocell button) to precise location of its points. In addition, using the model of probability mixture, we obtain a convenient tool for further theoretical researches in the important field of modern medicine. The methods could be used for processing various types of human brain signals.

REFERENCES

1. C.D. McGillem and J.I. Aunon, “Analysis of event-related potentials,” in: Methods of Analysis of Brain Electrical and Magnetic Signals: EEG Handbook, A.S. Gevins and A. Remond (eds.), Elsevier Science Publishers, Amsterdam (1987), pp. 131–169.

2. M. Fabiani, G. Gratton, and K. Federmeier, “Event-related brain potentials: methods, theory and applications,” in: Handbook of Psychophysiology, J.T. Cacioppo, L.G. Tassinary, and G.G. Berntson (eds.), Cambridge University Press, Cambridge (2007), pp. 85–119.

3. T. Zakharova, S. Nikiforov, M. Goncharenko, M. Dranitsyna, G. Klimov, M. Khaziakhmetov, and N. Chayanov, “Signal processing methods for localization of nonrenewable brain regions,” Syst. Means Inform., 22, No. 2, 157–175 (2012).

4. V. Yu. Korolev, Probabilistic and Statistical Methods of Decomposition of Volatility of Chaotic Processes, Moscow University Publishing House, Moscow (2011).

5. A.K. Gorshenin, V.Yu. Korolev, and A.M. Tursunbayev, “Median modification of EM- and SEM-algorithms for separation of mixtures of probability distributions and their application to the decomposition of volatility of financial time series,” Inform. Appl., 2, No. 4, 12–47 (2008).

6. A. Gorshenin and V. Korolev, “Modeling of statistical fluctuations of information flows by mixtures of gamma distributions,” in: Proceedings of 27th European Conference on Modelling and Simulation, Digitaldruck Pirrot GmbH, Dudweiler (2013), pp. 569–572.

7. A.K. Gorshenin, “The information technology to research the fine structure of chaotic processes in plasma by the analysis of the spectra,” Syst. Means Inform., 24, No. 1, 116–125 (2014).

8. V. Yu. Korolev and A. Yu. Korchagin, “Modified grid-based method for decomposition of normal variance-mean mixtures,” Inform. Appl., 8, No. 4, 11–19 (2014).

9. V. Yu. Korolev, “Generalized hyperbolic distributions as limit laws for random sums,” Theor. Probab. Appl., 58, No. 1, 117–132 (2013).

10. L. M. Zaks and V. Yu. Korolev, “Generalized variance gamma distributions as limit laws for random sums,” Inform. Appl., 7, No.1, 105–115 (2013).

11. B. Jørgensen, Statistical Properties of the Generalized Inverse Gaussian Distribution, Springer, Berlin (1982).

12. E. W. Stacy, “A generalization of the gamma distribution,” Ann. Math. Stat., 33, 1187–1192 (1962).