- Good morning.

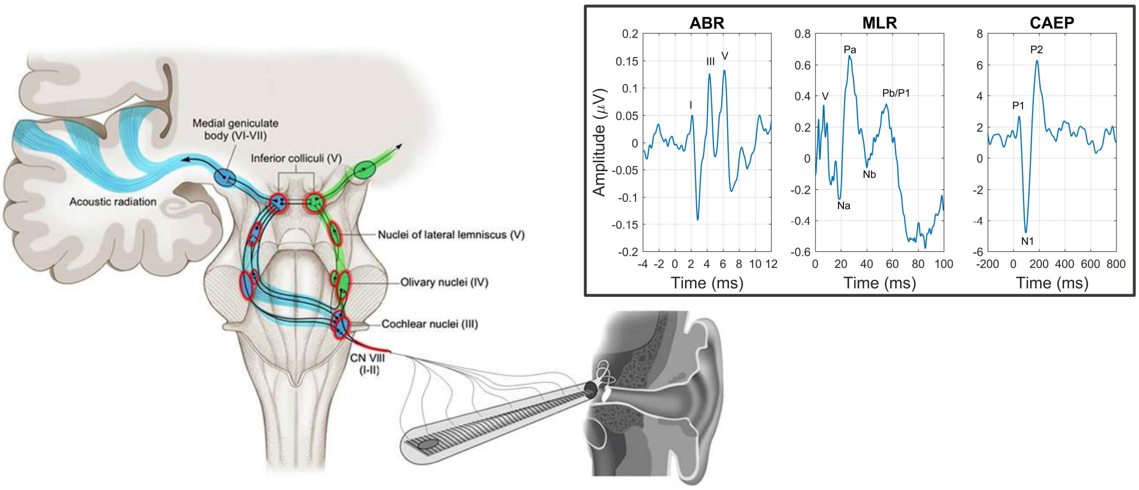
- Today I will talk about emerging technologies that have potential to inspire new developments in the field of auditory evoked potentials.

- The work that I will present here today has been developed by my research team, which is headed by Prof. Angel de la Torre.

- The Team is composed of researchers from the University of Granada and medical personnel from the San Cecilio University Hospital.

- We are located in Granada, in the south of Spain, and these are some photographs from our Technical School, the University Hospital, and our Research Centre.

Auditory evoked potentials (AEPs)



UNIVERSIDAD DE GRANADA

3

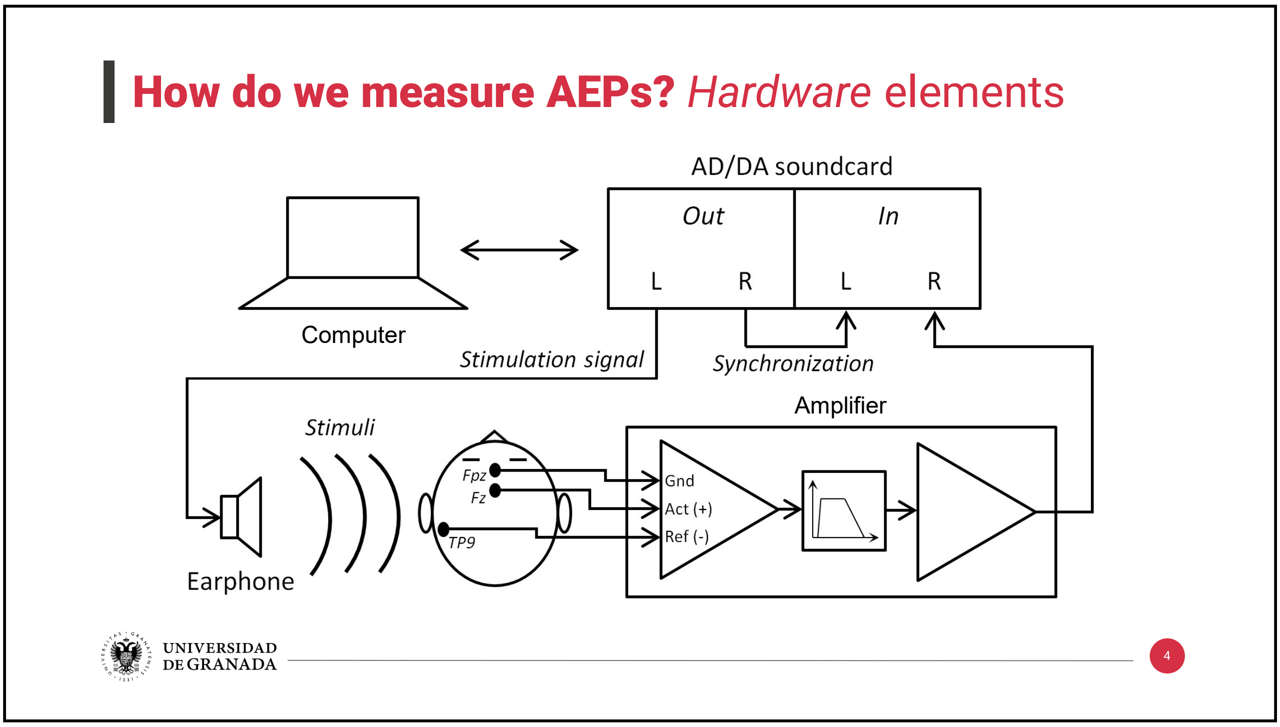
- Auditory evoked Potentials (AEPs) are voltage peaks that represent the activation of the neurons at different stages of the ascending auditory pathway.

- For example, the wave I of auditory brainstem responses (ABRs) shows neural activity elicited in the cochlea, the brainstem and the midbrain.

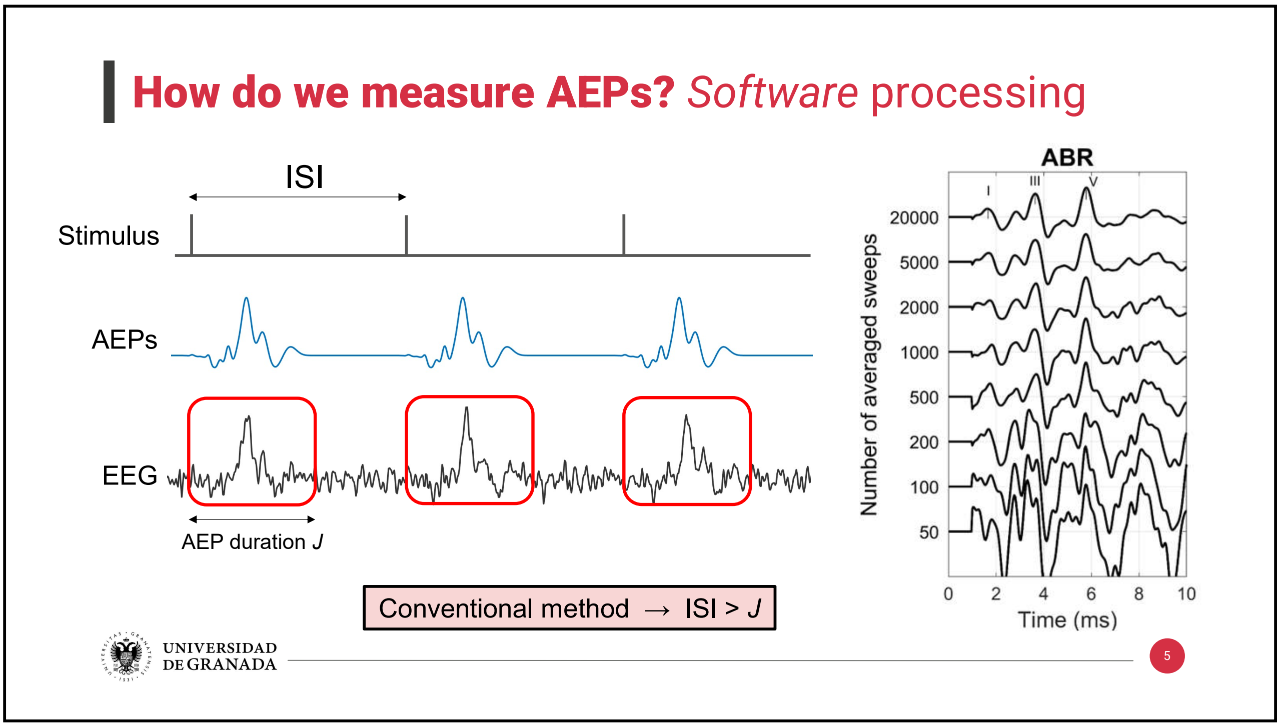
- Middle latency responses (MLRs) show activity from the medial geniculate body and the primary auditory cortex.

- And cortical auditory evoked potentials (CAEPs) from the primary and secondary auditory cortex.

- By analysing these signals we can (1) understand our neural structures better and (2) we can also evaluate someone´s hearing objectively.



- This figure shows a schematic of a simple recording system of 1 EEG channel.
- A soundcard is connected to a laptop, through which a stimulus signal is delivered to an earphone, which could be an insert earphone or a speaker.
- The stimulus evokes a neural response, which is recorded by electrodes placed on the head. This signal is amplified and sent back to the computer through one input line of the soundcard.
- In addition, there is also a synchronization signal that is sent synchronously with the stimulus to determine the time instants in which the stimuli were sent.



- From the synchronization signal we obtain the time instants at which the stimuli were presented. We call Inter-Stimulus Interval (ISI) to the separation between stimuli.

- It is common to assume that all stimuli evoke the same AEP.

- However, the EEG not only records the neural response, but also a large amount of noise from different sources.

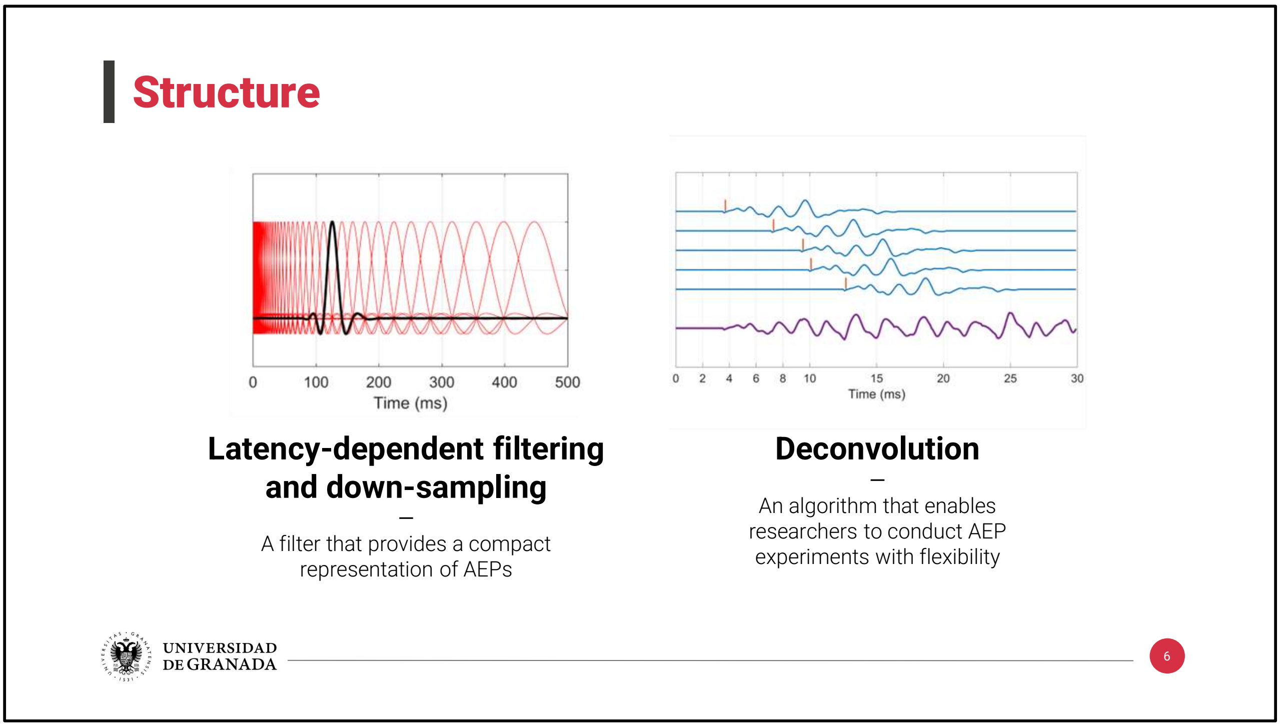
- In this example, the EEG is illustrated just with a bit of noise, but a real recorded EEG is much more contaminated by noise.

- The conventional method to improve the quality of the response consists of averaging the EEG sections that contain the AEPs (red segments).

- The more averaged responses, the higher the quality of the response.

- We can observe in the right figure that averaging only 50 segments leads to a low-quality estimation of the response, and that the quality of the response increases as the number of averaged segments increases.

- At 20,000 averaged segments the quality is high, but this requires a long test time.
- Importantly, the conventional method uses ISIs larger than the duration of the AEP (J) to avoid responses to be contaminated by adjacent responses.

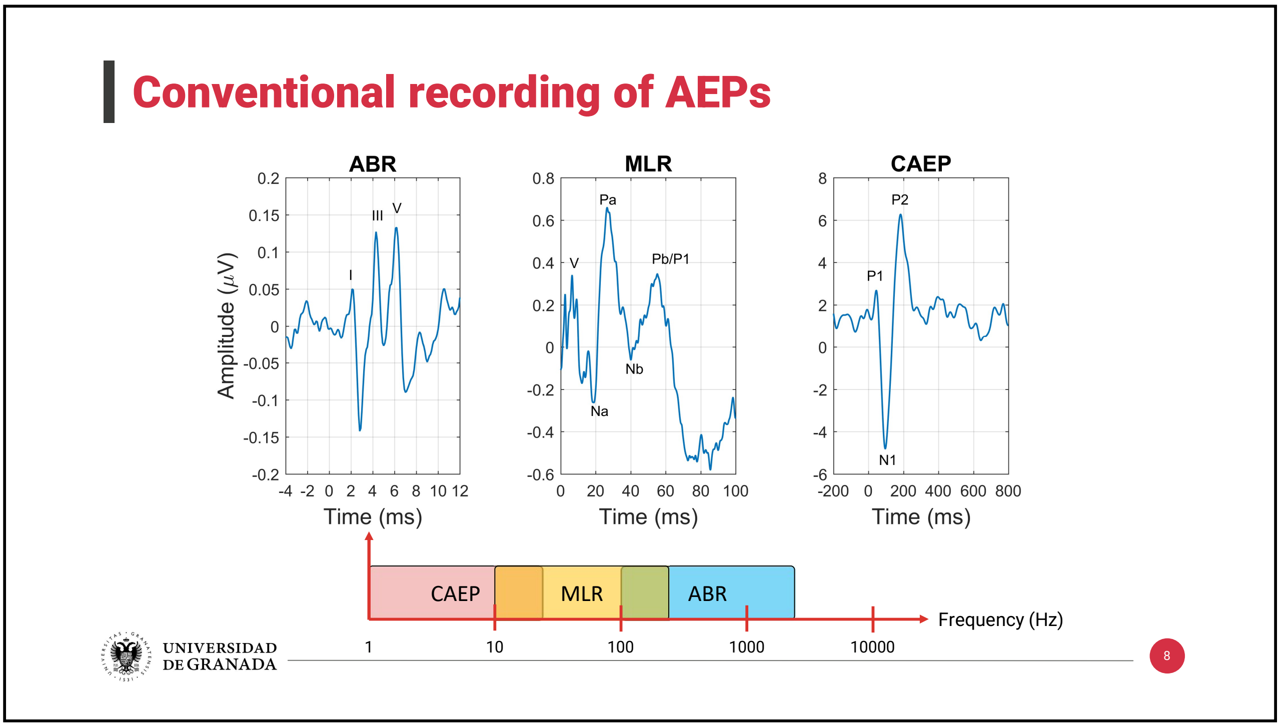


- In this context, I will present two methods developed by our research team that expand clinical and research possibilities in the field of AEPs.

1. A filter that provides a compact representation of AEPs.

2. An algorithm that allows deconvolution of overlapping AEPs, which increases flexibility in the design of experiments.

Latency-dependent filtering & Down-sampling



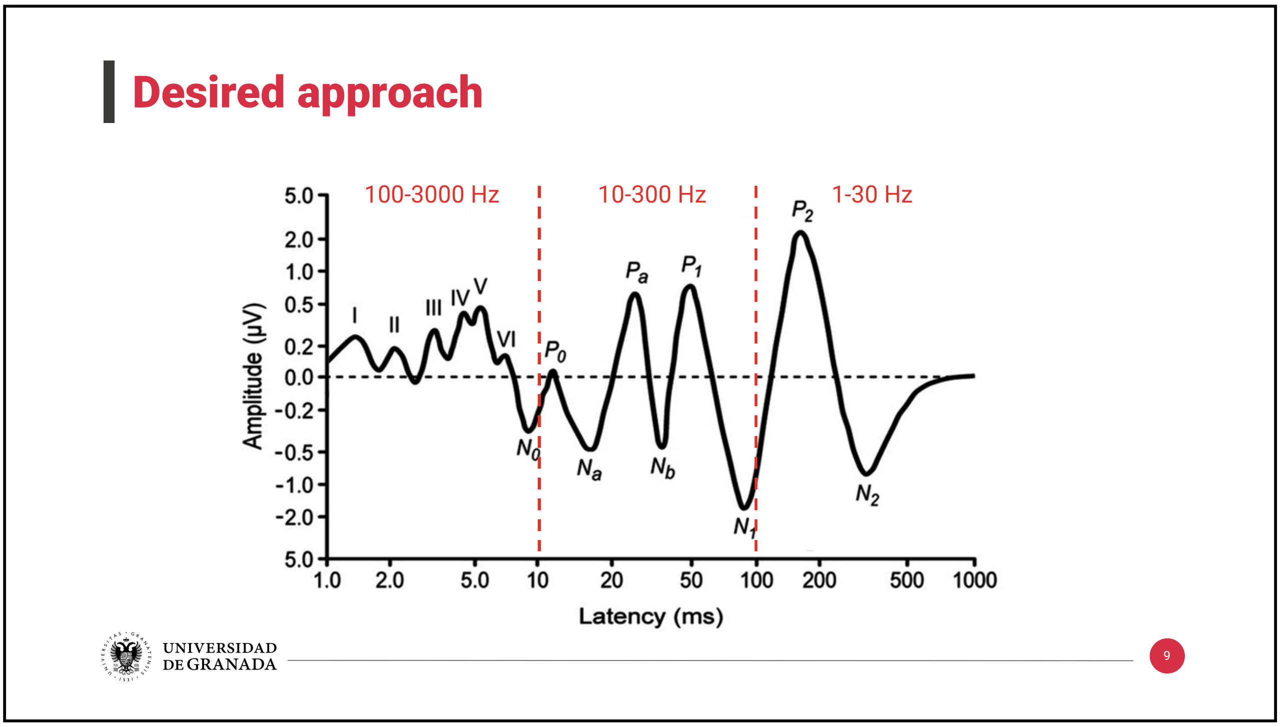
- The neurophysiological activity along the ascending auditory pathway is conventionally recorded via ABR, MLR and CAEP.

- ABRs appear within the first 10 ms from the stimulus onset. Their energy is in between 100 and 3000 Hz.

- MLRs have a duration of 100 ms, and their energy is between 10 and 300 Hz.

- CAEPs present a typical duration of about 400 ms, and their energy is in between 1 and 30 Hz.

- Since ABRs, MLRs and CAEPs present energy in different frequency bands, they are conventionally recorded as separate responses.

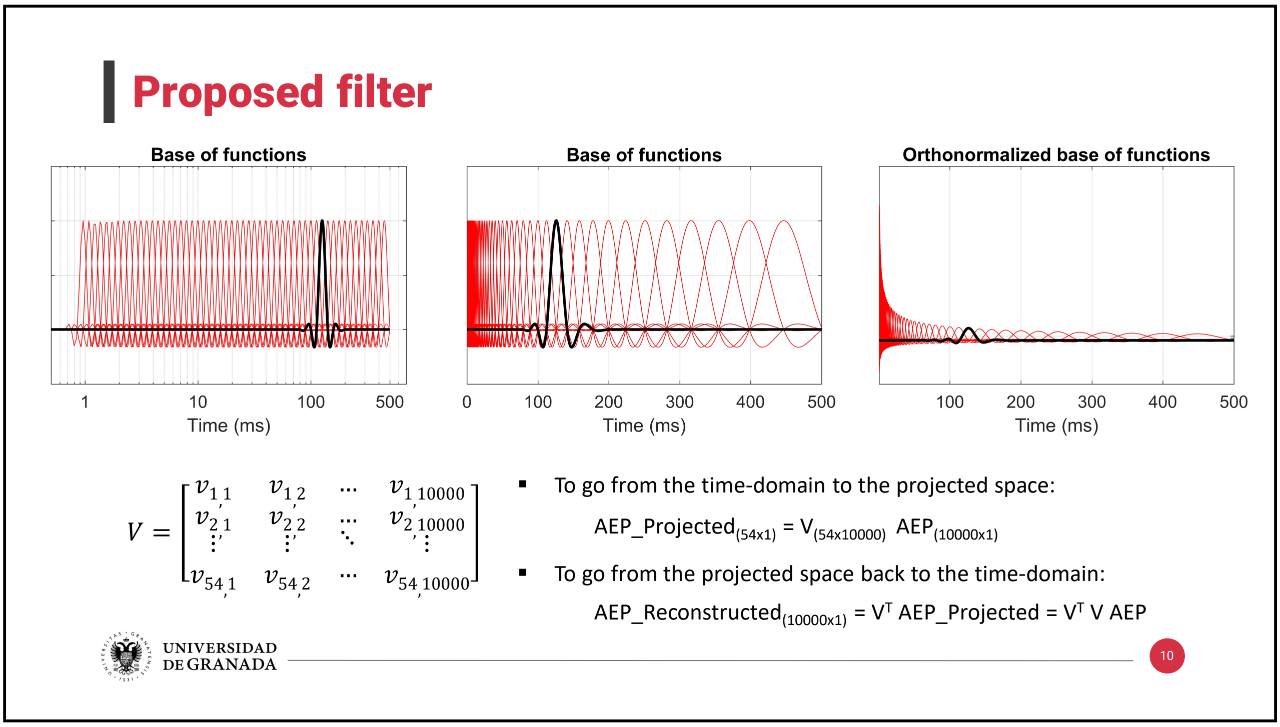


- However, it would be desired to represent all the components in the same figure.

- 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*.



- To achieve latency-dependent filtering, we have built a base of functions that are uniformly distributed in the LOGARITHMIC time-scale.

- Please note that here that we can define the number of functions per decade. In this example, there are 20 functions per decade, covering from 1 to 500 ms – thus there are 54 functions.

- The figure in the center presents the functions in the linear-time scale. We observe that earlier components are represented with more functions than later components.

- Finally, we normalize the base of functions in a way that all functions are orthonormal vectors (i.e. their scalar product is 0), and also, their module is 1.

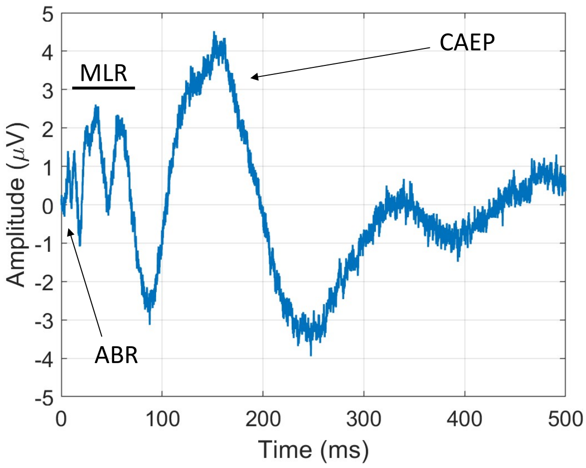
- We organise our base of functions in a matrix that, in this example, has 54 rows (one row for each function) and 10,000 columns (as many samples as the AEP – in this case this is representing 500 ms of an AEP sampled at 20 kHz)

- By applying matrix processing, we can project an AEP represented in the time domain in the projected space by applying the V matrix over the AEP.

- Importantly, it should be noted that the AEP is represented in this space with only 54 coefficients, rather than 10,000 samples as in the time domain. There is an important dimensionality reduction.

- Once projected, we can apply V transposed the projected AEP to represent this signal back in the time domain – this will be the *Reconstructed* (or filtered) AEP.

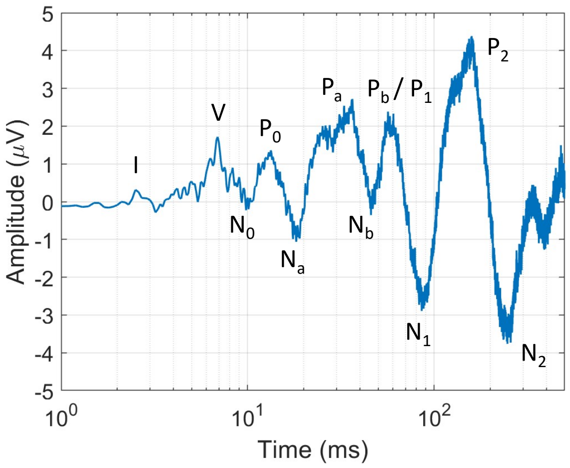
AEP example



UNIVERSIDAD DE GRANADA

- Let´s see this process with an example.
- This signal represents 500 ms of an AEP where ABR, MLR and CAEP components can be identified.

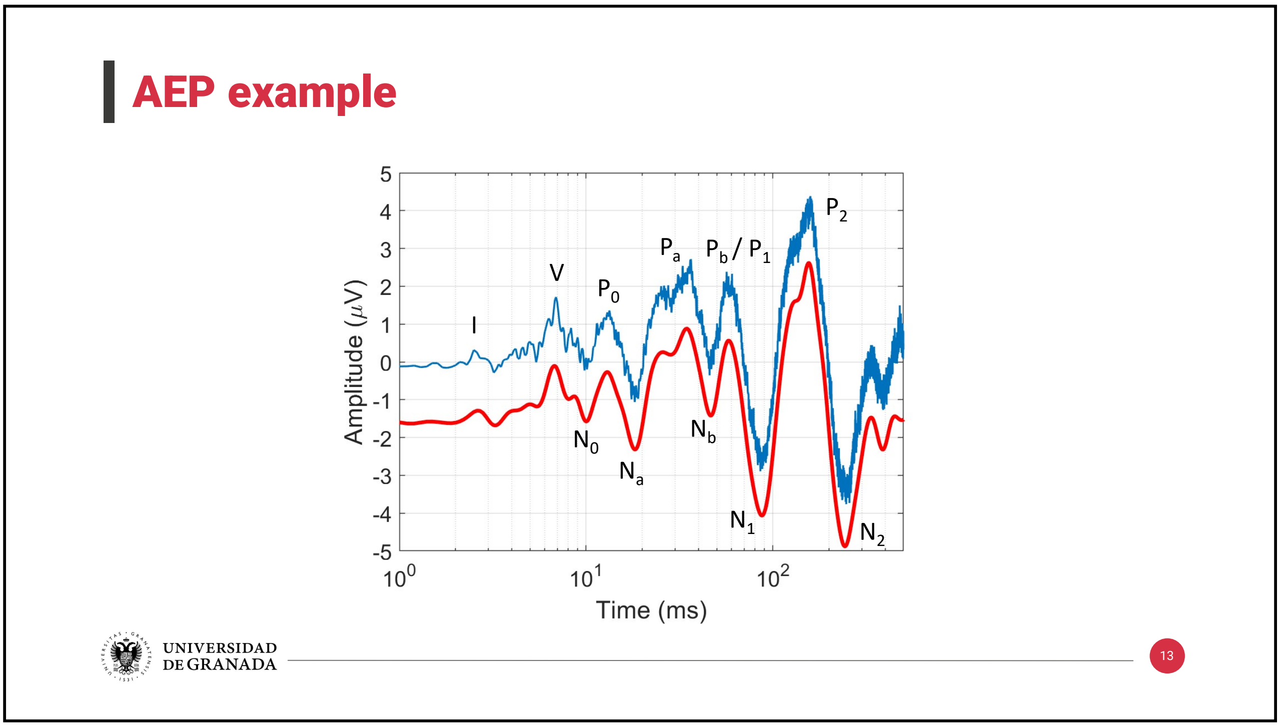
AEP example



- This is the same representation of that signal, but this time, in the logarithmic time scale.

- We see that representing AEPs in the logarithmic time scale facilitates the identification of the components, but …

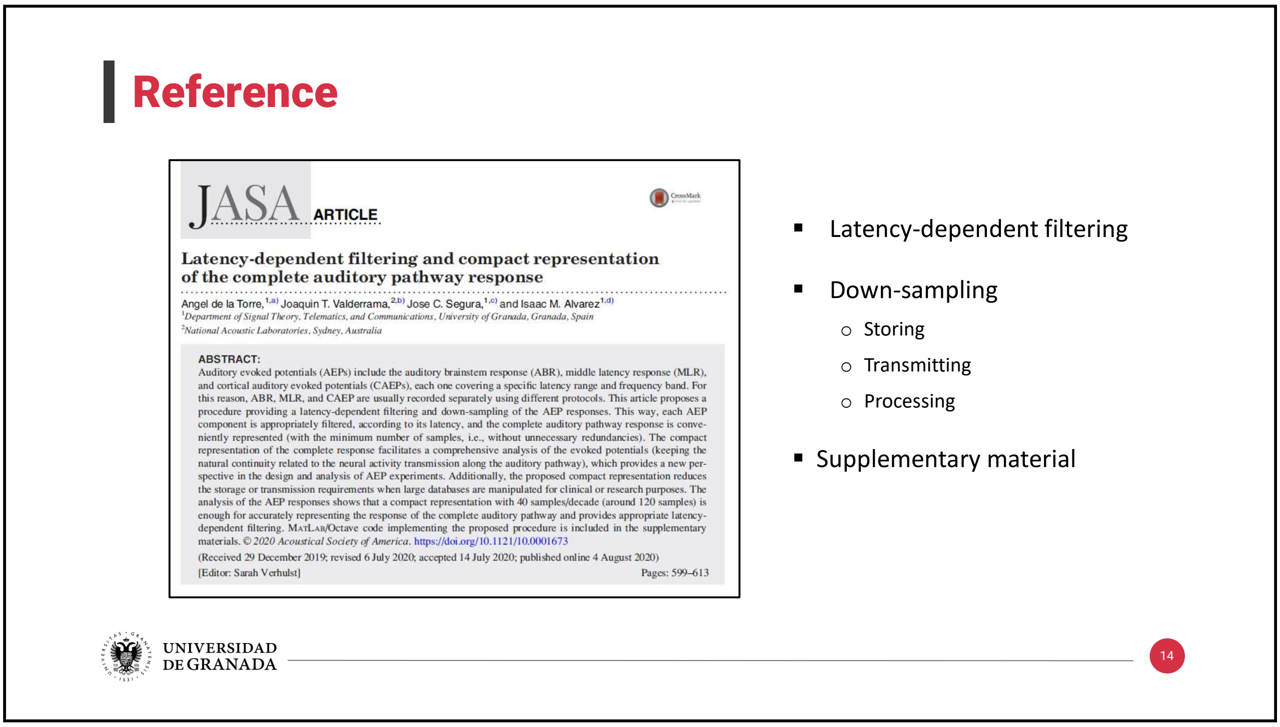
- … it also leads to high-frequency noise, particularly in the longer latencies.



- By projecting this signal in the reduced space and projecting back to the time domain, we achieve a latency-dependent filtering that is adequate to represent all components of the auditory pathway in the same figure, from wave I in the cochlea to cortical components.

- This novel representation removes the traditional discontinuities between peripherical, middle and central components, and in our view, is the natural way of representing AEPs.

- In fact, since we developed this algorithm in 2019, this is the way we represent AEPs in our studies.



- In addition to providing *latency-dependent filtering*, we mentioned earlier that this algorithm also provides *DOWN-SAMPLING.*

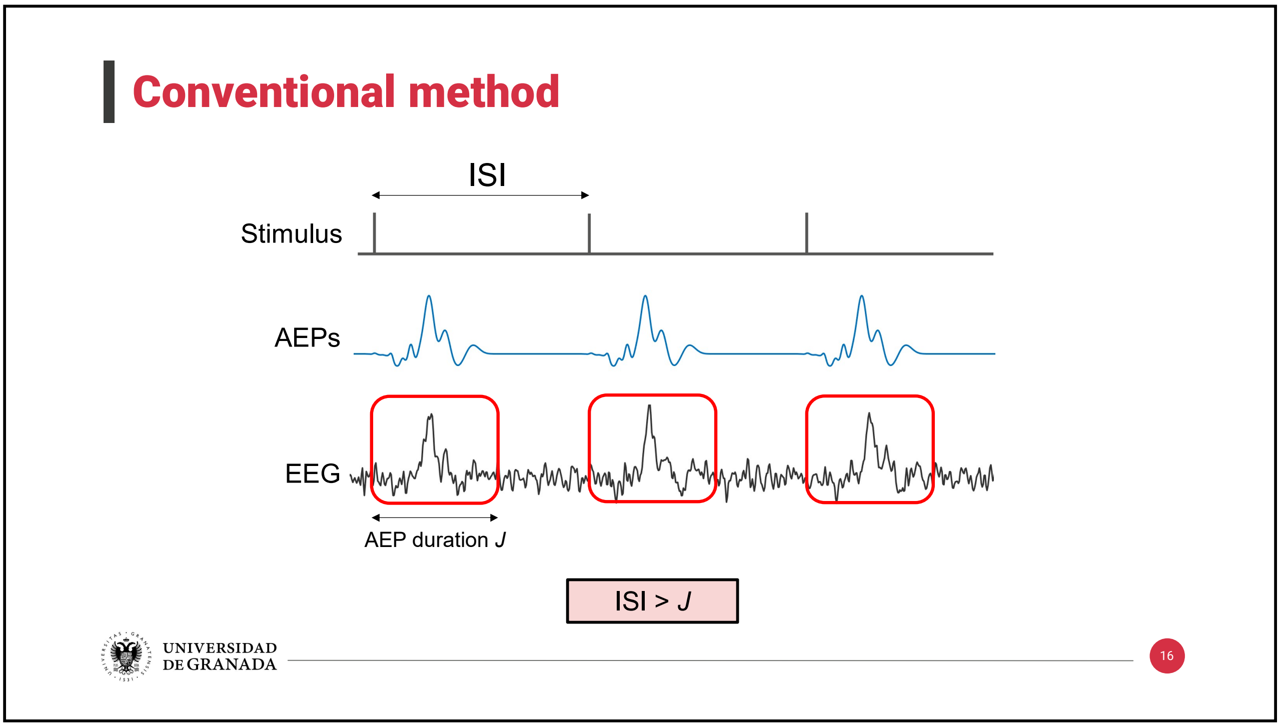
- When we project the AEP on the transformed domain, we are able to represent the AEP with only a few coefficients without losing information – in our example, 54 coefficients rather than 10,000 coefficient as in the time domain.

- Representing AEPs with fewer coefficients has important implications. For example, for storing a database of AEPs (it´s not the same to store AEPs of 10,000 coefficients than AEPs of 54 coefficients), for transmitting AEPs, and for processing AEPs such as for automatic classification.

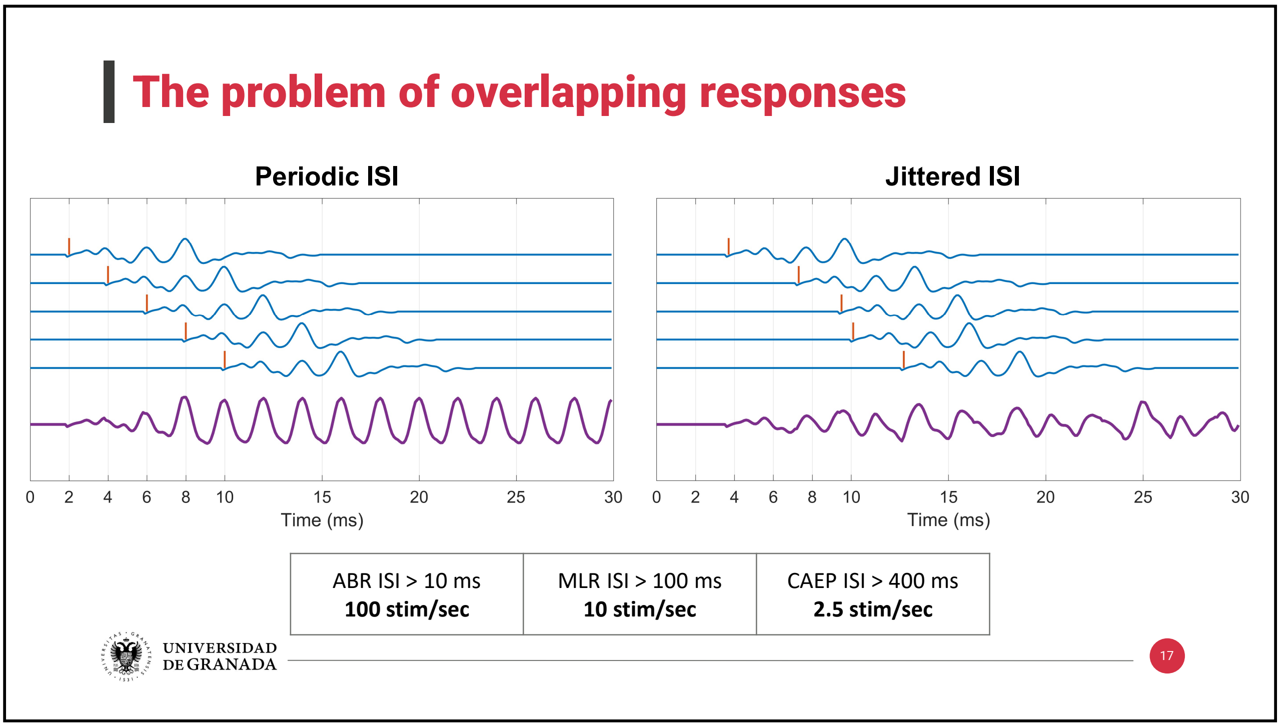
- The mathematical algorithm is described in detail in this publication, which also includes extensive supplementary material with Matlab/Octave toolboxes that run simulations and implement the methodologies. This is to facilitate the implementation of these methods by anyone interested in using them.

Deconvolution

- The second methodology is Deconvolution, which consists of a mathematical algorithm that allows the estimation of the AEP in situations where responses overlap.



- We mentioned earlier that in the conventional approach to record AEPs, the ISI must be larger than the duration of the response to avoid overlapping responses.



- This is an example of what we would get if we use a periodic ISI shorter than the duration of the response. In this case, we are using a periodic ISI of 2 ms, while the duration of the ABR is 10 ms.

- This figure shows that the signal recorded at the electrodes is a steady-state signal, from which it is not mathematically possible to estimate the original response.

- This means that there is a maximum presentation rate to avoid responses to be overlapped.

- Since ABRs present a duration of 10 ms, the ISI must be longer than that time, which leads to a maximum presentation rate of 100 stimuli per second.

- Likewise, MLRs present a duration of 100 ms, and the maximum rate would be 10 stimuli per second.

- And considering a duration of 400 ms for CAEP, their maximum rate would be 2.5 stimuli per second.

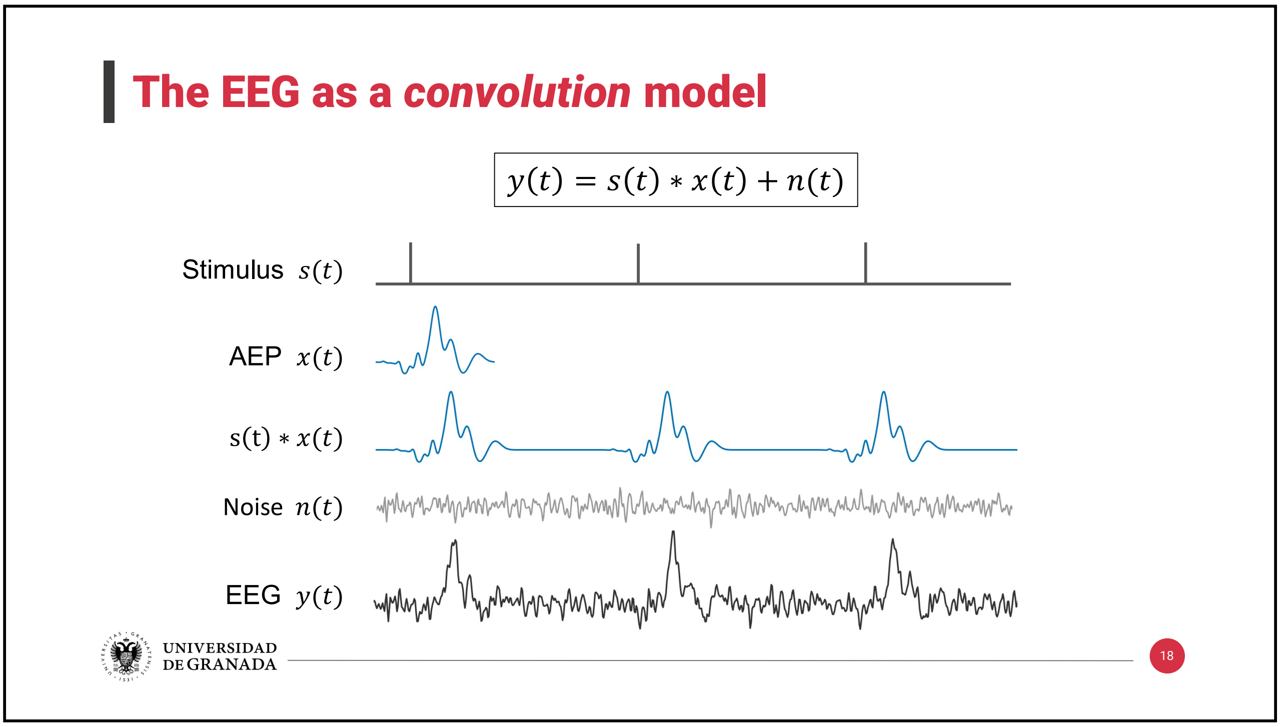
- These maximum presentation rates are an important barrier for scientists, particularly to advance knowledge in how our auditory system encodes the sounds we usually hear (speech, music, etc.).

- The way we can estimate AEPs that are overlapped is through DECONVOLUTION.

- And the first requirement to be able to deconvolve (or disentangle) responses that are overlapping is by jittering the stimulus sequence,

- such as in this example, where we can visually see that some ISIs are shorter than others.

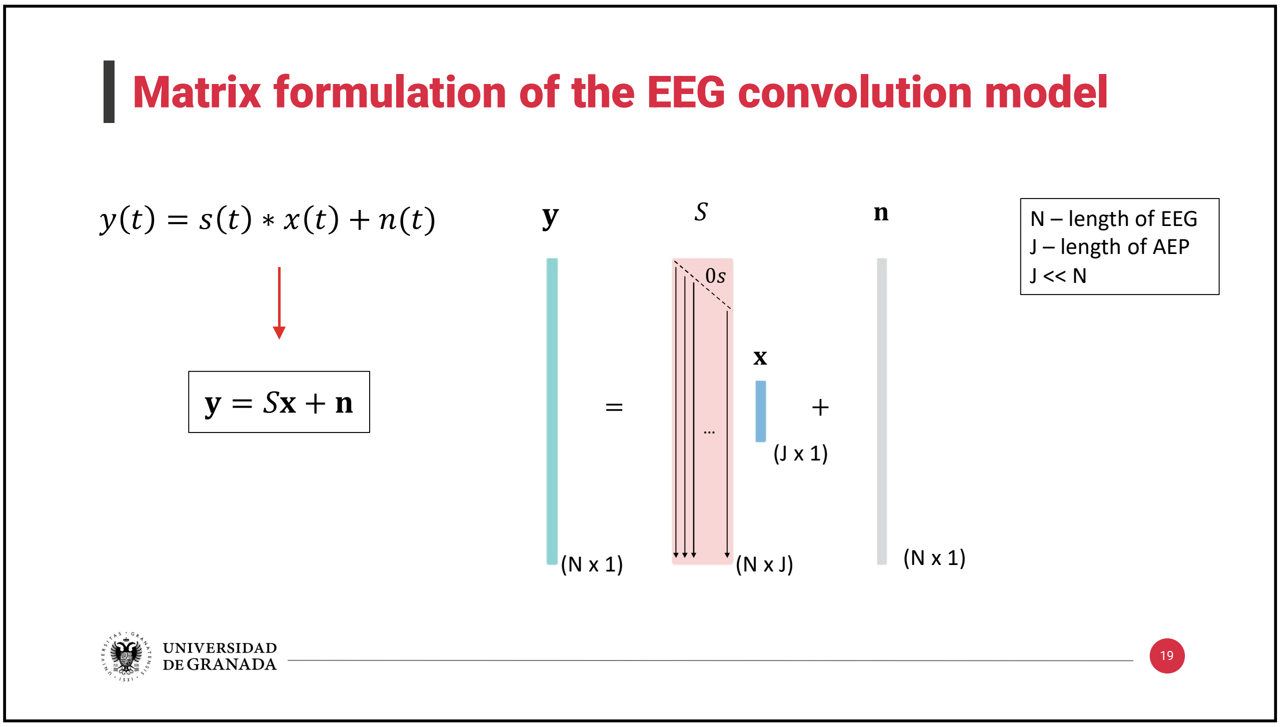
- The result of overlapping AEPs is a quasi-periodic signal, from which it is now mathematically possible to estimate the original response.



- I will present now the fundamentals of our algorithm to deconvolve overlapping AEPs.

- The EEG y(t) can be represented mathematically as the convolution of the stimulus sequence s(t) and the AEP x(t) plus noise.

- Let´s keep in mind that the signal of interest that we are interested in estimating is the AEP x(t).



- The EEG convolution model can be represented as a matrix operation.

- This way, the EEG **y** is a column vector of N samples (N being the number of samples of the EEG, several millions samples for example).

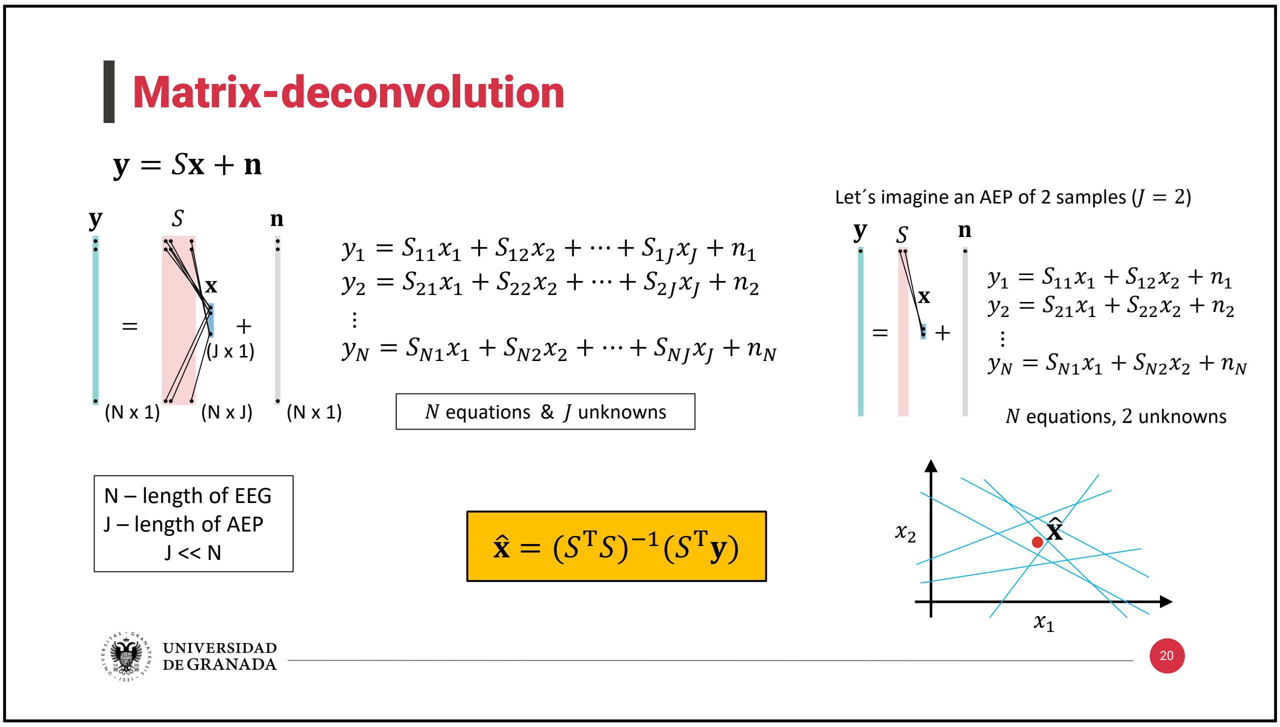
- *S***x** is the convolution operation.

- *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 the stimuli), and shifting this vector one sample every column until we complete the matrix.

- **x** is the AEP, which is a column vector of J samples.

- And **n** represents the noise, and has the same size as the EEG.



- 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 **x** vector plus the first element of the noise vector, 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). This is an over-defined system of equations.

- Since the AEP has several dimensions, it is difficult to visualize the solutions of this system of equations, but let´s imagine that our AEP only has 2 samples (x1 and x2).

- This way we would have (again) *N* equations, but this time, only 2 unknowns.

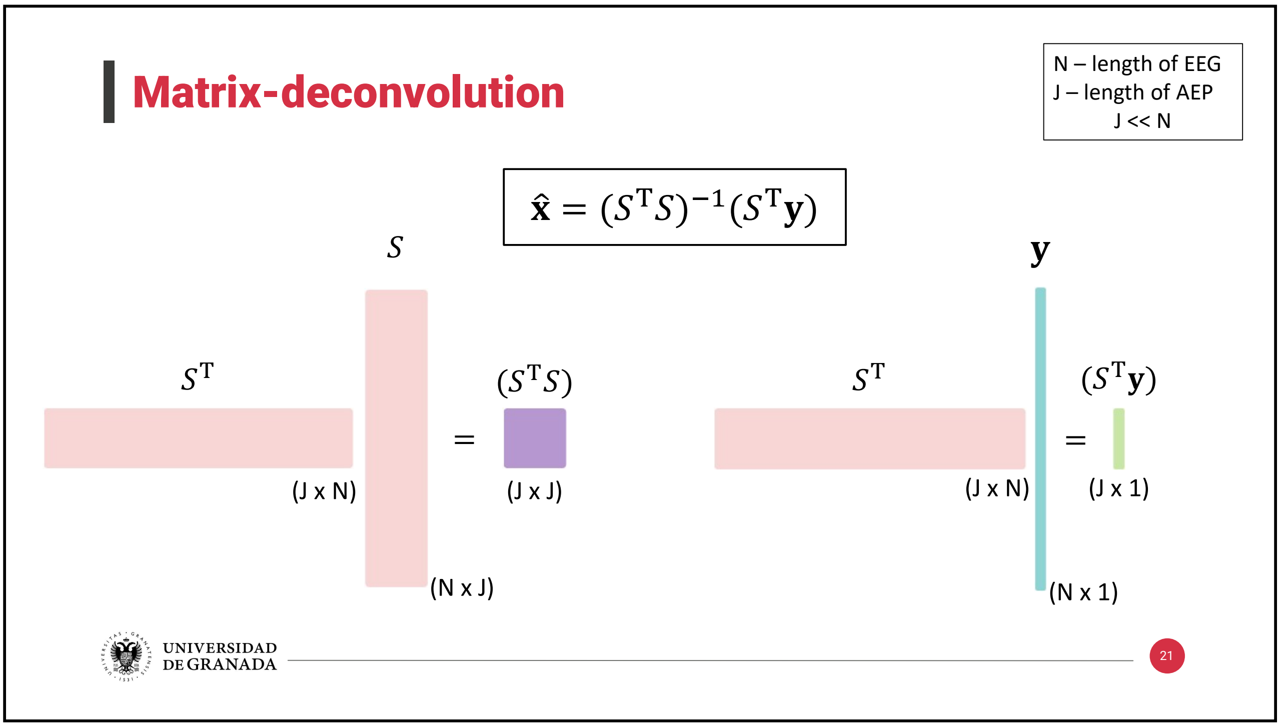
- It is now easier to visualize that each equation would lead to a line in the 2-dimensional space.

- We can also observe that this is an over-dimensioned system of equations, and that there is not a single solution.

- However, there is one unique solution that minimizes the error – that is the least-squares solution,

- And 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)

- This is the deconvolved AEP



- 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 square matrix J x J (being J the length of the AEP, much smaller than N)

- And S^{T} applied to *S* leads **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. This signal has the same dimension as the AEP **x.**

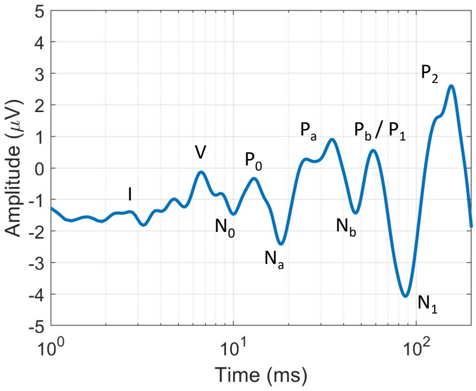
- This means that to estimate the deconvolved response, first we need to invert the $(S^{\mathrm{T}}S)$ square matrix, and multiply it by the $(S^{\mathrm{T}}\mathbf{y})$ vector

- It should be noted that when there is no overlapping responses, the $(S^{\mathrm{T}}S)$ 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.

- When responses overlap, the $(S^{\mathrm{T}}S)$ square will not be an identity matrix, and we will have to invert that matrix to deconvolve the AEP.

Matrix-deconvolution

$$\hat{\mathbf{x}} = (S^{\mathrm{T}}S)^{-1}(S^{\mathrm{T}}\mathbf{y})$$



- AEP of 200 ms @ 16,384 Hz
- J = 3,277 samples → $(S^{\mathrm{T}}S)_{(3277 \times 3277)}$
- How long does the matrix division take?
- 9 seconds

UNIVERSIDAD DE GRANADA

22

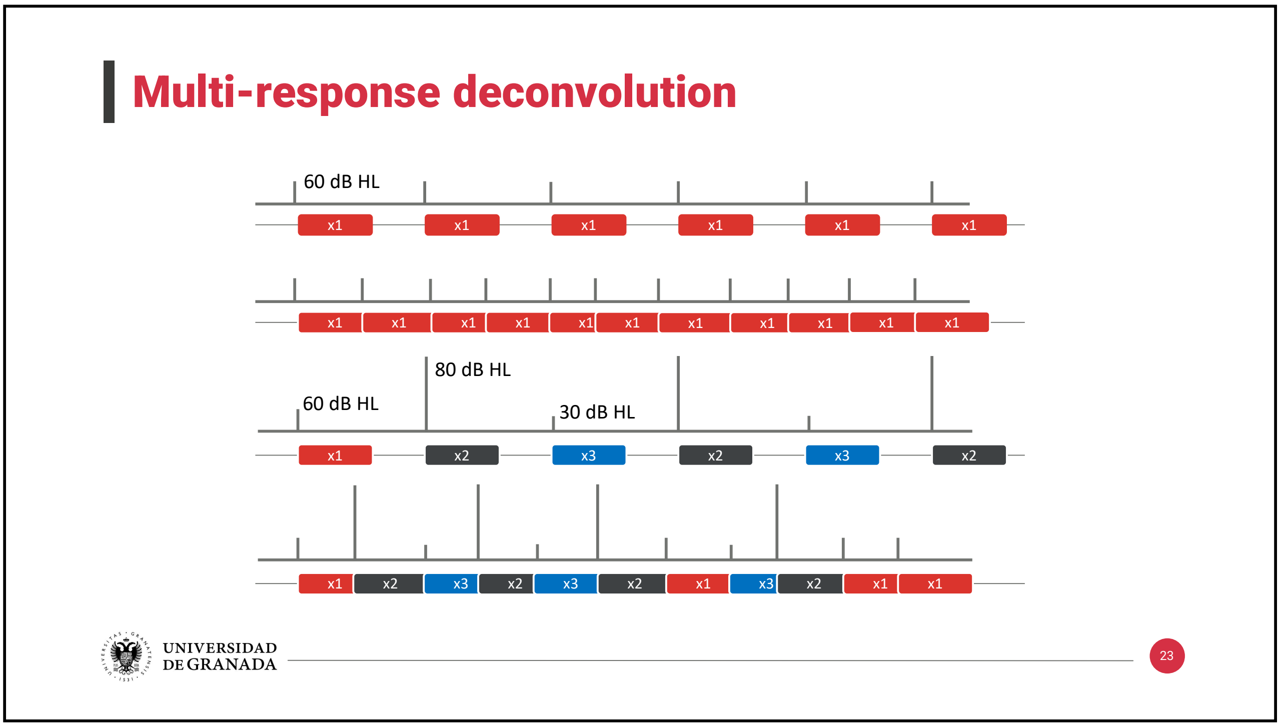
- Let´s see how this process works with an AEP signal of 200 ms duration, sampled at 16384 Hz.

- This AEP has 3,277 samples.

- Considering that the matrix $(S^{\mathrm{T}}S)$ has a dimension 3,277 x 3,277, how long would it take to execute the matrix division?

- The complexity of a matrix division increases with the size of the matrix.

- For this example, using a personal computer, it takes around 9 seconds to invert, which is a feasible processing time in most applications.



- For now we have considered only one type of stimulus that evokes one response (x1).

- And we have described the deconvolution process to estimate x1 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 x1, x2 and x3 when they overlap – this is called *multi-response deconvolution*.

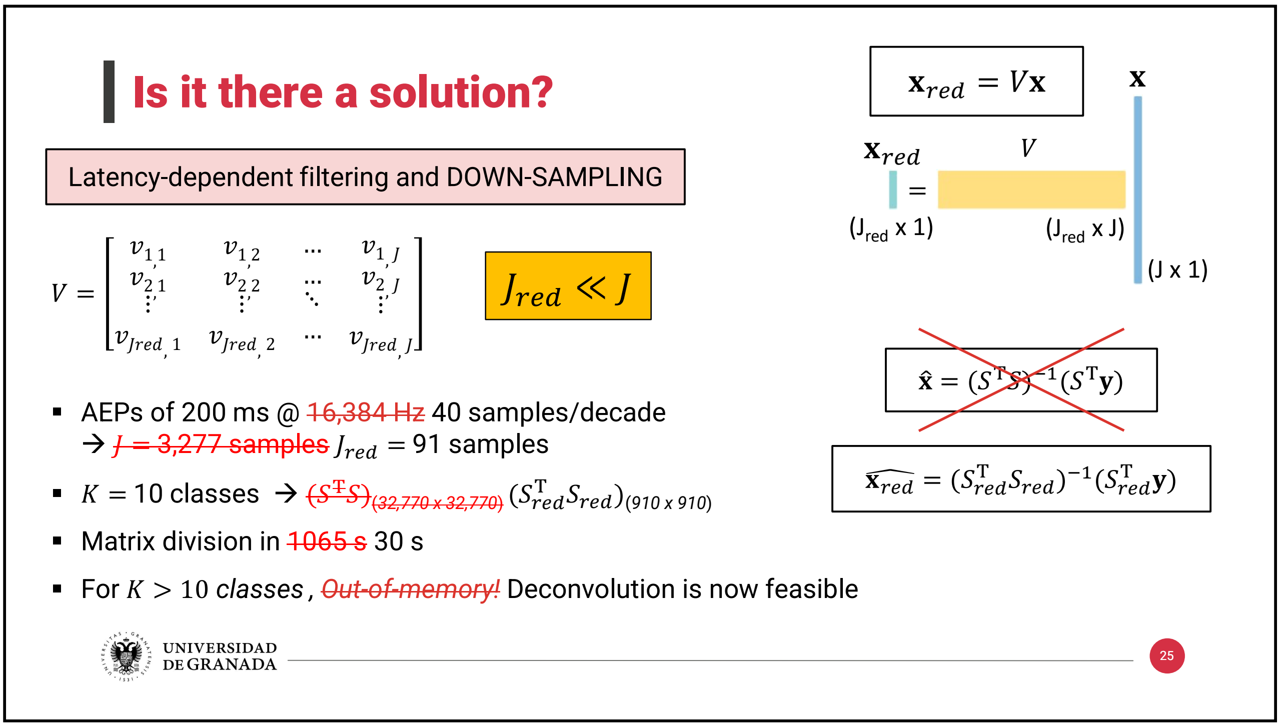
Multi-response deconvolution

$$\hat{\mathbf{x}} = (S^{\mathrm{T}}S)^{-1}(S^{\mathrm{T}}\mathbf{y})$$

- AEPs of 200 ms @ 16,384 Hz → $J = 3{,}277$ samples
- $K = 10$ classes → $(S^{\mathrm{T}}S)_{(32{,}770 \text{ x } 32{,}770)}$
- $(S^{\mathrm{T}}S)_{(32{,}770 \text{ x } 32{,}770)}$ → 1,073,872,900 numbers * 8 bytes → 8,6 GB
- Matrix division in 1065 s
- For $K > 10$ classes, *Out-of-memory!*

UNIVERSIDAD DE GRANADA

- In our previous example of an AEP of 3277 samples, if we had 10 different classes, the size of $(S^{\mathrm{T}}S)$ would increase to 32,770 x 32,770
- Please note that a matrix of this size would have around 1 million 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 1000 seconds.
- And importantly, for more than 10 classes, the memory requirements to perform deconvolution are not manageable for a personal computer.



- Is it there any possible way to accelerate the deconvolution process and overcome the memory-requirements limitation for large number of classes?

- The solution comes by applying latency-dependent filtering and DOWN-SAMPLING.

- This way, the AEP is no longer represented with *J* samples, but with $J_{reduced}$

- This can be visually observed by applying the *V* matrix to the AEP **x**, which leads to a reduced representation of the AEP.

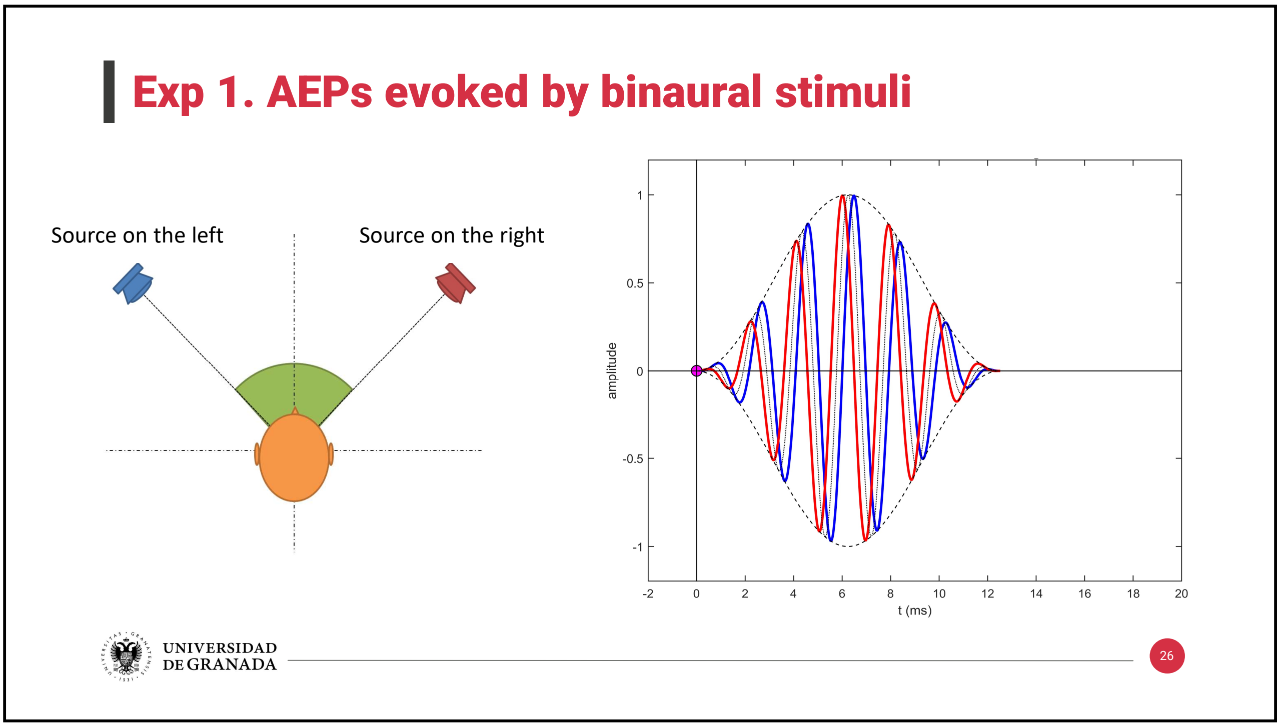
- As a result of this dimensionality reduction, to estimate the AEP via matrix deconvolution we no longer need to invert a matrix J x J, but a matrix Jred x Jred.

- In our previous example, AEPs sampled at 40 samples/decade now have 91 samples, and for 10 classes, the size of $(S_{red}^T S_{red})$ would be 910 x 910.

- Now the matrix division could be performed in just 30 seconds.

- And for a large number of classes, deconvolution is feasible.

- Let´s now see some examples where multi-response deconvolution has been useful.



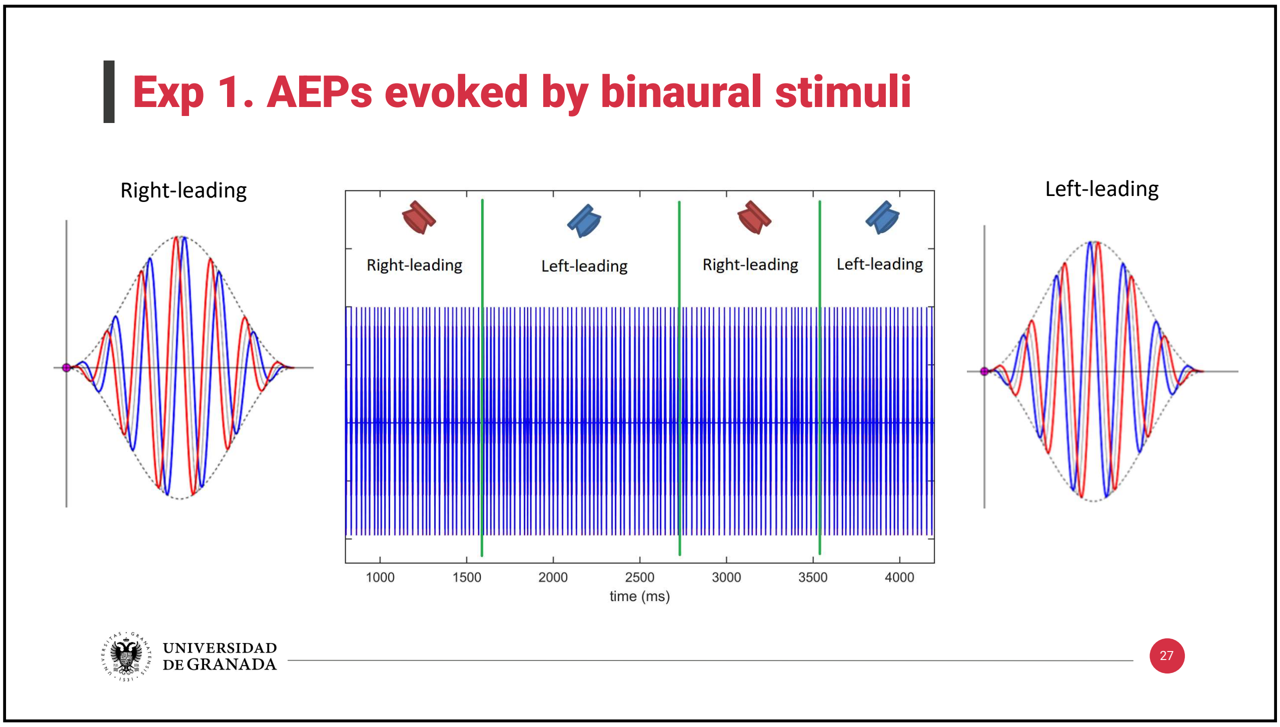
- In a different experiment we recreated an acoustic scenario in which the position of a sound source moved from one side of the head to the other.

- To do this, we used a stereo audio stimulus, delivered to the subject via insert earphones.

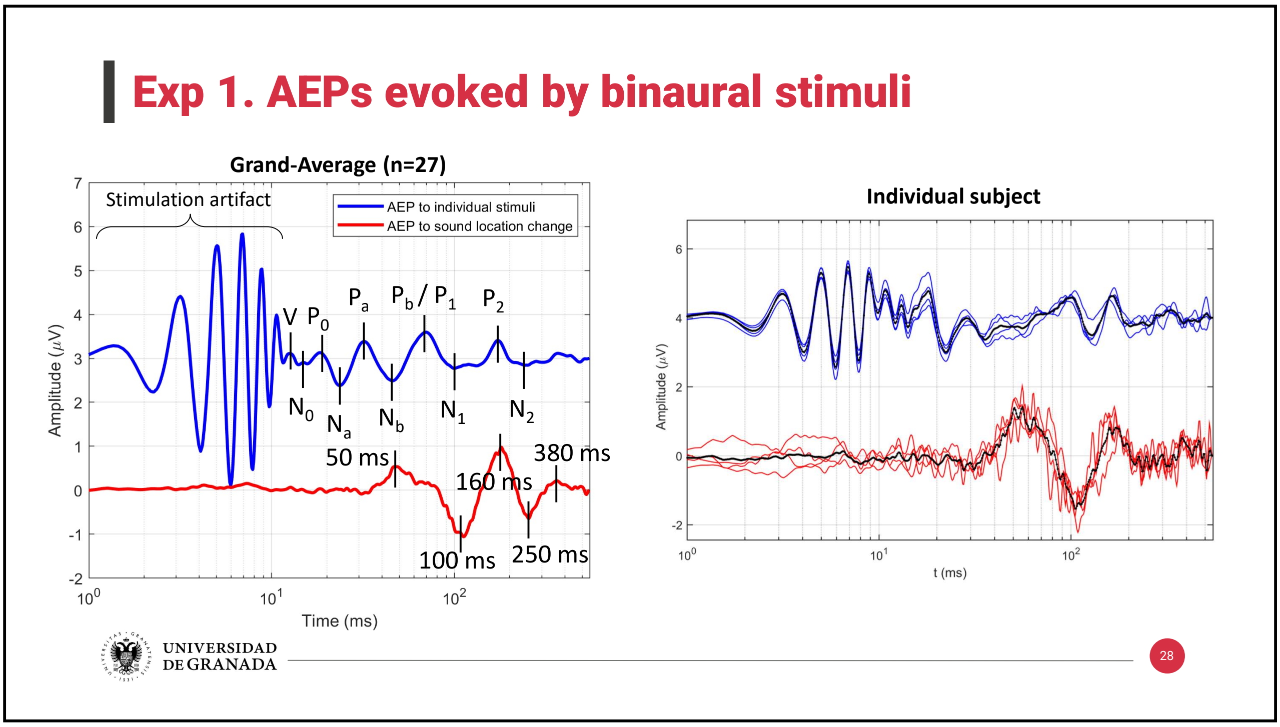
- The stimulus consisted of a windowed tone of 500 Hz.

- We varied the phase of the left and right stimulus signals to induce an interaural time difference which allowed participants perceive the sound coming from their left or from their right side.

- In this example, we observe that the stimulus of the right ear (in red) arrives the participant´s ear sooner than the left stimulus (in blue), which is perceived as if the sound came from their right side.



- The stimulus sequence consisted of several repetitions of the binaural windowed tones.
- Randomly between 1 and 2 seconds there was a change on the source location, which is represented by the vertical green lines and the coloured speaker icons.
- This way, during 1 to 2 seconds participants perceived that the sound came from their right, then from their left, then from their right again, and so on.
- We hypothesized that each windowed tone would evoke a neurophysiological response, and that each change of sound location would evoke an additional response.
- We used multi-response deconvolution and latency-dependent filtering to obtain 2 AEPs: (1) an AEPs to each individual burst & (2) an additional AEP associated with the change of the sound location



- The left figure presents the grand-average responses across the 27 participants that we tested.

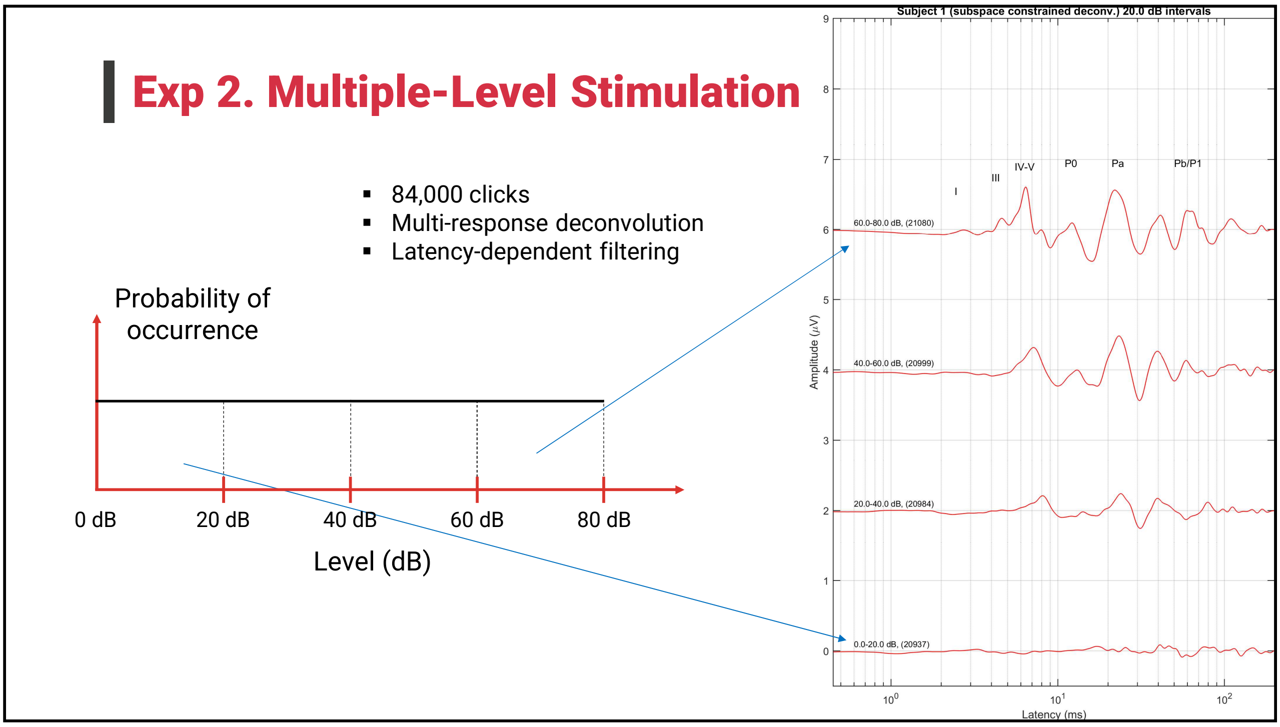
- The blue signal shows the neurophysiological response to each individual stimulus.

- In addition to the stimulus artifact, which has the same morphology as the stimulus, this signal presents components consistent with conventional ABR, MLR and CAEP components.

- The red signal represents the AEP elicited by the change of sound location. Interestingly, due to the long latencies of the components of this signal, we can conclude that the binaural stimuli used in this experiment evoke a series of cortical components with latencies ranging from 50 to almost 400 ms.

- This experiment is another example of the flexibility provided by deconvolution, and the comprehensive analysis of AEPs enabled by the latency-dependent filtering.

- It´s noteworthy mentioning that these components could be obtained both at group level (as shown in the grand-average figure) and also at individual level.

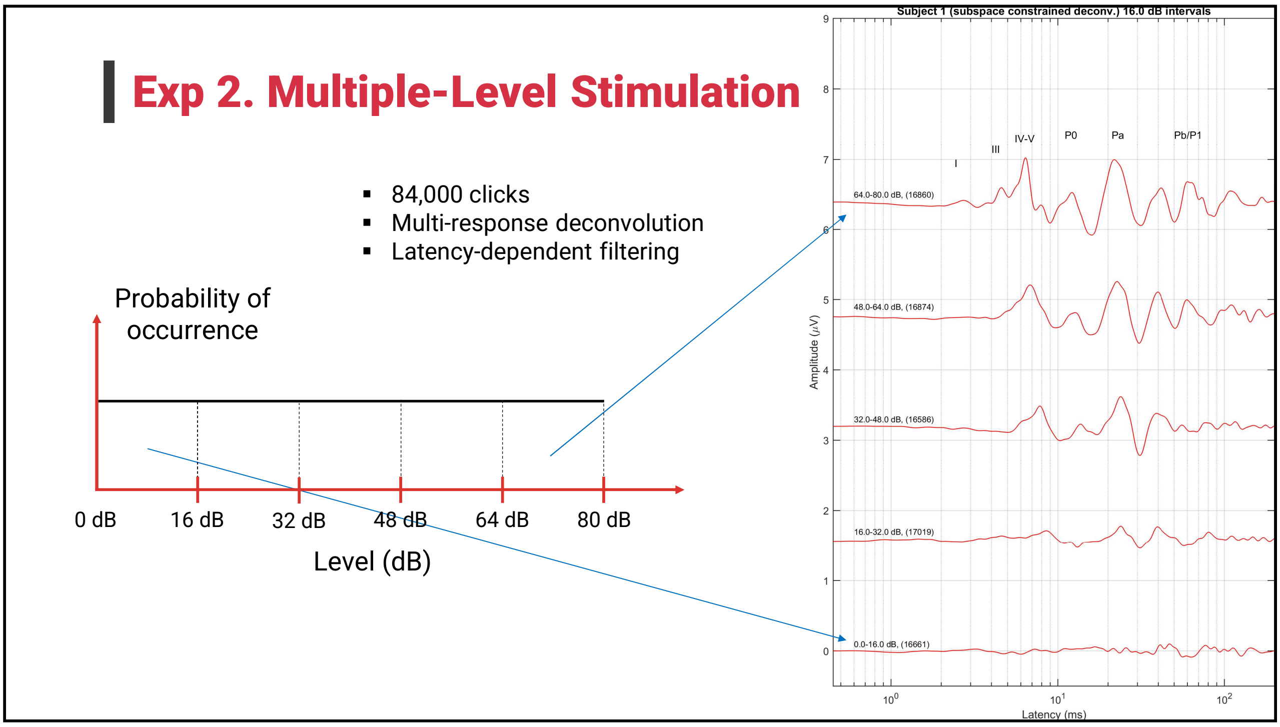


- In another experiment, we designed a stimulus sequence consisting of clicks, in which their level varied linearly as shown in this diagram.

- This multiple-level stimulation enabled a flexible categorization of events based on level.

- The figure on the right presents AEPs from one participant obtained by categorizing click events in 4 groups of 20 dB.

- To obtain these signals, we used multi-response deconvolution considering 4 classes, and then applying latency-dependent filter to represent the full-range response in the logarithmic time scale.



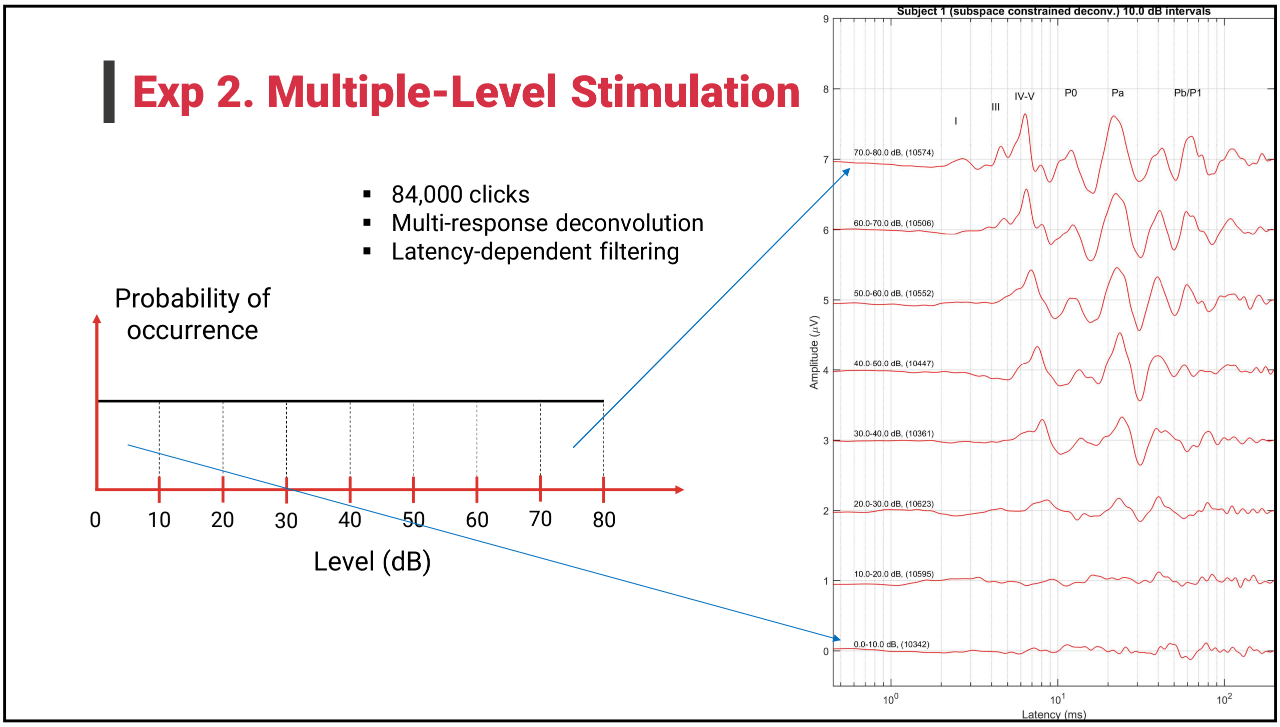
- AEPs evoked by this stimulation sequence could also be categorised in more classes.

- This time, we categorised the evoked responses in 5 classes – in groups of 16 dB.

- It should be noted that when we make categories, we assume all the responses in one class have the same morphology, and that assumption might not be met if we consider groups of large deviations in level.

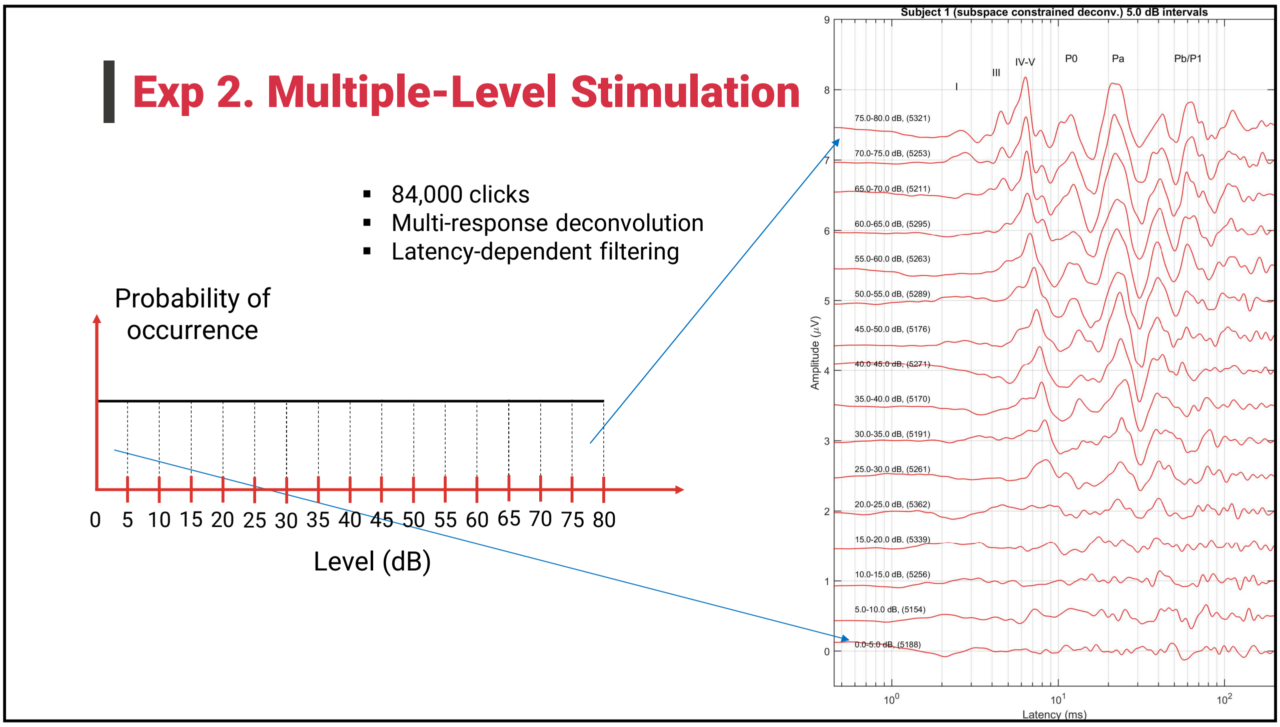
- Does this mean that the more categories the better? On one hand, yes – because we can track morphology changes with more resolution; but on the other, we shall consider the number of available responses, since making too many categories could lead to AEP estimations of poor quality if the number of responses is not sufficiently large.

- In this particular example, since we have 84,000 responses (which are a lot), we can proceed further and split the available responses in more categories.



- These are the AEPs when we make 8 categories of 10 dB range each.

- We see that components from cochlea (wave I) to cortex (P1) can be easily tracked, and that the most robust components are wave V from the ABR and the Pa of the MLR.



- The evoked potentials presented in this slide are the result of making 16 categories in groups of 5 dB, each of them obtained from around 5,000 responses.

- This figure shows that splitting the available responses in more categories on one hand leads to sharper peaks (because the time-invariant assumption is better accomplished for narrow level distributions), …

- but also the responses present lower quality resulting from the lower number of responses in each group.

- It is interesting to highlight that the constraint to make many categories in this experiment comes from an *audiological* limitation (which is the number of available responses), and not a *mathematical* limitation – as we can apply deconvolution in the reduced space to simultaneously estimate AEPs of several categories.

- These three experiments also show how useful is to represent the full-range of evoked responses in the logarithmic time scale using latency-dependent filtering, because we can provide a comprehensive audiological interpretation, considering all components of the auditory pathway, rather than isolated sections.

Matrix-based formulation of the iterative randomized stimulation and averaging method for recording evoked potentials

Angel de la Torre,[1] Joaquin T. Valderrama,[2,a)] Jose C. Segura,[1] and Isaac M. Alvarez[1]
[1]*Department of Signal Theory, Telematics, and Communications, University of Granada, Granada, Spain*
[2]*National Acoustic Laboratories, Sydney, Australia*

(Received 4 June 2019; revised 16 October 2019; accepted 15 November 2019; published online 26 December 2019)

The iterative randomized stimulation and averaging (IRSA) method was proposed for recording evoked potentials when the individual responses are overlapped. The main inconvenience of IRSA is its computational cost, associated with a large number of iterations required for recovering the evoked potentials and the computation required for each iteration [involving the whole electroencephalogram (EEG)]. This article proposes a matrix-based formulation of IRSA, which is mathematically equivalent and saves computational load (because each iteration involves just a segment with the length of the response, instead of the whole EEG). Additionally, it presents an analysis of convergence that demonstrates that IRSA converges to the least-squares (LS) deconvolution. Based on the convergence analysis, some optimizations for the IRSA algorithm are proposed. Experimental results (configured for obtaining the full-range auditory evoked potentials) show the mathematical equivalence of the different IRSA implementations and the LS-deconvolution and compare the respective computational costs of these implementations under different conditions. The proposed optimizations allow the practical use of IRSA for many clinical and research applications and provide a reduction of the computational cost, very important with respect to the conventional IRSA, and moderate with respect to the LS-deconvolution. MATLAB/Octave implementations of the different methods are provided as supplementary material. © *2019 Acoustical Society of America.*
https://doi.org/10.1121/1.5139639

[BLM] Pages: 4545–4556

JASA ARTICLE

Latency-dependent filtering and compact representation of the complete auditory pathway response

Angel de la Torre,[1,a)] Joaquin T. Valderrama,[2,b)] Jose C. Segura,[1,c)] and Isaac M. Alvarez[1,d)]
[1]*Department of Signal Theory, Telematics, and Communications, University of Granada, Granada, Spain*
[2]*National Acoustic Laboratories, Sydney, Australia*

ABSTRACT:
Auditory evoked potentials (AEPs) include the auditory brainstem response (ABR), middle latency response (MLR), and cortical auditory evoked potentials (CAEPs), each one covering a specific latency range and frequency band. For this reason, ABR, MLR, and CAEP are usually recorded separately using different protocols. This article proposes a procedure providing a latency-dependent filtering and down-sampling of the AEP responses. This way, each AEP component is appropriately filtered, according to its latency, and the complete auditory pathway response is conveniently represented (with the minimum number of samples, i.e., without unnecessary redundancies). The compact representation of the complete response facilitates a comprehensive analysis of the evoked potentials (keeping the natural continuity related to the neural activity transmission along the auditory pathway), which provides a new perspective in the design and analysis of AEP experiments. Additionally, the proposed compact representation reduces the storage or transmission requirements when large databases are manipulated for clinical or research purposes. The analysis of the AEP responses shows that a compact representation with 40 samples/decade (around 120 samples) is enough for accurately representing the response of the complete auditory pathway and provides appropriate latency-dependent filtering. MATLAB/Octave code implementing the proposed procedure is included in the supplementary materials. © *2020 Acoustical Society of America.* https://doi.org/10.1121/10.0001673
(Received 29 December 2019; revised 6 July 2020; accepted 14 July 2020; published online 4 August 2020)
[Editor: Sarah Verhulst] Pages: 599–613

- As in the previous reference, these publications also provide the mathematical formulation of the algorithms and supplementary material with Matlab and Octave toolboxes that run simulations and implement the methodologies, aimed at facilitating any interested researcher, clinician or industry to use these techniques.

- In addition to these two references, an article presenting multi-response deconvolution is currently in progress.

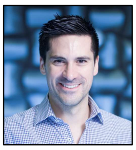
Take-home message & Acknowledgments

- The compact representation of AEPs enabled by *latency-dependent filtering and down-sampling* facilitates the audiological analysis of all the components of the ascending auditory pathway, from cochlea to cortex.
- The representation of AEPs with fewer coefficients enables multi-response deconvolution via matrix processing – a method that substantially increases flexibility in the design advanced audiological experiments.
- Publications provide extensive supplementary materials with simulations and toolboxes to help any interested scientist use and implement these methodologies.


Ángel de la Torre
UGR


José C. Segura
UGR


Isaac M. Álvarez
UGR


Joaquín T. Valderrama
UGR


José L. Vargas
UGR, SCUH

BINAURAL-EVAL / B.TIC.382.UGR20

UNIÓN EUROPEA
Fondo Europeo de Desarrollo Regional
Junta de Andalucía
Consejería de Economía, Conocimiento, Empresas y Universidad

Speech-AEPs / PID2020-119073GB-I00 / AEI / 10.13039/501100011033

MINISTERIO DE CIENCIA E INNOVACIÓN
AGENCIA ESTATAL DE INVESTIGACIÓN

María Zambrano Fellowship

Funded by the European Union
NextGenerationEU
MINISTERIO DE UNIVERSIDADES

UNIVERSIDAD DE GRANADA

34