Tutorial of Parametric Signal Fitting by Exponentially Modified Gaussian (PSF-EMG)

Tutorial of the EMG program developed in the work:

Parametric signal fitting of highly asymmetric voltammograms by using the exponentially modified gaussian (EMG) function

S. Cavanillas, N. Serrano, J.M. Díaz Cruz*, C. Ariño and M. Esteban

Departament de Química Analítica. Universitat de Barcelona
Martí i Franquès, 1-11, E – 08028 – Barcelona (Spain)
E-mail: josemanuel.diaz@ub.edu

FILES INVOLVED

In order to be able to run **EMG** from Matlab, the following files must be copied to a directory included in the Matlab path:

EMG.m   Matlab program for the fitting of experimental data matrices by means of EMG

EMGcalc.m   Matlab programs which are used by **EMG.m**

negtozero.m

SAMPLE DATA

The workspace Matlab files **simulatedataI.mat** contains simulated data to practise some aspects of the EMG procedure according to the present tutorial.
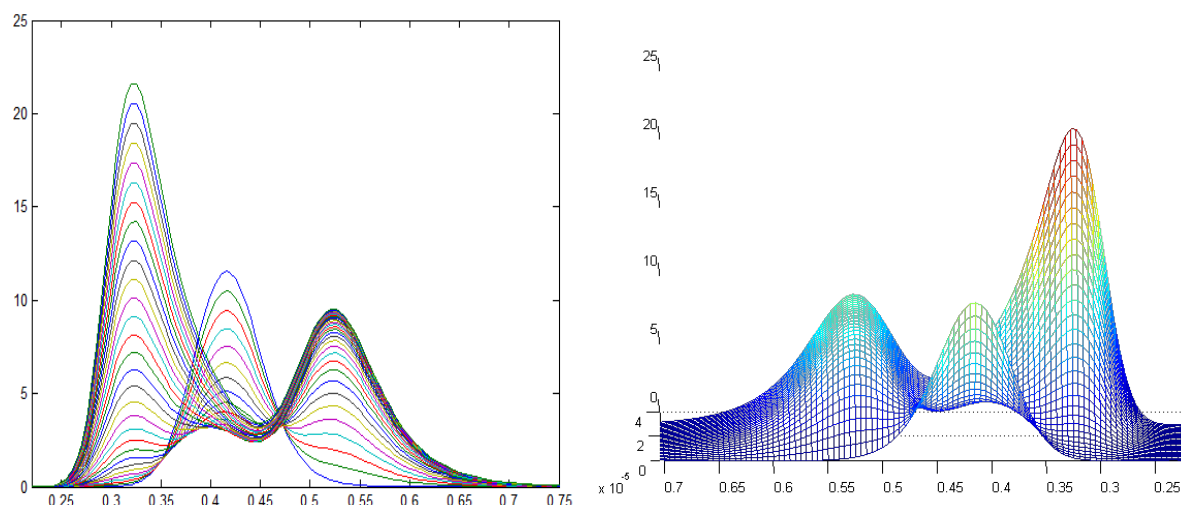
ANALYSIS OF SAMPLE DATA

When you load the file sample.mat in Matlab you can find the following matrices:

I   (141 rows, 30 columns)   current data matrix including 30 voltammograms simulated for 141 potential values
E   (141 rows, 1 column)   column vector with the potential values
var   (30 rows, 1 column)   column vector with the number of the voltammograms

You can look at the data matrix in both 2D and 3D ways by means of the commands

```
plot(E,I)                              mesh(var,E,I)
```


Before starting the analysis of the **I** matrix, it is necessary to provide a selectivity for every component. This can be done visually. In this case, 3 components (peaks) are necessary and the selectivity of each peak is provided inside the **csel** matrix. In the **simulatedataI**, the elements of **csel** are:
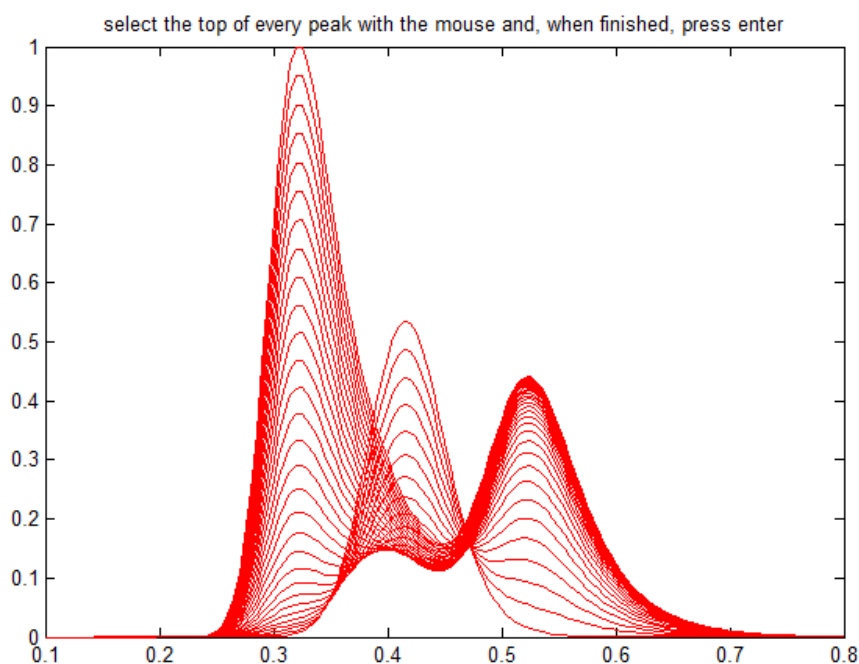
```
4    1    2
30   30   30
```

Where every column refers a component, and the first and second row define its existence range (i.e., the scan numbers when the peak appears and disappears). For instance, component 1 appears in voltammogram number 4 and exists until voltammogram 30. The numbers in these two rows have to be deduced by a careful inspection of the data matrix, voltammogram by voltammogram.
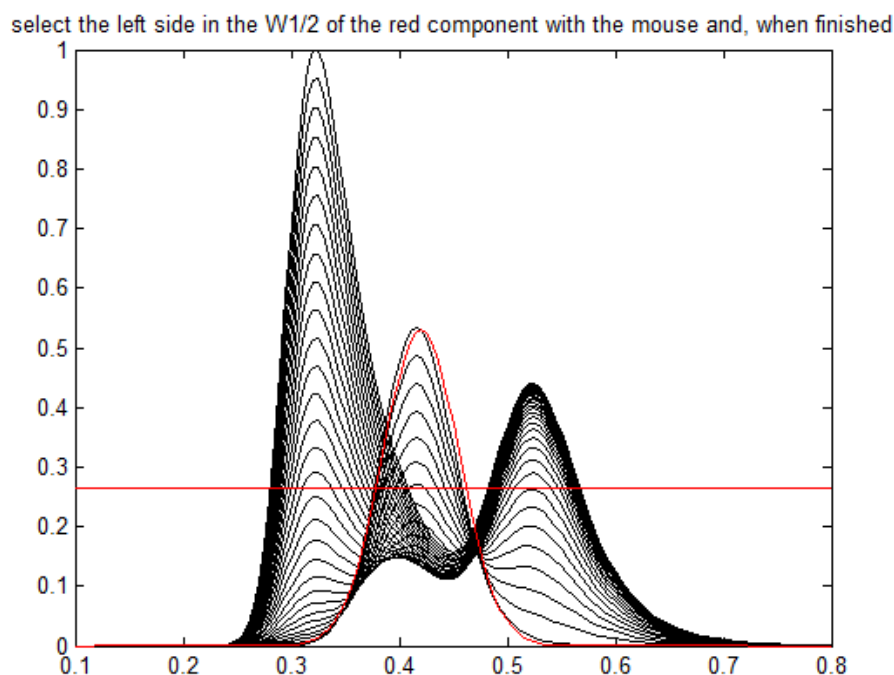
Once **csel** is ready, **EMG** program can be launched with the command:

>> [Irep,a,b,c,d,Area,algorithm]=EMG(I,E,var,csel)

Firstly, you get a Figure with a red 2D plot of the experimental matrix. Then you have to click with mouse the top of every component that you want to consider, which defines both height and position of the peak. Then, In the case of the **simulatedataI**, 3 components should be defined, which produce a graph like this (it can change, depending on where you arbitrarily choose the position and the height of the peaks. The best is to try in these parts of the experiment where the peaks are less overlapping to other signals, but not too far from the region where they appear for the first time):

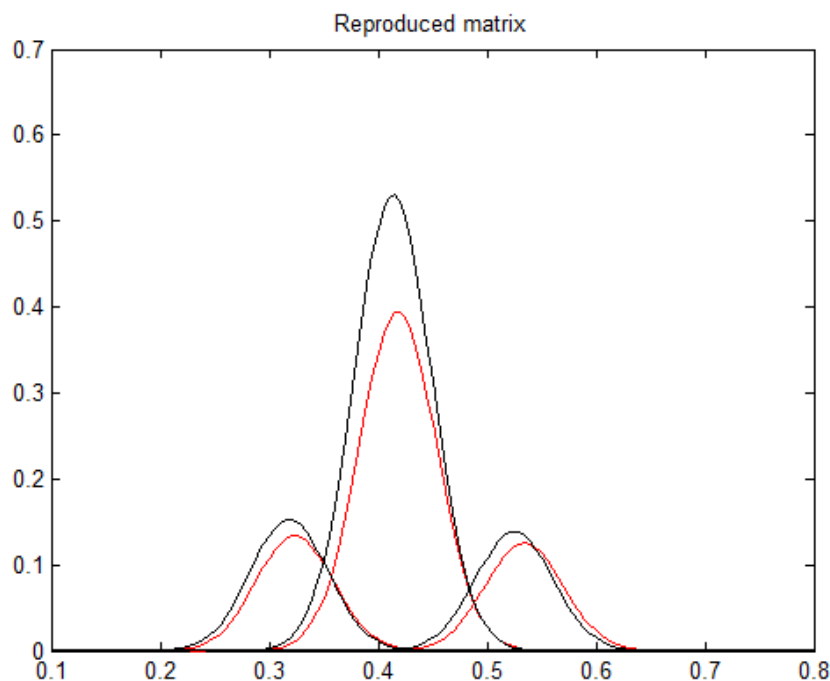select the top of every peak with the mouse and, when finished, press enter

When you have completed the selection, just press enter and a first estimation of all peaks will be displayed over imposed to the black-coloured plot of the matrix.

select the left side in the W1/2 of the red component with the mouse and, when finished
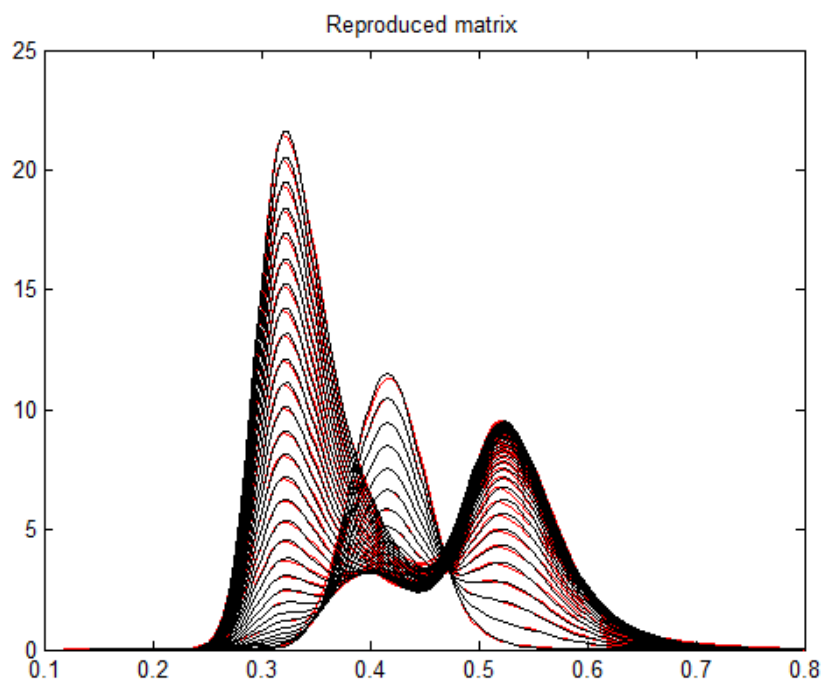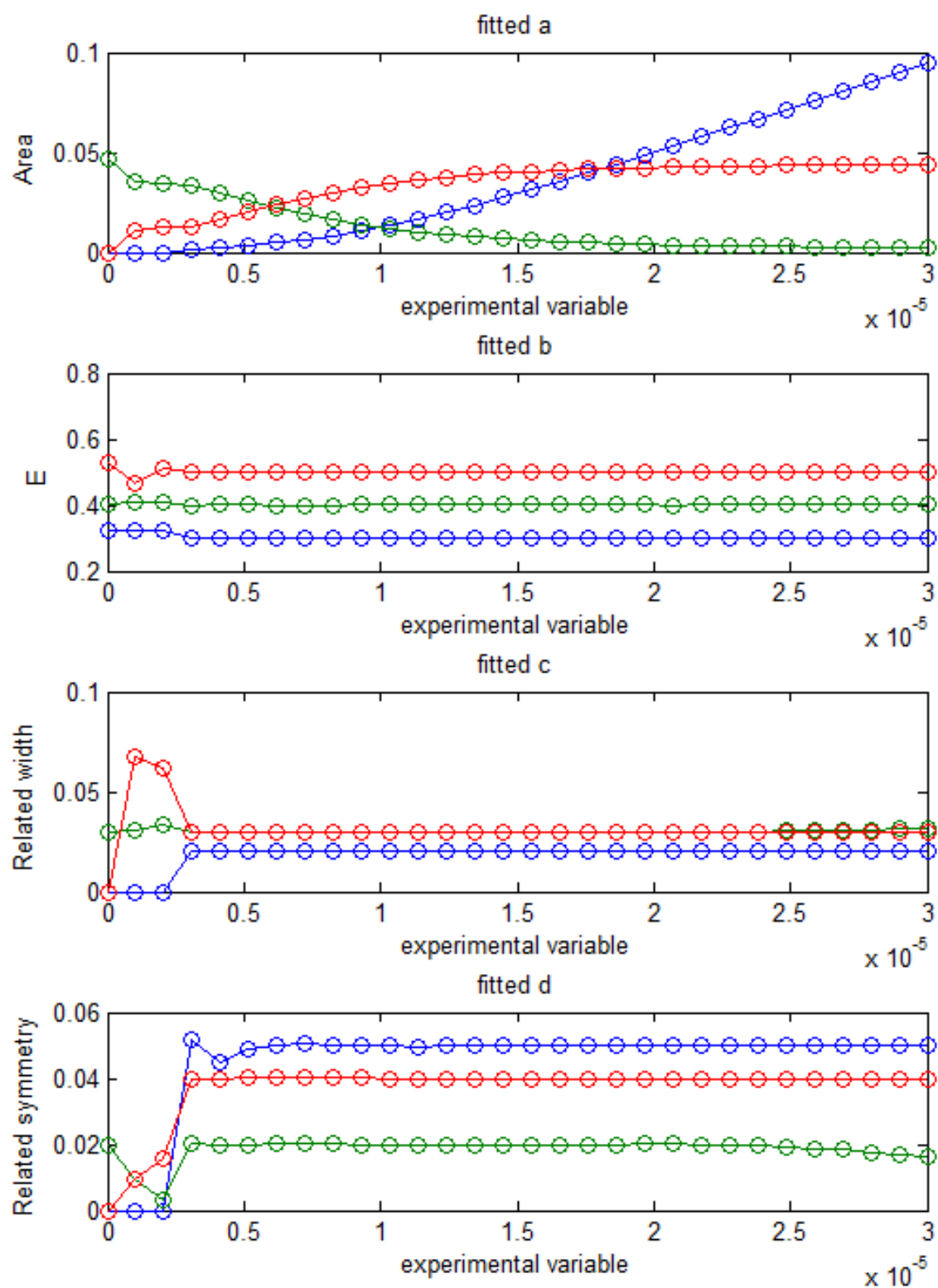
In this graph, peaks are red and denote the experimental points with red line. A horizontal line is drawn at the half height of the peak. Then you have to click twice around every peak, to define the left and right boundaries, respectively.

Finally, the program gets the initial estimation to start the iteration and it selects the best algorithm to fit each peak. Also, you get a Figure with a red 2D plot of initial estimation of each peak (red line).
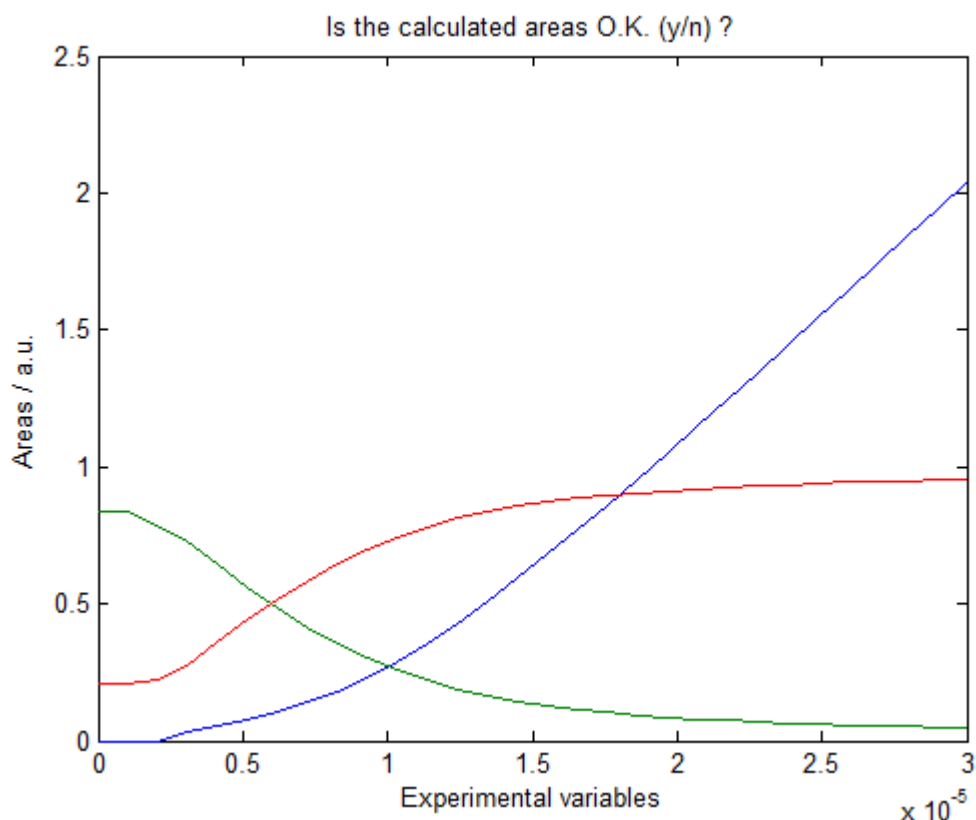

Reproduced matrix

After you press a key, two new graphs appear: i) comparison between the reproduced data (red lines) and the experimental data (black lines); ii) the analytical parameters computed from the optimised values of a, b, c, d. These are the area, the parameter related with peak potential, the parameter related with the width and the parameter related with the symmetry, respectively.

Reproduced matrix

After pressing a key again, you get the concentration profiles by area.

Is the calculated areas O.K. (y/n) ?

Finally, the program asks if you agree with the adjusted parameters and the concentration profiles. Then select **y** or **n** and press enter. If you press **n** the program restart again.

*Is the calculated areas O.K. (y/n) ? **y***

If you press **y** the lack of fit appears:

>> lack of fit, in %

      lof =   0.4843

The program generates the following matrices:

Irep (30 x 141):   reproduced matrix.
a (30 x 3):   fitted "a" values for each component at each voltammogram.
b (30 x 3):   fitted "b" values for each component at each voltammogram.
c (30 x 3):   fitted "c" values for each component at each voltammogram.
d (30 x 3):   fitted "d" values for each component at each voltammogram.
C (30 x 3):   concentration profiles for every component.
Area (30 x 3):   peak potential E for every component.
Algotihm (1x3):   algorithm used for every component.