Text of the presentation at:

International Evoked Response Audiometry Study Group (IERASG-2023)

Deconvolution for flexible recording of transient evoked potentials

Joaquin T. Valderrama[1,2,3], Angel de la Torre[1,2], Isaac M. Alvarez[1,2], Jose C. Segura[1,2], Jose L. Vargas[4,5]

[1]*Department of Signal Theory, Telematics and Communications, University of Granada, Spain.*
[2]*Research Centre for Information and Communication Technologies (CITIC-UGR), University of Granada, Spain.*
[3]*Department of Linguistics, Macquarie University, Sydney, Australia.*
[4]*Department of Surgery and its Specialties, University of Granada, Spain , Granada, Spain.*
[5]*ENT Service, Hospital Universitario Clinico San Cecilio, Servicio Andaluz de Salud, Granada, Spain.*
E-mail: jvalderrama@ugr.es (J.T. Valderrama)

Slide 1 (Title and authors):

Good morning. Today I will present the fundamentals of a mathematical algorithm that is able to deconvolve overlapping auditory evoked potentials (AEPs).

Slide 2 (The team):

The work that I will present here today has been developed by my research team, which is directed by Prof. Angel de la Torre. The team is composed of researchers from the University of Granada and medical personnel from the San Cecilio Clinical University Hospital. We are located in Granada, in the south of Spain, and these are some photographs from the University Hospital, our Technical School, and our Research Centre.

Slide 3 (Conventional recording of AEPs):

The conventional recording of evoked potentials consists of averaging the EEG segments that contain the evoked responses. This approach requires the inter-stimulus interval (ISI) to be larger than the duration of the response, in order to avoid contamination from adjacent responses. This is an important limitation if we want to study how the auditory system responds to complex stimuli, such as real speech or music, since the acoustic events that evoke a neurophysiological response certainly do not meet this criterion. So how can we estimate responses that overlap with each other?

Slide 4 (Matrix deconvolution):

One approach is through deconvolution.

Slide 5 (The EEG as a convolution model):

The recorded EEG can be modeled as the convolution of a stimulus sequence with an evoked response plus noise.

Slide 6 (Matrix formulation of the convolution):

Importantly, convolution can be represented as a matrix operation. This way:

- The EEG $\mathbf{y}$ is a column vector of N samples (N being the number of samples of the EEG, several millions samples for example).
- $S\mathbf{x}$ is the convolution operation.
- S is a matrix of N columns and J rows, being J the number of samples of the AEP (e.g. 100 samples for an ABR of 10 ms sampled at 10 kHz). This matrix is built by presenting the stimulus sequence in the first column (mostly 0s, and 1s in the start of each stimulus), and shifting this vector one sample every column until we complete the matrix.
- $\mathbf{x}$ is the AEP, which is a column vector of J samples.
- And $\mathbf{n}$ represents the noise, and has the same size as the EEG.

Slide 7 (Matrix deconvolution):

In fact, the matrix formulation of the EEG convolution model can be seen as a system of equations. The first component of the EEG y is the matrix multiplication of the first row of S and the AEP $\mathbf{x}$ vector plus the first element of the noise vector $\mathbf{n}$, and so on.

We can see that we have a system with N equations (one equation for each sample of the EEG, meaning a large number of equations) and J unknowns (as many unknowns as the size of the AEP response). This is an over-determined system of equations.

Since the AEP has many components (i.e. J dimensions), it is difficult to visualize the solutions of this system of equations, but let's imagine that our AEP only has 2 samples (x_1 and x_2). This way we would have (again) N equations, but this time, only 2 unknowns. It is now easier to visualize that each equation would lead to a line in the 2-dimensional space (x_1, x_2). We can also observe that due to the noise there is not a single solution. However, there is one unique solution that minimizes the error (that is the least-squares solution). It is well known that the least-squares solution to this system of equations is the matrix division of $(S^T\mathbf{y})$ by (S^TS), or in other words, the inverse of (S^TS) applied to $(S^T\mathbf{y})$. This solution is the deconvolved AEP $\hat{\mathbf{x}}$.

Slide 8 (Matrix deconvolution, 2):

How is the AEP estimated?

On one hand, S-transposed (S^T) applied to S leads to the autocorrelation matrix of the stimulus sequence, which is a $J \times J$ square matrix (being J the length of the AEP, much smaller than N). And on the other hand S^T applied to $\mathbf{y}$ leads to the averaged EEG, i.e. the signal resulting from averaging the EEG segments in an equivalent way as in the conventional method. The averaged EEG ($S^T\mathbf{y}$) has the same dimension (J) as the AEP response $\mathbf{x}$.

This means that to estimate the deconvolved response, first we need to invert the $(S^T S)$ square matrix, and multiply it by the $(S^T \mathbf{y})$ vector.

It should be noted that when the responses do not overlap, the $(S^T S)$ square matrix is the identity matrix and (since the inversion of the identity is still an identity matrix) the least-squares solution is the synchronous average of the response. However, when responses overlap, the $(S^T S)$ square matrix will not be an identity matrix, and we will have to invert that matrix to deconvolve the AEP.

Slide 9 (Example of matrix deconvolution):

Let's see how this process works with an AEP signal of 200 ms duration, sampled at 16,384 Hz. This AEP has 3,277 samples.

Considering that the matrix $(S^T S)$ has a dimension $3,277 \times 3,277$, how long would deconvolution take? The complexity of a matrix division increases with the size of the matrix. For this example, using a personal computer, deconvolution takes around 9 seconds, which is a feasible processing time in most applications.

Slide 10 (Multi-response deconvolution, definition):

For now we have considered only one type of stimulus that evokes one response ($\mathbf{x}$), and we have described the deconvolution process to estimate $\mathbf{x}$ when the responses overlap.

But what would happen if different stimuli are presented? In this case, it would be reasonable to assume that different stimuli would evoke AEPs of different morphology. And the challenge is to estimate these different responses $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$ when they overlap. This is called multi-response deconvolution.

Slide 11 (Multi-response deconvolution, complexity):

We have already seen that for 1 class of AEPs, deconvolution involves inverting a matrix of size J by J.

However, when we are considering K different classes, you may remember from the presentation of my colleague Professor de la Torre (on the first day of this conference) that the size of the matrix to be inverted is, this time, KJ by KJ.

Slide 12 (Example of multi-response deconvolution):

This means that in our previous example of an AEP of 3,277 samples, if we had 10 different classes, the size of the matrix to invert would increase to 32,770 $\times$ 32,770. Please note that a matrix of this size would have around 1 billion numbers, each of them represented with 8 bytes would lead to a matrix that needs 8,6 GB just to store the matrix.

In this example, we should note that the time required for matrix division increases to over 1,000 seconds. And importantly, for more than 10 classes, the memory requirements to perform deconvolution are not manageable for a personal computer.

The solution we found to overcome this problem consists of reducing the dimensionality, I will show how.

Slide 13 (LDFDS):

The method is called ”Latency-dependent filtering and down-sampling”.

Slide 14 (Representation of evoked potentials):

The neurophysiological activity along the ascending auditory pathway is conventionally recorded via ABR, MLR and CAEP. ABRs present energy between 100 and 3000 Hz, MLRs between 10 and 300 Hz, and CAEPs between 1 and 30 Hz. Since ABRs, MLRs and CAEPs present energy in different frequency bands, they are conventionally recorded as separated responses.

However, it would be desired to represent all the components in the same figure. Please note that this picture is a diagram, not a real response, because obtaining a signal like this is not straightforward. This type of representation would require the signal to be filtered according to its latency.

An optimal filter would let pass frequencies between 100 and 3000 Hz in the ABR section, 10 and 300 Hz in the MLR section, and 1 and 30 Hz in the CAEP section.

This is what we have called Latency-dependent filtering.

Slide 15 (Filters):

Before describing the filter we have designed, please allow me to review what a filter is like. A filter is equivalently described by its impulsive response (in time) or its frequency response. A low-pass filter provides an average, in time, of the input signal, according to its impulsive response or equivalently attenuates high frequency components, and presents this type of morphology, in a way that the width of the filter defines the cut-off frequency.

Narrower filters have a broader bandwidth, which are appropriate for ABRs. Broader filters have a narrower bandwidth, which is appropriate for MLRs. And very broad filters with narrow bandwidths are appropriate for CAEPs.

Slide 16 (Conventional filtering):

Filtering consists of convolving a signal with a filter, and as we mentioned earlier, convolution can be represented as a matrix operation.

The values of the filtered signal result from the matrix product of each row of the H matrix with the evoked potential $\mathbf{x}$. We observe that conventional filtering uses the same filter for all the samples of the signal, and therefore, all the samples are filtered with the same bandwidth.

Slide 17 (Latency dependent filtering):

Interestingly, the matrix formulation of filtering enables the implementation of a latency-dependent filtering by modifying the properties of the filters in each row. We can use narrower filters (with a bandwidth appropriate for ABRs) in earlier latencies, and broader filters (with lower bandwidths) in higher latencies. And that is how we implement latency-dependent filtering.

Slide 18 (LDF and down-sampling):

Further, the reduction of the bandwidth as latency increases lets us also implement latency-dependent down-sampling by appropriately selecting rows of the convolution matrix. The large bandwidth in earlier latencies require a high sampling rate, but as the bandwidth reduces, the sampling rate decreases.

In fact, the way we build this filter is by distributing the functions linearly in the logarithmic time domain, which is inspired by the fact that evoked potentials present around 3 to 4 large oscillations per decade. In this example, there are 15 functions per decade, and since we cover approximately 2 and a half decades, there are a total of 43 functions.

When we represent the base of functions in the linear time domain, we can observe how the proposed filter implements a latency-dependent filtering (because the width of the filter changes with latency) and latency-dependent down-sampling (because sampling rate changes in a non-uniform way according to the latency).

Finally, the base of functions is orthonormalized so all the functions have the same energy.

Slide 19 (LDFDS example):

Let's see how this process works with an example of a full-range auditory evoked response, including ABR, MLR and CAEP components. This AEP has a duration of 500 ms, sampled at 20 kHz, thus leading to 10,000 samples. This figure shows that representing this evoked potential in the linear time scale is not appropriate, because ABR and MLR are squeezed in earlier latencies, and the identification of their components is not clear.

Slide 20 (LDFDS example):

In contrast, all the components can be easily identified in the logarithmic time scale. However, this figure also shows that representing the evoked potential in the logarithmic time scale involves high-frequency noise in later latencies (which is not part of the biological response).

The first step in Latency-dependent filtering is projecting the evoked potential on the reduced space defined by applying the V matrix to the evoked potential. This matrix has as many rows as functions (43 in our example) and the same number of columns as samples of the evoked potential (10,000 in this example). Importantly, it should be noted that the AEP is represented in this space with only 43 coefficients, rather than 10,000 samples as in the time domain, which involves an important dimensionality reduction.

Slide 21 (LDFDS example):

Once projected, we can apply V-transposed to the projected AEP to represent this signal back in the time domain (that is the filtered response).

An important contribution of this methodology is the definition of a vectorial subspace of a very low dimensionality where we can represent evoked potentials. Representing evoked potentials with just a few coefficients could be interesting for many applications, such as for features extraction in big data analyses, automatic classification, data storage, data transmission, etc.

Slide 22 (Optimized deconvolution):

Another important advantage of reducing the dimensionality is that we can perform deconvolution in the reduced space to reduce computational load.

Slide 23 (Subspace constrained deconvolution):

Compared to matrix deconvolution performed in the time domain, where the size of the matrix to be inverted is $J \times J$, we can perform deconvolution in the reduced space defined by latency-dependent filtering and down-sampling.

By working out the matrix operations, we see that deconvolution in the reduced space involves inverting a matrix of a much lower size than in the full domain, as J-reduced is much smaller than J.

Slide 24 (Example of multi-response deconvolution in subspace):

In our previous example, the AEP is no longer represented by 3,277 samples, but with just 91 samples using 40 functions/decade. If we considered 10 different classes, the size of the matrix to invert would be 910 by 910, and deconvolution takes only 30 seconds. Further, for more than 10 classes, deconvolution is feasible as we no longer have memory limitations.

Slide 25 (Research possibilities):

Overcoming the dimensionality limitation in deconvolution substantially increases flexibility in experimental design. Deconvolution can be applied in a broad range of applications and experiments. I will briefly present some of the fields where we have applied deconvolution, but the possibilities are countless.

Slide 26 (Multi-level stimulation and threshold estimation):

We have applied deconvolution with clicks presented at fast rates whose level was randomized between 0 and 80 dB HL. This allowed us to categorize the click events in terms of their level, for example, by making 8 categories, each with a range of 10 dB. This experiment allowed us to use multi-response deconvolution to obtain the full trace of AEPs at different levels simultaneously. This figure also shows the value of presenting all the components of the auditory pathway in a single plot, from cochlea to cortex.

Slide 27 (Binaural hearing):

Another field where we have used deconvolution is to study binaural-hearing processes.

Binaural hearing is often studied via the Binaural Interaction Component, which is the response resulting from suppressing the sum of monaural responses to the binaural response. The problem of the BIC is well known: due to its low amplitude, the test-retest reproducibility of this component is not high, and often it cannot be recorded in all participants.

In contrast, we used a tone burst presented binaurally, in which the stimulus of one ear was delayed over the other ear to recreate a sound source being on the left or on the right of the participant. We presented several repetitions of these stimuli in a way that the source location changed from time to time. This experimental paradigm led to a model in which each stimuli would evoke a neurophysiological response (that would be sequence 1, in blue), and on top of that, another additional response would be evoked every time the sound source changed (which is sequence 2, in red).

By applying multi-response deconvolution considering these two categories and LDFDS, we obtained the two responses presented in the bottom left figure. These results demonstrate the coexistence of multiple overlapping neurophysiological processes. On one hand, the blue signal is associated with hearing detection, and we see that the tone-burst stimuli evoke neurophysiological activity all along the auditory pathway. On the other hand, the red signal shows neurophysiological activity associated with binaural-hearing processes, and we see that these processes are manifested at central stages, with the first component appearing at 50 ms.

Also, these results were also reproducible at individual level.

Slide 28 (Neural adaptation):

One more field appropriate for deconvolution is to study neural adaptation processes via bursts of clicks.

In this stimulus paradigm, we proposed a model in which the response evoked by the onset click of the burst would be different from the responses evoked by the remaining clicks.

We used multi-response deconvolution and LDFDS to obtain the full-range responses evoked by these two categories of events. Results showed large adaptation at cortical level, and importantly, cortical components could be recorded at that fast rates. Adaptation at brainstem level was not manifested, probably because the clicks presentation rates was not high enough. Results were clear both in the grand-average signal resulting from 10 participants, and also in all participants at individual level.

These results are also a clear example of the potential of LDFDS to provide a comprehensive representation of the auditory pathway, presenting in a single plot components from the cochlea to the cortex.

Slide 29 (Responses to speech):

Finally, deconvolution can be applied to understand how the auditory system encodes complex stimuli such as real speech. The conventional approach of estimating the neurophysiological response evoked by the envelope of speech assumes that there is a unique response evoked by speech.

However, it may be reasonable to assume that complex stimuli such as speech evoke multiple responses of different morphology, as we observed in previous examples. Therefore, we could use multi-response deconvolution to provide a better model of the non-linearities of the auditory system. We could consider a simple model of 2 categories as shown in this diagram, similar to the bursts of clicks presented earlier, or we could make more complex categories. This is ongoing research in our team, and we hope we to some progress in the next IERASG conference in Colorado.

Slide 30 (References):

The matrix deconvolution, LDFDS, and deconvolution in the subspace methods have been published in a number of publications. Importantly, these publications provide not only the mathematical formulation of the algorithms, but also extensive supplementary material including Matlab / Octave toolboxes that run simulations and implement the methodologies, aimed at facilitating any interested researcher, clinician or industry to understand and use these techniques.

In addition, an article presenting multi-response deconvolution is currently under review.

Slide 31 (Summary, projects and web):

To summarize my presentation:

- Matrix Deconvolution enables evoked potentials to be recorded at fast rates, increasing flexibility in experimental design.
- The compact representation provided by latency-dependent filtering and down-sampling (LDFDS) facilitates (1) a comprehensive representation of evoked potentials along the auditory pathway, and (2) an important dimensionality reduction.
- Performing deconvolution in the reduced space defined by LDFDS significantly reduces computational load.
- Multi-response deconvolution is appropriate to model multiple neurophysiological processes evoked by complex stimuli.
- MATLAB / Octave toolkits with functions and simulations are available to help understand and use these methodologies.

This slide includes the projects providing financial support to our research. Some complementary information can be found in our web, https://wpd.ugr.es/~sig.proc.audiology/. Thank you very much for your attention.