





Multi-response deconvolution of auditory evoked potentials in a reduced representation space

Angel de la Torre,^{1,2,a)}  Inmaculada Sanchez,^{1,3} Isaac M. Alvarez,^{1,2}  Jose C. Segura,^{1,2,b)} 
 Joaquin T. Valderrama,^{1,2,4}  Nicolas Muller,⁵ and Jose L. Vargas⁵

¹Department of Signal Theory, Telematics and Communications, University of Granada, Granada, Spain

²Research Centre for Information and Communication Technologies (CITIC-UGR), University of Granada, Granada, Spain

³AVVALE España S.L., Madrid, Spain

⁴Department of Linguistics, Macquarie University, Sydney, Australia

⁵ENT Service, Hospital Universitario Clinico San Cecilio, Servicio Andaluz de Salud, Granada, Spain

ABSTRACT:

The estimation of auditory evoked potentials requires deconvolution when the duration of the responses to be recovered exceeds the inter-stimulus interval. Based on least squares deconvolution, in this article we extend the procedure to the case of a multi-response convolutional model, that is, a model in which different categories of stimulus are expected to evoke different responses. The computational cost of the multi-response deconvolution significantly increases with the number of responses to be deconvolved, which restricts its applicability in practical situations. In order to alleviate this restriction, we propose to perform the multi-response deconvolution in a reduced representation space associated with a latency-dependent filtering of auditory responses, which provides a significant dimensionality reduction. We demonstrate the practical viability of the multi-response deconvolution with auditory responses evoked by clicks presented at different levels and categorized according to their stimulation level. The multi-response deconvolution applied in a reduced representation space provides the least squares estimation of the responses with a reasonable computational load. matlab/Octave code implementing the proposed procedure is included as supplementary material.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1121/10.0026228>

(Received 5 July 2023; revised 3 April 2024; accepted 16 May 2024; published online 5 June 2024)

[Editor: Colleen G. Le Prell]

Pages: 3639–3653

I. INTRODUCTION

Auditory Evoked Potentials (AEPs) are commonly applied to study the auditory system in the context of both research and clinical perspectives (Burkard and Don, 2007). Due to the low amplitude of AEP responses and the presence of background noise, the estimation of evoked responses from the electroencephalogram (EEG) usually includes a repetitive presentation of stimuli and the synchronous averaging of epochs (Thornton, 2007).

Synchronous averaging requires the inter-stimulus interval (ISI) to be longer than the response duration in order to avoid overlapping adjacent responses. The conventional protocols for AEP recording suggest appropriate configuration of the EEG filtering, response duration, and maximum stimulation rate specific for the estimation of auditory brainstem responses (ABR), middle latency responses (MLR), or cortical auditory evoked potentials (CAEPs) in order to avoid overlapping of responses (Hall, 2007).

If the response duration is longer than the ISI (because of the high stimulation rate) a deconvolution-based

estimation (instead of synchronous averaging) is required to obtain the responses (Bohorquez and Ozdamar, 2006; Eysholdt and Schreiner, 1982; Valderrama *et al.*, 2014b). Deconvolution methods for recovering AEP responses include maximum length sequences (MLS) (Eysholdt and Schreiner, 1982; Thornton and Slaven, 1993), adjacent-responses (ADJAR) (Woldorff, 1993), quasi-periodic sequence deconvolution (QSD) (Jewett *et al.*, 2004), continuous loop averaging deconvolution (CLAD) (Bohorquez and Ozdamar, 2006; Ozdamar and Bohorquez, 2006), linear deconvolution for baseline correction (LDBC) (Lütkenhöner, 2010), randomized stimulation and averaging (RSA) (Valderrama *et al.*, 2012), iterative randomized stimulation and averaging (IRSA) (de la Torre *et al.*, 2019; Valderrama *et al.*, 2014b; Valderrama *et al.*, 2016) and randomized stimulation with least squares deconvolution (RSLSD) (Bardy *et al.*, 2014a; Bardy *et al.*, 2014b; Bardy *et al.*, 2014c; de la Torre *et al.*, 2019). Alternatively, the use of uncorrelated stimulation sequences (i.e., with ISI following a Poisson distribution) allows to deconvolve the responses by synchronous average (Maddox and Lee, 2018; Polonenko and Maddox, 2019, 2022; Stoll and Maddox, 2023). However, for specific experimental configurations

^{a)}Email: atv@ugr.es

^{b)}Deceased.

and ISI distributions, and because of short- and long-term adaptation mechanisms (Valderrama *et al.*, 2014a; Valderrama *et al.*, 2016), the responses associated with different ISI values could be significantly different, which would lead to an inconsistent convolutional model and to an inaccurate estimation of the responses. For this reason, the use of a Poisson ISI distribution (with a large spread in the ISI values) makes it recommendable to verify that the magnitude of the ISI effect is moderate (as analyzed in Maddox and Lee, 2018; Polonenko and Maddox, 2019, 2022).

While some deconvolution methods impose constraints to the stimulation pattern (like MLS or CLAD), others (like IRSA and RSLSD) provide much more flexibility in the stimulus design, because they only require the autocorrelation matrix of the stimulation sequence to be invertible (Bardy *et al.*, 2014a; Bardy *et al.*, 2014c; de la Torre *et al.*, 2019; de la Torre *et al.*, 2022; Valderrama *et al.*, 2014a; Valderrama *et al.*, 2016). Thanks to this flexibility, audiological experiments with more ecologically-valid stimuli (i.e., more closely related to natural stimulation) are possible (Burkard *et al.*, 2018; de la Torre *et al.*, 2019; Finneran *et al.*, 2019; Martinez *et al.*, 2021; Valderrama *et al.*, 2014b; Valderrama *et al.*, 2019; Valderrama *et al.*, 2016). Deconvolution-based AEP estimation methods allow the study of neural adaptation mechanisms (Gillespie and Muller, 2009; Thornton and Coleman, 1975; Thornton and Slaven, 1993; Valderrama *et al.*, 2014b), estimation of the response from extensive portions of the auditory pathway (de la Torre *et al.*, 2020; de la Torre *et al.*, 2022; Holt and Ozdamar, 2016; Kohl *et al.*, 2019) and the analysis of the neurophysiological response to complex stimuli more natural than repetitive sequences of clicks, like natural speech (Maddox and Lee, 2018; Valderrama *et al.*, 2019) or responses associated with binaural activity (Martinez *et al.*, 2021).

The convolutional model assumes the hearing system to be linear and time invariant (LTI), which is an acceptable approach only under certain circumstances (Lütkenhöner, 2016; Valderrama *et al.*, 2014a). Under specific conditions (for example, if a stimulation pattern is presented repetitively, always at the same stimulation level) the responses evoked by each stimulus are expected to be identical and the hearing system can be considered LTI-like. However, if the expected response is not always the same, the behavior of the hearing system cannot be considered LTI-like (for example, if the stimuli are presented at different levels, the latency and amplitude of the waves change according to the stimulation level). Under a multi-response convolutional model (Bardy *et al.*, 2014a; Bardy *et al.*, 2014b; Valderrama *et al.*, 2016), different categories of stimuli can be defined, each category evoking a specific response. An appropriate categorization of the stimuli (for example, a categorization based on stimulation level) provides an LTI-like behavior of the hearing system without the requirement of repetitive stimulation patterns. This categorization increases the flexibility in the design of audiological experiments (Bardy, 2022; Bardy *et al.*, 2014a; Bardy *et al.*, 2014b; Bardy *et al.*,

2014c; Martinez *et al.*, 2021; Metzger *et al.*, 2020; Valderrama *et al.*, 2019; Valderrama *et al.*, 2016). In the multi-response paradigm, the deconvolution problem consists in the estimation of the response associated with each category of stimuli (Bardy *et al.*, 2014a; Bardy *et al.*, 2014b; Bardy *et al.*, 2014c; Dale, 1999).

Convolution can be mathematically described in terms of matrix formulation. If the response length is J samples, deconvolution involves the inversion of a $J \times J$ matrix. Multi-response deconvolution can also be described as a matrix-inversion problem. If the model involves M independent responses, each one with a length of J samples, the size of the matrix to be inverted is $(J \times M) \times (J \times M)$. Therefore, the computational cost of multi-response deconvolution increases very fast with the number of responses (since the complexity of matrix algorithms increases faster than $N_d^{2,3}$ for $N_d \times N_d$ matrices; Davie and Stothers, 2013), both if the deconvolution is performed as a matrix division (Bardy *et al.*, 2014a) or via an iterative procedure (Martinez *et al.*, 2021; Valderrama *et al.*, 2019; Valderrama *et al.*, 2016).

The size of the matrix to be inverted (and therefore, the computational cost) can significantly be reduced by the application of the recently proposed latency-dependent filtering and downsampling (LDFDS) procedure to the AEP responses (de la Torre *et al.*, 2020). LDFDS is based on the local bandwidth required to adequately represent each component of the AEP, progressively decreasing as the latency increases. A filtering method that is locally adapted to the required bandwidth reduces the noise of the responses, and the locally-adapted downsampling provides a substantial reduction of the dimensionality required to represent the evoked responses. This idea is particularly useful for processing the complete auditory pathway response (i.e., including brainstem, middle latency, and cortical components simultaneously): for example, the complete response usually requiring around 10 000 samples at a constant sampling-rate (one second sampled at 10 kHz) can be correctly represented after LDFDS with around 120 samples (three decades of latency represented with 40 samples per decade) (de la Torre *et al.*, 2020). As proposed by Lütkenhöner (2016) and de la Torre *et al.* (2022), the deconvolution problem can be constrained to a reduced representation space (defined according to the expected behavior of the responses). The deconvolution constrained to the reduced representation space provided by LDFDS provides AEP responses less affected by noise (thanks to the latency-dependent filtering) and substantially reduces the computational cost (thanks to the dimensionality reduction) (de la Torre *et al.*, 2022).

Based on the previous studies about deconvolution constrained to a reduced representation space and about multi-response deconvolution, in this work, we combine both concepts and propose the multi-response deconvolution constrained to the reduced representation space provided by LDFDS. As in the case of the single-response deconvolution, the proposed subspace-constrained deconvolution provides the least squares estimation of the evoked response

(de la Torre *et al.*, 2022). Additionally, the dimensionality reduction, from J (typically several thousands) to J_r (around one hundred), substantially reduces the size of the matrix to be inverted, $(J_r \times M) \times (J_r \times M)$, which significantly reduces the computational cost associated with the deconvolution. Therefore, multi-response AEP experiments which are usually difficult (or even impossible) to be deconvolved in the complete representation space, can easily be deconvolved in the reduced representation space.

This article presents the mathematical formulation of multi-response deconvolution performed in the reduced representation space. We demonstrate the viability of the proposed method using auditory responses evoked by clicks with randomized stimulation levels (Martinez *et al.*, 2023), in which the stimulus presentation level is uniformly distributed in the interval 0–80 dB normal hearing level (nHL), and the multi-response model is defined by categorizing the stimuli according to their level.

II. SUBSPACE-CONSTRAINED MULTI-RESPONSE DECONVOLUTION

A. Least-squares deconvolution

In an AEP recording process, the digital EEG signal $y(n)$ is usually modeled as the convolution of the stimulation sequence $s(n)$ (consisting of impulses at the beginning of each stimulation event) and the response signal $x(n)$,

$$y(n) = s(n) * x(n) + n_0(n), \tag{1}$$

where $n_0(n)$ is the noise affecting the EEG (de la Torre *et al.*, 2022; Jewett *et al.*, 2004; Ozdamar and Bohorquez, 2006).

If N is the number of samples in the EEG (at a given sampling rate) and J the length of the evoked response, this convolutional model can be rewritten using matrix notation as

$$\mathbf{y} = S\mathbf{x} + \mathbf{n}_0, \tag{2}$$

where \mathbf{y} , $S\mathbf{x}$ and \mathbf{n}_0 are N -component column vectors (representing, the EEG signal, the convolution of the stimulation signal with the response, and the noise, respectively), \mathbf{x} is a J -component column vector (representing the evoked response) and S is a $(N \times J)$ matrix (N rows and J columns) with $S(n, j) = s(n - j)$ providing the convolution $s(n) * x(n)$ as a matrix operation (Bardy *et al.*, 2014a; de la Torre *et al.*, 2019; de la Torre *et al.*, 2022).

The deconvolution of \mathbf{y} (i.e., the estimation of \mathbf{x}), can be formulated as an over-determined system of linear equations (with N equations and J unknowns, being $N \gg J$). In the context of linear algebra, the least squares (LS) criterion minimizes the squared residual $\|\mathbf{y} - S\mathbf{x}\|^2$, and provides the LS deconvolution of the response as

$$\hat{\mathbf{x}} = (S^T S)^{-1} S^T \mathbf{y}, \tag{3}$$

where S^T is the transpose of S (Bardy *et al.*, 2014a; de la Torre *et al.*, 2022; Gentle, 1998; Hayashi, 2000; Press *et al.*, 2002).

By defining the matrix S_k as S normalized and transpose (i.e., $S_k = S^T / K$, where K is the number of impulses in the stimulation sequence), and taking into account that $R_s = S_k S$ is the normalized $(J \times J)$ autocorrelation matrix of the stimulation sequence $s(n)$, the LS deconvolution can be rewritten as

$$\hat{\mathbf{x}} = R_s^{-1} S_k \mathbf{y} = R_s^{-1} \mathbf{z}_0, \tag{4}$$

where $\mathbf{z}_0 = S_k \mathbf{y}$ is a J -component vector obtained as the synchronous average of the EEG.

Therefore, the LS estimation of the evoked response requires the synchronous averaging of the EEG (\mathbf{z}_0) and the inversion of the $(J \times J)$ normalized autocorrelation matrix of the stimulation sequence (R_s). This LS estimation can be obtained by matrix division (as proposed in RSLSD; Bardy *et al.*, 2014a; Bardy *et al.*, 2014b; Bardy *et al.*, 2014c) or by an iterative procedure (as proposed by IRSA; de la Torre *et al.*, 2019). It is important to note that, according to Eq. (4), the synchronous average \mathbf{z}_0 is a biased estimator of the response. The unbiased estimation of the response requires the application of the inverse of the autocorrelation matrix R_s^{-1} in order to compensate for the effect of the response overlapping. Additionally, the application of the inverted matrix could amplify some components of the noise in the estimated response, due to low eigenvalues in R_s . The condition number of the matrix (i.e., the ratio of the largest to the smallest eigenvalues, in absolute value) provides an approximate evaluation of the magnitude of both effects (the bias of the synchronous average \mathbf{z}_0 and the potential amplification of specific components of the noise).

B. Multi-response deconvolution

Let us suppose that the auditory system is stimulated with M categories of stimulus, each one evoking a different response $x_m(n)$ (with $1 \leq m \leq M$). Assuming a multi-response convolutional model, the contribution of each category of stimulus m to the EEG is the convolution of its stimulation sequence $s_m(n)$ with the corresponding response $x_m(n)$, and the EEG is described as

$$y(n) = s_1(n) * x_1(n) + s_2(n) * x_2(n) + \dots + s_M(n) * x_M(n) + n_0(n), \tag{5}$$

which can be rewritten using matrix notation as

$$\mathbf{y} = S_1 \mathbf{x}_1 + S_2 \mathbf{x}_2 + \dots + S_M \mathbf{x}_M + \mathbf{n}_0, \tag{6}$$

where \mathbf{y} , $S_m \mathbf{x}_m$ and \mathbf{n}_0 are N -component column vectors, \mathbf{x}_m are J -component column vectors, and S_m are $(N \times J)$ matrices, providing the convolution of each category of response as a matrix operation. This matrix equation can be compacted as

$$\mathbf{y} = [S_1 S_2 \dots S_M] \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_M \end{bmatrix} + \mathbf{n}_0 = S_{all} \mathbf{x}_{all} + \mathbf{n}_0, \tag{7}$$

where S_{all} is the horizontal concatenation of the matrices S_m and \mathbf{x}_{all} the vertical concatenation of the vectors \mathbf{x}_m representing the responses. This matrix equation represents an over-determined system of linear equations, with N equations and $J \times M$ unknowns, and the formal LS solution is

$$\hat{\mathbf{x}}_{all} = [S_{all}^T S_{all}]^{-1} S_{all}^T \mathbf{y}. \quad (8)$$

If we expand the concatenated matrices and vectors, we can write,

$$\begin{bmatrix} \hat{\mathbf{x}}_1 \\ \hat{\mathbf{x}}_2 \\ \vdots \\ \hat{\mathbf{x}}_M \end{bmatrix} = \left(\begin{bmatrix} S_1^T \\ S_2^T \\ \vdots \\ S_M^T \end{bmatrix} [S_1 S_2 \cdots S_M] \right)^{-1} \begin{bmatrix} S_1^T \\ S_2^T \\ \vdots \\ S_M^T \end{bmatrix} \mathbf{y}, \quad (9)$$

or equivalently,

$$\begin{bmatrix} \hat{\mathbf{x}}_1 \\ \hat{\mathbf{x}}_2 \\ \vdots \\ \hat{\mathbf{x}}_M \end{bmatrix} = \begin{bmatrix} R_{s11} & R_{s12} & \cdots & R_{s1M} \\ R_{s21} & R_{s22} & \cdots & R_{s2M} \\ \vdots & \vdots & \ddots & \vdots \\ R_{sM1} & R_{sM2} & \cdots & R_{sMM} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{z}_{01} \\ \mathbf{z}_{02} \\ \vdots \\ \mathbf{z}_{0M} \end{bmatrix}, \quad (10)$$

where

$$R_{smn} = \frac{1}{K_m} S_m^T S_n \quad (11)$$

is the correlation matrix of the m th and n th stimulation sequences, normalized by the number of stimuli in the m th sequence, and

$$\mathbf{z}_{0m} = \frac{1}{K_m} S_m^T \mathbf{y} \quad (12)$$

is the synchronous average of the EEG for the m th stimulation sequence.

Therefore, the multi-response deconvolution procedure involves the following: (a) from the stimulation sequences $s_m(n)$ and the cross correlation among them, the estimation of the correlation matrices R_{smn} ; (b) from the stimulation sequences $s_m(n)$ and the EEG $y(n)$, the estimation of the synchronous average for each sequence \mathbf{z}_{0m} ; (c) the matrices R_{smn} are concatenated into R_{sall} and the vectors \mathbf{z}_{0m} into \mathbf{z}_{0all} ; (d) the concatenation of the responses $\hat{\mathbf{x}}_{all}$ is estimated by applying the inverted matrix R_{sall}^{-1} to the concatenated synchronous averaging \mathbf{z}_{0all} ,

$$\hat{\mathbf{x}}_{all} = \begin{bmatrix} \hat{\mathbf{x}}_1 \\ \hat{\mathbf{x}}_2 \\ \vdots \\ \hat{\mathbf{x}}_M \end{bmatrix} = R_{sall}^{-1} \mathbf{z}_{0all}, \quad (13)$$

and finally, (e) the responses to each category of stimulus $\hat{\mathbf{x}}_m$ can be extracted from the concatenated response $\hat{\mathbf{x}}_{all}$.

This multi-response deconvolution can be obtained by matrix division (Bardy *et al.*, 2014a), or alternatively, with an iterative procedure (Valderrama *et al.*, 2016). However, the numerical algorithms for matrix division (or the proposed iterative alternatives) show a computational complexity growing faster than $N_d^{2.3}$ (being N_d the number of dimensions, in this case, $J \times M$), which limits the practical use of multi-response deconvolution if the length of the response J or the number of responses M are large.

C. Deconvolution in a reduced representation space

The specific bandwidth of the AEPs for each portion of the response (range 100–3000 Hz for ABR; 10–300 Hz for MLR; 1–30 Hz for CAEPs) suggests a dimensionality reduction of the representation, based on a latency dependent filtering and downsampling (LDFDS) of the AEP responses (de la Torre *et al.*, 2020). The LDFDS operation reduces the noise of the AEP responses (most of the energy of the evoked responses is below ten oscillations/decade and therefore the signal-to-noise ratio (SNR) of the estimated responses improves by removing those components out of the band of interest at each latency) and provides a significant dimensionality reduction (typically from several thousand of samples to around one hundred) thanks to the latency-dependent downsampling. The LDFDS is applied to an AEP response \mathbf{x} as a $(J_r \times J)$ matrix operator V_r , where J and J_r are, respectively, the dimensionality of the original representation space and that of the reduced representation space (i.e., after filtering and downsampling),

$$\mathbf{x}_r = V_r \mathbf{x}, \quad (14)$$

where the impulsive response of the filter (implemented in the rows of V_r) changes from row to row to adapt to the bandwidth required at each latency, and the non-uniform downsampling is implemented by selecting appropriate rows in V_r according to the latency-specific bandwidth.

The LDFDS matrix is proposed as an orthonormal projector (the rows of the V_r matrix are orthonormal), which preserves the metrics (i.e., the distances and energies) in the reduced subspace and allows the recovery of the latency-dependent filtered response in the original representation (at the original uniform sampling rate and with J components) by applying the transpose of the V_r matrix to the response in the subspace representation \mathbf{x}_r (de la Torre *et al.*, 2020).

If the response to be estimated \mathbf{x} is appropriately represented with a reduced representation \mathbf{x}_r given by V_r , then we can write $\mathbf{x} = V_r^T \mathbf{x}_r$ and the convolutional model can be written as

$$\mathbf{y} = S V_r^T \mathbf{x}_r + \mathbf{n}_0, \quad (15)$$

and the formal LS solution of this deconvolution problem is (de la Torre *et al.*, 2022),

$$\begin{aligned} \hat{\mathbf{x}}_r &= \left((S V_r^T)^T (S V_r^T) \right)^{-1} (S V_r^T)^T \mathbf{y} \\ &= (V_r S^T S V_r^T)^{-1} V_r S^T \mathbf{y} \\ &= (V_r R_s V_r^T)^{-1} V_r \mathbf{z}_0 = R_{sr}^{-1} \mathbf{z}_{0r}, \end{aligned} \tag{16}$$

where R_{sr} and \mathbf{z}_{0r} are, respectively, R_s and \mathbf{z}_0 projected into the subspace.

In summary, subspace-constrained LS deconvolution of the EEG can be obtained with the following steps: (a) the autocorrelation matrix R_s and the synchronous averaging of the EEG \mathbf{z}_0 are transformed to the subspace using the transformation V_r ; (b) the autocorrelation matrix in the reduced representation is inverted; and (c) the inverted reduced autocorrelation matrix is applied to the reduced synchronous averaging. The LS deconvolution in Eq. (16) is similar to that in Eq. (4), with the difference that the problem is solved in the reduced representation space, and therefore it involves the vectors \mathbf{z}_{0r} and $\hat{\mathbf{x}}_r$ of J_r components (instead of J), and the inversion of the $(J_r \times J_r)$ matrix R_{sr} [instead of the inversion of the $(J \times J)$ matrix R_s], which provides a substantial reduction of computational cost (de la Torre *et al.*, 2022). Again, this subspace-constrained LS deconvolution can be implemented via matrix division or by an iterative procedure.

D. Multi-response deconvolution in a reduced representation space

As previously described, multi-response deconvolution is usually difficult to implement due to the high dimensionality $(J \times M)$ involved in the matrix division. On the other hand, subspace-constrained deconvolution provides a substantial dimensionality reduction (from J , typically several thousands, to J_r , around one hundred). The obvious solution for a practical implementation of multi-response deconvolution consists in its formulation in a J_r dimension reduced representation space, i.e., involving a $J_r \times M$ dimensionality in the matrix division. This section provides the mathematical formulation of multi-response deconvolution constrained to the reduced representation space.

Let us suppose a multi-response convolutional problem where each response \mathbf{x}_m can be described without information loss in a reduced representation space given by an orthonormal projector V_r . The reduced responses can therefore be expressed as $\mathbf{x}_{rm} = V_r \mathbf{x}_m$ (the responses can be recovered from the reduced responses as $\mathbf{x}_m = V_r^T \mathbf{x}_{rm}$), and the multi-response convolutional model can be written as

$$\begin{aligned} \mathbf{y} &= [S_1 S_2 \cdots S_M] \begin{bmatrix} V_r^T \mathbf{x}_{r1} \\ V_r^T \mathbf{x}_{r2} \\ \vdots \\ V_r^T \mathbf{x}_{rM} \end{bmatrix} + \mathbf{n}_0 \\ &= [S_1 V_r^T S_2 V_r^T \cdots S_M V_r^T] \begin{bmatrix} \mathbf{x}_{r1} \\ \mathbf{x}_{r2} \\ \vdots \\ \mathbf{x}_{rM} \end{bmatrix} + \mathbf{n}_0, \end{aligned} \tag{17}$$

which is an over-determined system of linear equations with N equations and $J_r \times M$ unknowns whose formal LS solution is,

$$\begin{bmatrix} \hat{\mathbf{x}}_{r1} \\ \hat{\mathbf{x}}_{r2} \\ \vdots \\ \hat{\mathbf{x}}_{rM} \end{bmatrix} = \begin{bmatrix} R_{sr11} & R_{sr12} & \cdots & R_{sr1M} \\ R_{sr21} & R_{sr22} & \cdots & R_{sr2M} \\ \vdots & \vdots & \ddots & \vdots \\ R_{srM1} & R_{srM2} & \cdots & R_{srMM} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{z}_{0r1} \\ \mathbf{z}_{0r2} \\ \vdots \\ \mathbf{z}_{0rM} \end{bmatrix}, \tag{18}$$

where

$$R_{srmm} = \frac{1}{K_m} V_r S_m^T S_m V_r^T = V_r R_{smm} V_r^T \tag{19}$$

is the correlation matrix of the m th and n th stimulation sequences, normalized by the number of stimuli in the m th sequence and projected to the subspace, and

$$\mathbf{z}_{0rm} = \frac{1}{K_m} V_r S_m^T \mathbf{y} = V_r \mathbf{z}_{0m} \tag{20}$$

is the synchronous average of the EEG for the m th stimulation sequence projected to the subspace.

In summary, multi-response deconvolution in the reduced representation space requires (a) the estimation of the correlation matrices R_{srmm} and the synchronous averages \mathbf{z}_{0m} ; (b) their projection into the subspace, R_{srmm} and \mathbf{z}_{0rm} (using the V_r transformation); (c) the concatenation of the reduced correlation matrices $R_{sr\text{all}}$ and the reduced synchronous averages $\mathbf{z}_{0r\text{all}}$; and (d) the application of the inverted matrix $R_{sr\text{all}}^{-1}$ to $\mathbf{z}_{0r\text{all}}$, providing the concatenated solutions in the subspace representation,

$$\hat{\mathbf{x}}_{r\text{all}} = \begin{bmatrix} \hat{\mathbf{x}}_{r1} \\ \hat{\mathbf{x}}_{r2} \\ \vdots \\ \hat{\mathbf{x}}_{rM} \end{bmatrix} = R_{sr\text{all}}^{-1} \mathbf{z}_{0r\text{all}}. \tag{21}$$

In order to obtain the responses in the original representation space $\hat{\mathbf{x}}_m$, the responses in the reduced representation are expanded by applying the transposed transformation matrix,

$$\hat{\mathbf{x}}_m = V_r^T \hat{\mathbf{x}}_{rm}. \tag{22}$$

The supplementary material includes a detailed explanation of the matrix operations involved in single- and multi-response convolution and deconvolution (supplementary material Sec. 1); MATLAB/Octave code (scripts and functions) implementing the subspace-constrained multi-response deconvolution proposed in this manuscript with some examples and simulations (supplementary material Sec. 2); and the extension of the subspace-constrained multi-response deconvolution for the case in which a specific orthonormal projector V_m is applied to each response \mathbf{x}_m (supplementary material Sec. 3).

III. EXPERIMENTS AND RESULTS

The proposed subspace-constrained multi-response deconvolution has been evaluated using auditory responses evoked by clicks presented at different levels and categorized according to the stimulation level. The purpose of the experiments is to demonstrate the viability of multi-response deconvolution performed in a reduced representation and to compare this method with the multi-response deconvolution performed in the complete representation space.

A. Experimental design

In these AEP experiments, the stimulation consisted of rarefaction clicks of 0.1 ms, presented at a randomized stimulation rate with an ISI in the range 15–30 ms (average stimulation rate of 44.44 stim/s), and with a random stimulation level with continuous uniform distribution in the range 0–80 dB nHL. The 0 dB nHL reference level was estimated as described in [Martinez et al. \(2023\)](#), i.e., as the mean threshold level estimated in a sample of ten normal-hearing adults (five female, 23–38 years) who presented pure-tone threshold levels within the normal range in the 0.5–8 kHz frequency range and had no history of any type of auditory dysfunction. The clicks were presented diotically through ER-3A insert earphones (providing a flat response in the range 0.1–4 kHz) with a tube length of around 28 cm.

The EEGs were recorded using surface active electrodes ([MettingVanRijn et al., 1996](#)) located at the forehead (Fz, active), the two mastoids (Tp9 and Tp10, reference), and middle forehead (Fp1, common mode sense; Fp2, driven right-leg), using a Biosemi ActiveTwo AD-box and the ActiView EEG acquisition software. The EEG was digitized in bipolar mode (recommended configuration for ABR) at a sampling rate of 16 384 Hz with 24 bits/sample. The bipolar mode includes a bandpass filtering (100–3300 Hz bandwidth) which has been extended (with digital filtering using an appropriate inverse filter) to the range 20–3300 Hz in order to allow the acquisition of MLR responses. The EEG to be deconvolved has been obtained as the signal at the active channel (Fz) minus the average of both reference channels (Tp9 and Tp10). The EEG database contains recordings from eight subjects (aged 23–51 years, two female). All participants met the inclusion criteria of reporting no hearing difficulties and absence of a history of auditory dysfunction. Each EEG recording consists of 1890 s (i.e., 31.5 min) and includes the responses to a total of 84 000 stimuli.

The AEP responses include 200 ms of transient after the presentation of the stimuli, and therefore the length of the responses is $J = 3277$ samples. Taking into account that most of the energy of the evoked responses is below ten oscillations per decade of latency, the latency-dependent filtering and downsampling are performed using a resolution of 40 samples/decade ([de la Torre et al., 2020](#)), which preserves the frequency components below 3.37 kHz, 628 and 68.8 Hz at latencies of 1, 10, and 100 ms, respectively, and

also reduces the dimensionality of the responses in the subspace to $J_r = 91$ components.

In these multi-response deconvolution experiments, in addition to the deconvolution of the responses, the matrix involved in the matrix division has been analyzed in terms of its condition number in order to evaluate potential problems associated with the matrix inversion or the matrix division problem [a large condition number reports ill-conditioned deconvolution with potential amplification of noise in the estimated responses ([Bardy et al., 2014a](#))].

The multi-response deconvolutions performed in both the complete and the reduced representation spaces are compared taking into account the deconvolved responses, the condition number of the matrices involved in the deconvolution, and the computational cost, which is evaluated in terms of execution time, measured using a desktop computer with an Intel-Core i7-3770 CPU, 3.40 GHz, 8.00 GB RAM running the algorithms in MATLAB.

B. Multi-response deconvolution in the complete and the reduced representation space

Figure 1 represents the AEP responses obtained with multi-response deconvolution in subject 1 (the latency axis is represented in logarithmic scale in order to facilitate the simultaneous visualization of the ABR and MLR components). The figure includes the responses obtained with the multi-response conventional deconvolution (i.e., in the complete representation space) and with the subspace-constrained deconvolution (i.e., in the reduced representation space). These responses are estimated from the same EEG. In this example, the multi-response deconvolution has been performed by categorizing the stimuli into ten groups according to the stimulation level (in ranges of 8 dB) and each response has been obtained from approximately 8400 stimuli. In these plots, the expected changes in the responses (i.e., in the amplitude and latency of the components) can be observed as a function of the stimulation intensity.

There are several important differences between the deconvolution performed in the complete or the reduced representation space. First, the latency-dependent filtering applied in the second case reduces the high frequency noise according to the local bandwidth, which is more evident for later latencies (the noise reduction provided by the latency dependent filtering looks like a simple low-pass filtering due to the logarithmic compression of the latency axis in the figure). Second, there is a substantial reduction of the execution time when the deconvolution is performed in the reduced representation space: in this example, with ten categories, the dimension of the multi-response in the complete representation space is 32 770, while it is only 910 in the reduced representation space, and the deconvolution took 1064.2 s in the complete representation space, while only 30.3 s in the reduced representation space. Finally, the eigenvalues of the matrix to be inverted could not be computed in the case of the complete representation space due to a memory overflow error (the allocation of a $32\,770 \times 32\,770$ matrix requires 8.0 GB of free memory).

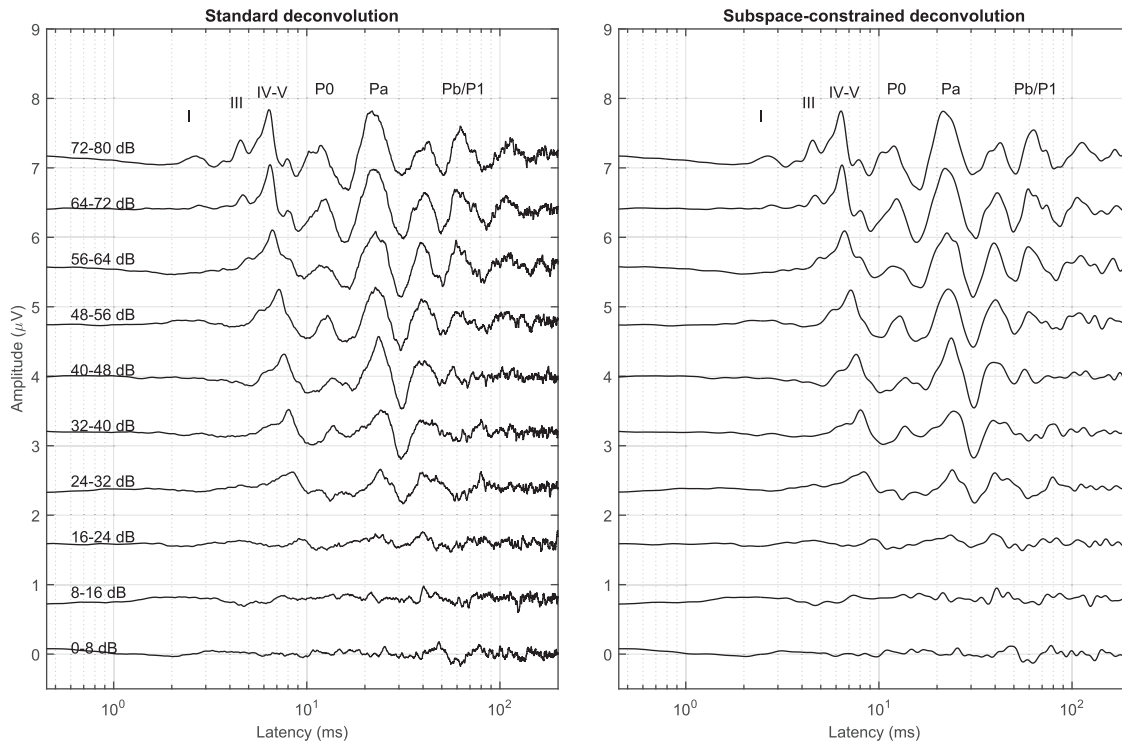


FIG. 1. AEP responses for subject 1, categorized by stimulation intensity in blocks of 8 dB (ten categories, approximately 8400 events/category). In the left panel, responses deconvolved in the complete representation space (standard deconvolution). In the right panel, responses deconvolved in the reduced representation space (subspace-constrained deconvolution). The main ABR and MLR components are labeled in the top trace. The stimulus level range per category is shown in the left panel.

Figure 2 shows the AEP responses for subject 1, for different categorizations based on the stimulation level, using between 4 and 32 categories (intervals between 20 and 2.5 dB). All these traces correspond to deconvolutions performed in the reduced representation space, and have been obtained from the same EEG (like those in Fig. 1). The total number of events (84000 with random stimulation level) are distributed among the different categories (21000, 10500, 5250, and 2625 events per category for 4, 8, 16, and 32 categories, respectively) according to the stimulation level. Each categorization and multi-response deconvolution was performed in an independent procedure after the EEG recording. These figures show that, as the intensity range decreases (the number of categories increases), the resolution in intensity improves, but fewer stimuli are involved in the estimation of each trace, and therefore the traces are more affected by noise. On the other hand, the categorization with wide intervals implies some spread of the latencies and amplitudes for the individual responses included in each category, making the LTI-like assumption weak and causing an amplitude reduction in the estimated responses, as observed for the intervals of 20 dB. Supplementary material (Sec. 4) includes similar plots for more categorizations (2, 3, 4, 5, 6, 8, 10, 12, 16, 20, 24, and 32 categories), for the multi-response deconvolution performed in the reduced representation space, and in the complete representation space for those cases in which it could be computed without a memory overflow error (i.e., ten categories or less).

Table I compares the computational cost when the deconvolution is performed in the complete or the reduced representation space. For each categorization of the stimuli (from 2 to 32 categories), the number of dimensions involved in the deconvolutions is specified, as well as the execution time for the estimation of the responses (including the initialization time, the deconvolution time, and the total execution time), the computation time devoted to the estimation of the eigenvalues, and the condition number of the matrix involved in the deconvolution. These results correspond to the deconvolution of the responses for subject 1. Figure 3 compares the execution times for the multi-response deconvolution performed in both the complete and the reduced representation space.

Since the representation of each response requires 3277 samples in the complete representation space but only 91 in the reduced representation space, there is a substantial reduction in the computational cost associated with the multi-response deconvolution, particularly as the number of categories increases. When the deconvolution is performed in the complete representation space, most of the computational cost is dedicated to the matrix division (except for categories 2 and 3). However, in the reduced representation space, the matrix division computational cost is ridiculous compared with that of the initialization (less than 0.2% in the worst case). The initialization time involves the computation of the cross-correlations (between the EEG and the stimulation sequences in order to estimate \mathbf{z}_{0all} , and among the stimulation sequences in order to estimate R_{sall}).

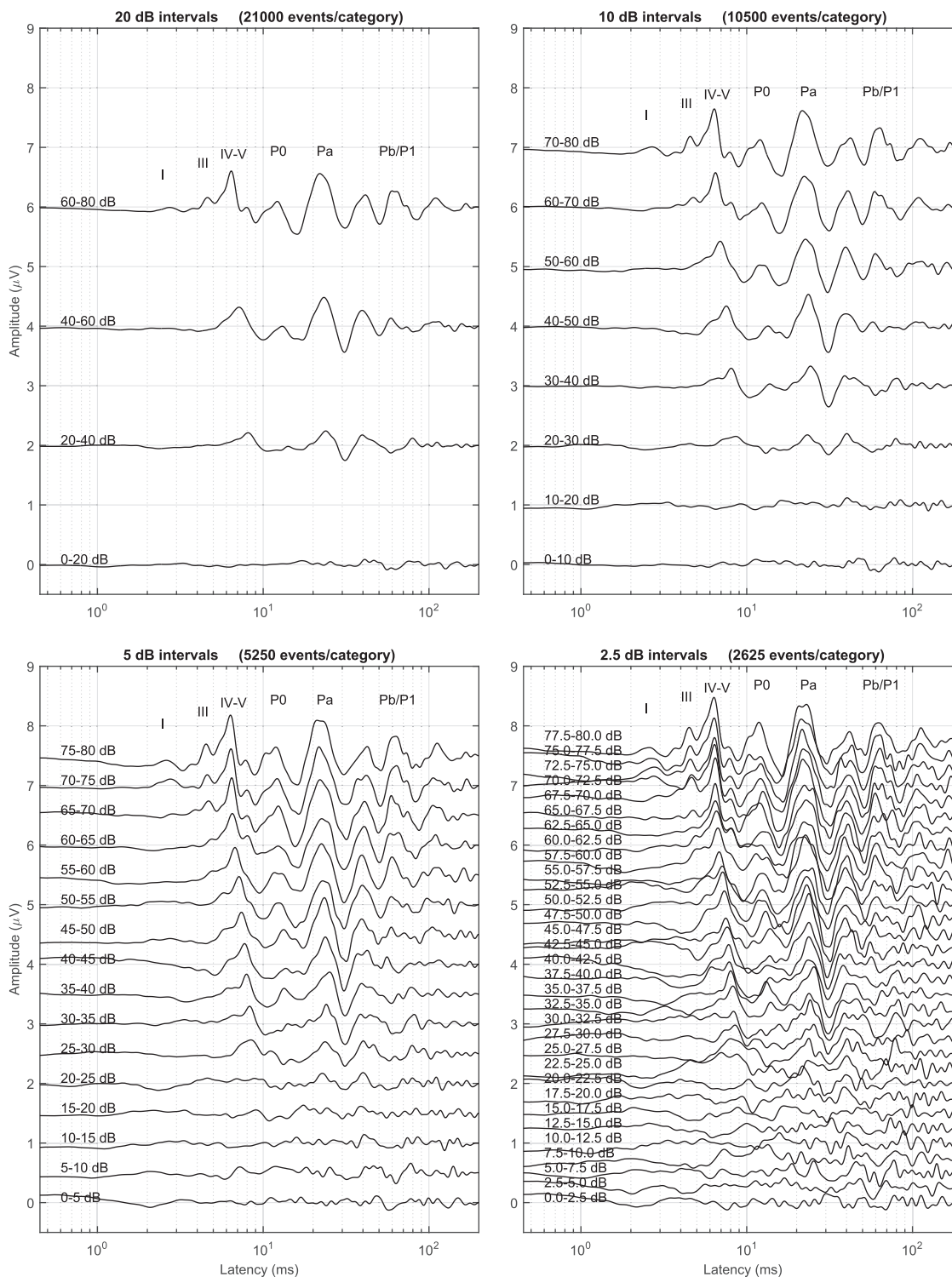


FIG. 2. AEP responses for subject 1, obtained with the multi-response deconvolution constrained to the reduced representation space, with different categorizations of the stimulation level: from top-left to bottom-right, 4, 8, 16, and 32 categories (intervals of 20, 10, 5, and 2.5 dB, respectively).

The initialization time is similar in the complete and reduced representation spaces, even though it is slightly smaller in the reduced representation space because the allocation of memory for $R_{s\text{all}}$ requires more resources than that for $R_{sr\text{all}}$.

In addition to the unpractical execution time when the multi-response deconvolution is performed in the complete

representation space, for 12 categories or more the size of the matrix involved in the deconvolution is too large and cannot be allocated in memory (leading to a memory overflow error).

The time devoted to the computation of the eigenvalues is also reported in Table I. While deconvolution requires a matrix division (which is computationally easier

TABLE I. Comparison of the computational cost for deconvolution performed in the complete or the reduced representation space. Execution time for (a) multi-response deconvolution (initialization, deconvolution and total) and (b) estimation of eigenvalues of the matrix, as a function of the number of categories (the width of the intensity interval and the number of dimensions involved in the deconvolution are also indicated for each experiment). In those cases where the eigenvalues could be computed, the computation time and the value of the condition number are also included.

Experiment		Complete representation space						Subspace-constrained deconvolution					
Numb. categ.	Interval (dB)	Numb. dimen.	Execution time (seconds)				Cond. number	Numb. dimen.	Execution time (seconds)				Cond. number
			init.	deconv.	total	eigenval.			init.	deconv.	total	eigenval.	
2	40.0	6554	7.80	2.54	10.34	104.05	230.0	182	7.15	0.0004	7.15	0.0085	220.7
3	26.7	9831	9.95	7.28	17.23	318.37	246.8	273	8.87	0.0008	8.87	0.0133	220.7
4	20.0	13 108	12.48	15.97	28.45	723.91	241.0	364	10.70	0.0017	10.70	0.0255	220.7
5	16.0	16 385	15.86	33.01	48.88	8642.96	256.2	455	12.99	0.0024	12.99	0.0547	220.8
6	13.3	19 662	20.60	59.16	79.76	mem. overflow		546	15.98	0.0049	15.99	0.0784	220.8
8	10.0	26 216	31.20	273.21	304.41	mem. overflow		728	22.78	0.0067	22.79	0.2239	220.8
10	8.0	32 770	46.17	1018.04	1064.21	mem. overflow		910	30.29	0.0120	30.30	0.2933	220.9
12	6.7	39 324	memory overflow			mem. overflow		1092	38.68	0.0218	38.70	0.5386	220.9
16	5.0	52 432	memory overflow			mem. overflow		1456	56.90	0.0531	56.95	1.1333	221.1
20	4.0	65 540	memory overflow			mem. overflow		1820	77.56	0.1038	77.67	2.4461	221.1
24	3.3	78 648	memory overflow			mem. overflow		2184	97.57	0.1114	97.68	3.6501	221.1
32	2.5	1 04 864	memory overflow			mem. overflow		2912	135.48	0.2279	135.71	7.8493	221.5

than a matrix inversion), the estimation of the eigenvalues of this matrix is significantly more expensive in terms of execution time and memory allocation. In the case of the complete representation space, the memory overflow error is reported for six categories or more, and the estimation of the eigenvalues takes around 2 h 24 min for five categories. As can be observed in the table, the condition number of the matrices (i.e., the ratio of the largest to the smallest eigenvalues) is smaller in the reduced representation space, meaning the LS deconvolution is better conditioned in the reduced representation space. This result is expected

according to the Cauchy interlacing theorem (also known as the Poincaré separation theorem), which gives upper and lower bounds of eigenvalues of a real symmetric matrix obtained as the orthogonal projection of a larger real symmetric matrix.

Supplementary material (Sec. 5) includes plots representing the execution time for the multi-response deconvolution and the eigenvalues computation (as a function of both, the number of estimated responses, and the number of components of the multi-response vector), an analysis of the memory requirements and an estimation of the computational cost expected for other computer architectures.

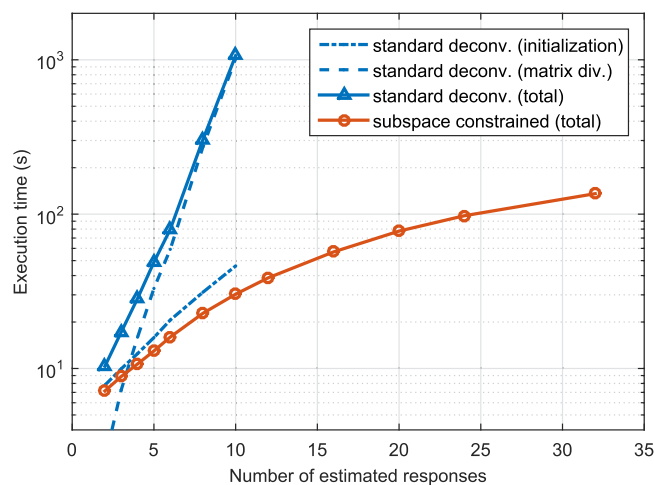


FIG. 3. (Color online) Execution time required for the estimation of all the responses with the multi-response deconvolution. Comparison of deconvolution performed in the complete representation space (standard deconvolution, blue line with triangles) and in the reduced representation space (subspace-constrained deconvolution, red line with circles). The execution times required by the initialization and the deconvolution are split in the case of the complete representation space.

C. Grand average and individual AEP responses

Figure 4 shows the individual AEP responses, using a multi-response deconvolution with 16 categories (stimulation intervals of 5 dB) estimated from the 31.5 min of EEG available for each subject. Each trace is obtained with approximately 5250 stimuli. The grand average AEP response (involving 42 000 stimuli per trace) is also included. The main ABR and MLR components are labeled in the first trace of each plot. Supplementary material (section 6) includes individual and grand average responses, estimated with several categorizations (intervals of 20, 10, and 5 dB).

The plots with individual traces (subjects 1 to 8) visually show that most of the AEP components from wave I (of the ABR) to the Pb (of the MLR) can be identified at the higher stimulation levels. The amplitude of the components decreases with the reduction of the stimulation intensity, and also an increase in the latencies, particularly in the ABR components, is observed. The comparison of the grand average and the individual AEP responses reveals a high consistency across subjects for most of the components.

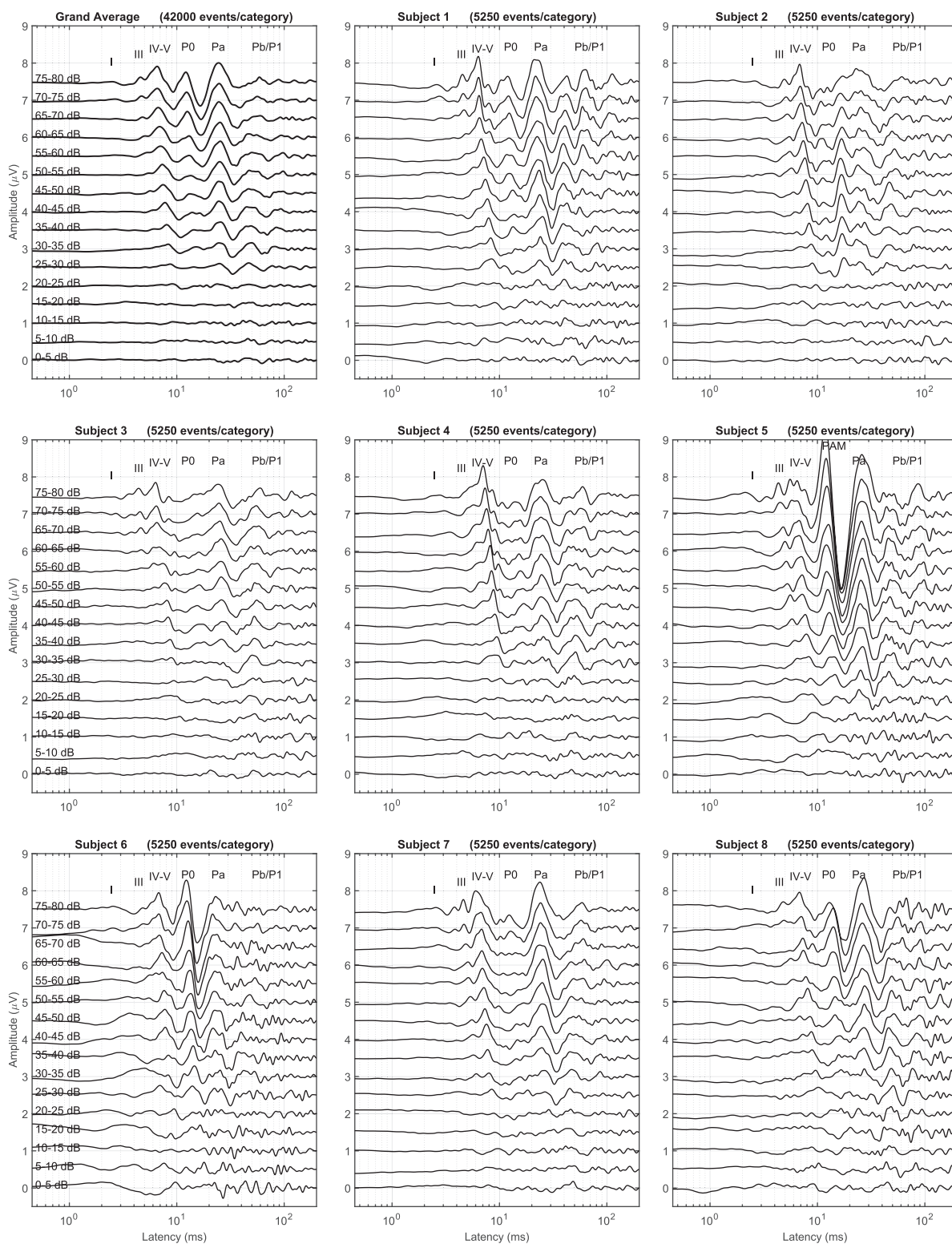


FIG. 4. AEP responses (grand average and individual responses for each subject) categorized by stimulation intensity in blocks of 5 dB (16 categories). The stimulus level range is shown in the left panels. A subspace-constrained deconvolution has been applied.

A post-auricular muscle response (PAMR, an action potential occurring at approximately 13 ms after the stimulus onset resulting from the contraction of the post-auricular muscle) is observed in subject 5 (Sec. 6 of the supplementary material compares the grand average responses including and excluding this subject).

D. Quality of the AEP estimations for different categorizations

The quality of the AEP responses was determined in terms of their SNR. The SNR was calculated as the ratio between the energy of the AEP waveform and the energy of the noise in the latency interval 0–200 ms. The SNR was

calculated in the reduced representation (since the transformation V_r between the complete and the reduced representation spaces is orthonormal, the energy is the same in both representations). The respective energies of the response and the noise are obtained according to the “plus-minus reference” procedure proposed by Schimmel (1967). In our implementation, the procedure was estimated using the first half and the second half of the available epochs (Martinez et al., 2023), specifically for each categorization and for each subject.

Figure 5 represents the average SNR estimated for each categorization and each subject as a function of the interval width used for the categorization. The figure includes individual traces for each subject (with triangles) and traces for the grand average (with circles). The number of categories is also indicated in the figure. A reduction in a factor 2 in the number of categories (for example, from 32 categories, 2.5 dB intervals, to 16 categories, 5 dB intervals) implies an increment in a factor 2 in the number of epochs per category contributing to the response estimation, and therefore, a SNR improvement of 3 dB would be expected (because the amplitude of the noise in the estimated response would decrease 3 dB as the number of epochs is doubled). This improvement is approximately achieved on the left side of the plot (many categories with narrow intervals), but the improvement is smaller than expected (the slope decreases) as we move to the right side of the plot. The reduction of the SNR improvement with respect to the expected one seems to be associated with the fact that wide categories cause a spread of latencies (and a weak verification of the LTI-like assumption) and the subsequent amplitude reduction in the response estimation (as observed in Fig. 2).

Supplementary material (Sec. 7) contains a detailed description of the procedure for the SNR estimation (including a simulation to verify that the procedure provides a non-biased SNR estimation) and detailed results (including also

the average signal and noise levels estimated for each subject and categorization).

E. Estimation of latencies and amplitudes

Even though a very detailed analysis of the morphology of the estimated AEP responses is not the main objective of this manuscript (because of its extension, such analysis will be considered in a future manuscript), in Fig. 6, we include measurements of amplitudes and latencies for the ABR wave V and the MLR wave Na-Pa (two of the most consistent components across subjects) as a function of the stimulation level. These measurements correspond to a categorization of the responses in intervals of 10 dB (which provides a reasonable trade-off between noise and resolution). The results presented in the figure include individual measurements for each subject (triangles) and mean across the subjects (circles). The latency results are corrected (reduced by 0.82 ms with respect to the measurements in the evoked responses) taking into account the length of the transducer tube (28 cm).

In the case of wave V, a clear dependence is observed in the plots, with a reduction of latency and an increase in the amplitude of the wave as the stimulation level increases. Due to the noise affecting the estimated evoked responses, the amplitude measurements are affected by fluctuations. The latency measurements are significantly more stable. In the case of wave Na-Pa, a strong dependence is observed in the amplitude, with a significant inter-subject variability. The plot with the largest amplitude corresponds to subject 5 (with a large PAMR component). In this case, the Na negative peak is influenced by the PAMR component (as observed in Fig. 4), which causes an overestimation of the Na-Pa amplitude. For this reason, the measurements corresponding to this subject were excluded from the average. The dependence on the stimulation level is significantly smaller for the latency of the peak Pa. Supplementary material (Sec. 8) includes measurements of waves V and Na-Pa (latencies and amplitudes) for categorizations in intervals of 20 dB (less noisy responses but with lower resolution) and 5 dB (higher resolution but more noisy responses).

IV. DISCUSSION AND CONCLUSIONS

In this article, we propose multi-response deconvolution of AEPs performed in the reduced representation space provided by the latency-dependent filtering and downsampling. The proposed method can be considered a natural and obvious combination of the multi-response deconvolution proposed by Bardy et al. (2014a) and the subspace-constrained deconvolution proposed by de la Torre et al. (2022). However, the proposed method provides a very useful and practical tool, because in the complete representation space, the multi-response deconvolution is difficult to be applied due to the computational cost, while this difficulty is alleviated in the reduced representation space. In the experiments presented in Fig. 1, the time required for the AEP estimation

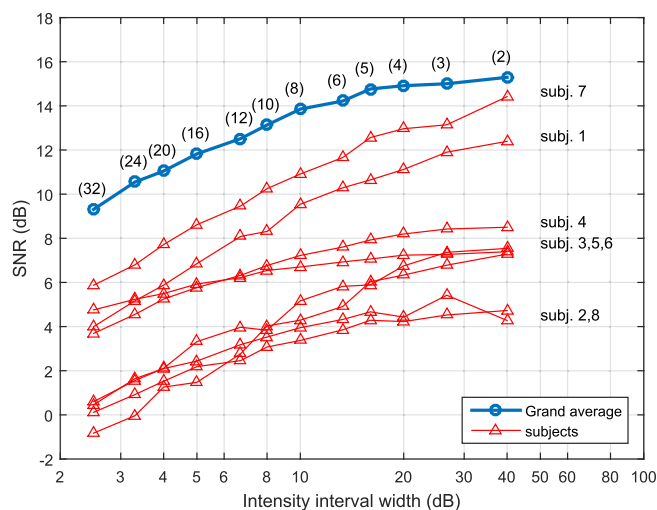


FIG. 5. (Color online) Average estimated SNR for each categorization and each subject as a function of the interval width used for the categorization (the number of categories is indicated in the top trace in parenthesis).

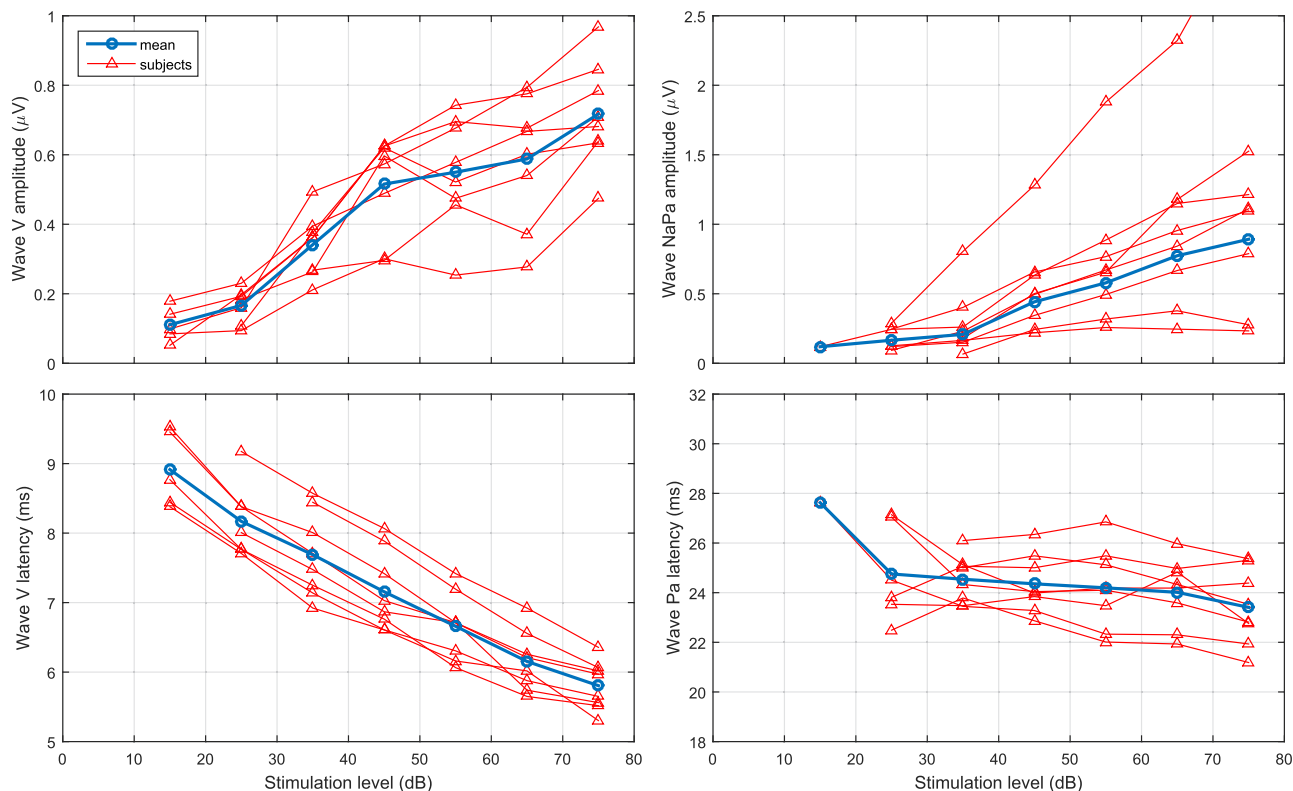


FIG. 6. (Color online) Measurements of the ABR wave V amplitude (top-left) and latency (bottom-left), and the MLR wave Na-Pa amplitude (top-right) and latency (bottom-right) as a function of the stimulation level. These measurements were obtained from the responses estimated with eight categorizations in intervals of 10 dB.

was reduced from 18 min to 30 s (and with more categories, the AEP estimation was not possible in the complete representation space due to memory overflow). Interestingly, the computational time required to deconvolve the responses (135.7 s for 32 categorizations in the most complex experiment) was significantly smaller than the duration of the EEG (i.e., the time devoted to EEG recording, 1890 s in the experiments). This fact could be applied in the future to estimate and represent the responses in real time during the recording session. Additionally, the latency-dependent filtering involved in the reduced representation provides a reduction of high frequency noise.

The reduced representation space also allows a detailed analysis of the matrix involved in the deconvolution (for instance, the estimation of the eigenvalues or the condition number). This analysis is computationally more expensive than the deconvolution. The eigenvalues and the condition number are useful to assess potential problems in the deconvolved responses (the matrix division amplifies those components with low eigenvalues, and the over-amplification of background noise increases the risk of highly noisy AEP estimations). On the other hand, the dimensionality reduction guarantees a reduction of the condition number according to the Cauchy interlacing theorem, as discussed in de la Torre *et al.* (2022), which is more important in the multi-response deconvolution (the condition number grows with the number of categories, particularly in the complete representation space, as can be observed in Table I). This

reduction of the condition number leads to a better conditioned deconvolution problem (less susceptible to noise amplification). The supplementary material (Sec. 9) includes an analysis (including examples and simulations) of the eigenvalues and the condition number and their utility to detect ill-conditioned deconvolution problems.

The evoked responses observed in Figs. 1, 2, and 4 are compatible with those described in de la Torre *et al.* (2019, 2020) and de la Torre *et al.* (2022) (for the ISI range 15–30 ms). These plots provide clear and consistent ABR and MLR responses. However, in the responses presented in this article, the CAEP components are difficult to observe due to the filter applied to the EEG (with a 20–3300 Hz bandwidth). The Pb/P1 wave (latency around 60 ms) is attenuated with respect to the amplitudes observed in the previous works, and the P2 wave, also attenuated, can be observed only in subjects 1 and 3. As in the previous works (even though with some limitations due to the bandwidth), the proposed comprehensive AEP representation facilitates the simultaneous exploration of peripheral and central components.

The proposed multi-response deconvolution allows an *a posteriori* categorization, as can be observed in Fig. 2. In these experiments, the stimuli can easily be categorized according to their intensity, and different categorizations can be defined using different sizes of the intensity intervals because the response morphology changes only subtly for small variations of the stimulation intensity.

Using few categories (with wide intensity intervals and many stimuli in each category) reduces the computational cost of the multi-response deconvolution and provides responses less affected by noise, as observed in Fig. 5, but affected by an amplitude reduction (associated with the spread of amplitudes and latencies within each category and the subsequent weak verification of the LTI-assumption), as observed in Fig. 2. With more categories, the amplitude of the components is usually larger (as observed in Fig. 2 and in the Secs. 4 and 6 in the supplementary material), because the spread in their latencies is smaller. Additionally, a better resolution in intensity is achieved (which would be important if the purpose of the exploration was, for example, to estimate click-based hearing thresholds). The optimal categorization would depend on the total number of available stimuli (which grows with the recording time) and the level of the background noise in the EEG. The flexibility in the multi-response categorization could be very useful for clinical applications of AEP explorations in order to optimize the information obtained from the recorded data since the categorization could easily be adapted after the exploration (or even during the exploration) according to the noise conditions and the purpose of the exploration (for example, identification of an electrophysiological response, or accurate determination of click-based thresholds).

One of the limitations of the convolutional model for the AEP responses is the LTI assumption of the hearing system. The hearing system is not linear (an increase in the stimulus level produces a non-linear but approximately logarithmic increase in the response, with threshold and saturation effects and with changes in the latencies and morphology). Similarly, the hearing system is not time invariant due to adaptation and fatigue mechanisms (Valderrama *et al.*, 2014a; Valderrama *et al.*, 2016). Under specific circumstances, the no-LTI hearing system behaves as an LTI-like system (for example for a quasi-stationary presentation of repetitive stimuli), if the response to each stimulus can be assumed to be identical. In such a case, the convolutional model can be applied to estimate the responses via deconvolution (or with synchronous averaging if the responses are not overlapped), but this limitation strongly constrains the stimulation patterns that can be used to evoke consistent responses. The proposed multi-response deconvolution allows the definition of different categories of stimuli (which are expected to evoke different responses) and therefore allows for the extension of the LTI-like assumption to stimulation patterns much more elaborated than the classical repetitive patterns. This flexibility can be used to make the acquisition sessions more comfortable for the explored subjects (some of them find annoying or boring the repetitive stimulation patterns, Martinez *et al.*, 2023), or for the assessment of specific aspects of the hearing perception (requiring complex stimulation patterns).

The multi-response estimation of evoked responses have some precedents in the literature. In the context of auditory steady state responses (ASSR) the simultaneous estimation of responses to different stimulation frequencies

is very common (John *et al.*, 1998; Lins and Picton, 1995; Luts *et al.*, 2006), and more recently, some interesting experiments have been proposed to evaluate the auditory attention (Manting *et al.*, 2023). Compared with those approaches, the multi-response deconvolution proposed in the present article has the advantage of providing the transient evoked responses, and therefore information about the source of the neural activity. In the context of multiple estimations of transient responses, the parallel auditory brainstem response (Polonenko and Maddox, 2019, 2022) proposes a multi-frequency stimulation procedure using non-correlated stimulation sequences and therefore allowing them to deconvolve the evoked response using the synchronous average (without the need of an elaborated deconvolution procedure and without limitations related to computational cost). This is possible (in spite of the response overlapping) because the autocorrelation matrix is the identity matrix for uncorrelated stimulation sequences. However, this approach is restricted to experiments in which the stimulation events are specifically configured to be statistically independent (which restricts the flexibility in the experimental design to stimulation sequences following a Poisson distribution). Additionally, the large ISI spread of uncorrelated stimulation sequences could compromise the verification of the LTI-assumption (weak when epochs with different expected responses are grouped under the same category), and makes it recommendable the verification of the size of the ISI effect (Maddox and Lee, 2018; Polonenko and Maddox, 2019, 2022; Stoll and Maddox, 2023) in order to guarantee an accurate estimation of the responses.

In general, a quasi-uncorrelated stimulation sequence (i.e., with large ISI spread) provides a low condition number in the autocorrelation matrix. This guarantees low amplification of the noise in the estimated response $\hat{\mathbf{x}}$ and a small bias between the response and the synchronous average \mathbf{z}_0 (as discussed and simulated in Sec. 9 of the supplementary material). However, a large spread of ISI compromises the verification of the LTI-assumption, and could lead to incorrect estimation of the responses. On the other hand, a narrow ISI range guarantees the verification of the LTI-assumption, but causes a large condition number (and amplification of noise, requiring an increment of the number of events to compensate for this effect). According to our previous experience with adaptation of evoked responses (de la Torre *et al.*, 2022; Valderrama *et al.*, 2014a; Valderrama *et al.*, 2016), a uniform ISI distribution with the upper limit twice the lower one (one octave ISI range) provides a reasonable trade-off between the verification of the LTI-assumption and the condition number.

In the present article, we have applied a relatively simple stimulation pattern (the stimuli are always the same except for the stimulation level), in order to describe the proposed procedure with a simple experiment. However, the proposed multi-response deconvolution can be extended to experimental designs using more complex stimulation patterns and more audacious and creative categorizations. In a previous work, we stimulated with bursts of clicks, and the

stimuli were categorized taking into account their position in the burst in order to study fast and slow adaptation mechanisms in the hearing system (Valderrama *et al.*, 2019). In another study, the stimuli were presented binaurally, where the stimulation pattern was affected by changes in the interaural delay, and different categories were defined for the stimulation pattern or the changes in the interaural delay (Martinez *et al.*, 2021). In these studies, an iterative (and time consuming) multi-response deconvolution was applied, and the experimental design was constrained by the computational cost required by the deconvolution. With the procedure proposed in this article, the computational cost is alleviated, and more complex AEP experiments can be designed, which will offer new perspectives for the research and assessment of the hearing system.

SUPPLEMENTARY MATERIAL

See the supplementary material for (Sec. 1) detailed description of convolution and deconvolution with matrix formulation including examples and simulations; (Sec. 2) description of the code (also provided as supplementary material) with algorithms, examples of use and simulations; (Sec. 3) extension of the subspace-constrained multi-response deconvolution for the case of response-specific transformation; (Sec. 4) detailed results of AEP responses for subject 1; (Sec. 5) detailed results about computational cost; (Sec. 6) detailed results about individual and grand average AEP responses; (Sec. 7) description of the procedure for SNR estimation in AEP responses and detailed results; (Sec. 8) detailed results about estimated latencies and amplitudes of ABR wave V and MLR wave Na-Pa; (Sec. 9) a detailed analysis about eigenvalues and matrix condition number, including simulations for single- and multi-response deconvolution. See also supplementary material for a zip file with code and data with algorithms, examples and simulations: (a) scripts implementing simple examples of convolution and deconvolution; (b) functions, scripts and data with the algorithms for multi-response deconvolution, with examples of use and simulations; (c) a script with a simulation verifying the accuracy of the procedure for SNR estimation; (d) functions and scripts implementing simulations related to the effect of the eigenvalues and the condition number on the deconvolved responses.

ACKNOWLEDGMENTS

This research was supported by the (1) PID2020-119073GB-I00 project grant, funded by MICIN / AEI / 10.13039/501100011033; (2) B.TIC.382.UGR20 project grant, funded by “Junta de Andalucía—Consejería de Economía, Conocimiento, Empresas y Universidad” and by the “European Union - ERDF A way to make Europe”; (3) ProyExcel_00152 project grant, funded by “Junta de Andalucía - Consejería de Universidad, Investigación e Innovación”; and (4) “Ayudas María Zambrano” post-doctoral fellowship, funded by the Ministry of Universities of the Spanish Government and by the “European Union—

Next Generation EU.” The authors would like to sincerely acknowledge the contribution and implication of our coauthor J. C. Segura, who unfortunately passed away September 10, 2023. The authors also acknowledge the constructive revision of the anonymous reviewers.

AUTHOR DECLARATIONS

Conflict of Interest

The authors declare no conflict of interest.

Ethics Approval

Ethics approval to conduct the present research was obtained from the Human Research Ethics Committee of the University of Granada (Ref. no. 961/CEIH/2019).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

- Bardy, F. (2022). “Deconvolution of ear’s activity (DEA): A new experimental paradigm to investigate central auditory processing,” *Front. Syst. Neurosci.* **16**, 892198.
- Bardy, F., Dillon, H., and Dun, B. V. (2014a). “Least-squares deconvolution of evoked potentials and sequence optimization for multiple stimuli under low-jitter conditions,” *Clin. Neurophysiol.* **125**, 727–737.
- Bardy, F., Dun, B. V., Dillon, H., and Cowan, R. (2014b). “Least-squares (LS) deconvolution of a series of overlapping cortical auditory evoked potentials: A simulation and experimental study,” *J. Neural Eng.* **11**, 046016.
- Bardy, F., Dun, B. V., Dillon, H., and McMahon, C. M. (2014c). “Deconvolution of overlapping cortical auditory evoked potentials recorded using short stimulus onset-asynchrony ranges,” *Clin. Neurophysiol.* **125**, 814–826.
- Bohorquez, J., and Ozdamar, O. (2006). “Signal to noise ratio analysis of maximum length sequence deconvolution of overlapping evoked potentials,” *J. Acoust. Soc. Am.* **119**, 2881–2888.
- Burkard, R. F., and Don, M. (2007). “The auditory brainstem response,” in *Auditory Evoked Potentials: Basic Principles and Clinical Application*, edited by R. Burkard, M. Don, and J. Eggermont (Lippincott Williams & Wilkins, Baltimore, MD), pp. 229–253.
- Burkard, R. F., Finneran, J. J., and Mulsow, J. (2018). “Comparison of maximum length sequence and randomized stimulation and averaging methods on the bottlenose dolphin auditory brainstem response,” *J. Acoust. Soc. Am.* **144**, 308–318.
- Dale, A. (1999). “Optimal experimental design for event-related fMRI,” *Hum. Brain Mapp.* **8**, 109–114.
- Davie, A. M., and Stothers, A. J. (2013). “Improved bound for complexity of matrix multiplication,” *Proc. R. Soc. Edinburgh Sect. A: Math.* **143A**(2), 351–370.
- de la Torre, A., Valderrama, J. T., Segura, J. C., and Alvarez, I. M. (2019). “Matrix-based formulation of the iterative randomized stimulation and averaging method for recording evoked potentials,” *J. Acoust. Soc. Am.* **146**, 4545–4556.
- de la Torre, A., Valderrama, J. T., Segura, J. C., and Alvarez, I. M. (2020). “Latency-dependent filtering and compact representation of the complete auditory pathway response,” *J. Acoust. Soc. Am.* **148**, 599–613.
- de la Torre, A., Valderrama, J. T., Segura, J. C., Alvarez, I. M., and Garcia-Miranda, J. (2022). “Subspace-constrained deconvolution of auditory evoked potentials,” *J. Acoust. Soc. Am.* **151**, 3745–3757.
- Eysholdt, U., and Schreiner, C. (1982). “Maximum length sequences: A fast method for measuring brain-stem-evoked responses,” *Audiology* **21**, 242–250.

- Finneran, J. J., Mulsow, J., and Burkard, R. F. (2019). "Signal-to-noise ratio of auditory brainstem responses (ABRs) across click rate in the bottlenose dolphin (*Tursiops truncatus*)," *J. Acoust. Soc. Am.* **145**, 1143–1151.
- Gentle, J. E. (1998). *Numerical Linear Algebra for Applications in Statistics* (Springer, New York).
- Gillespie, P. G., and Muller, U. (2009). "Mechanotransduction by hair cells: Models, molecules, and mechanisms," *Cell* **139**, 33–44.
- Hall, J. W. (2007). *New Handbook of Auditory Evoked Potentials* (Pearson Education, Boston, MA), pp. 58–108.
- Hayashi, F. (2000). *Econometrics* (Princeton University Press, Princeton, NJ).
- Holt, F., and Ozdamar, O. (2016). "Effects of rate (0.3–40/s) on simultaneously recorded auditory brainstem, middle and late responses using deconvolution," *Clin. Neurophysiol.* **127**, 1589–1602.
- Jewett, D. L., Caplovitz, G., Baird, B., Trumpis, M., Olson, M. P., and Larson-Prior, L. J. (2004). "The use of QSD (q-sequence deconvolution) to recover superposed, transient evoked-responses," *Clin. Neurophysiol.* **115**, 2754–2775.
- John, M. S., Lins, O. G., Boucher, B. L., and Picton, T. W. (1998). "Multiple auditory steady-state responses (MASTER): Stimulus and recording parameters," *Int. J. Audiol.* **37**, 59–82.
- Kohl, M. C., Schebsdat, E., Schneider, E. N., Niehl, A., Strauss, D. J., Ozdamar, O., and Bohorquez, J. (2019). "Fast acquisition of full-range auditory event-related potentials using an interleaved deconvolution approach," *J. Acoust. Soc. Am.* **145**, 540–550.
- Lins, O. G., and Picton, T. W. (1995). "Auditory steady-state responses to multiple simultaneous stimuli," *Electroencephalogr. Clin. Neurophysiol.* **96**, 420–432.
- Lütkenhöner, B. (2010). "Baseline correction of overlapping event-related responses using a linear deconvolution technique," *NeuroImage* **52**, 86–96.
- Lütkenhöner, B. (2016). "Estimation of a transient response from steady-state responses by deconvolution with built-in constraints," *J. Theor. Biol.* **404**, 143–159.
- Luts, H., Desloovere, C., and Wouters, J. (2006). "Clinical application of dichotic multiple-stimulus auditory steady-state responses in high-risk newborns and young children," *Audiol. Neurotol.* **11**, 24–37.
- Maddox, R. K., and Lee, A. K. C. (2018). "Auditory brainstem responses to continuous natural speech in human listeners," *eNeuro* **5**, e0441.
- Manting, C. L., Gulyas, B., Ullén, F., and Lundqvist, D. (2023). "Steady-state responses to concurrent melodies: Source distribution, top-down, and bottom-up attention," *Cereb. Cortex* **33**, 3053–3066.
- Martinez, M., Valderrama, J. T., Alvarez, I., Vargas, J. L., and de la Torre, A. (2021). "The transient response to interaural time differences," in *Proceedings of the XXVII International Evoked Response Audiometry Study Group (IERASG-2021)*, June 14–July 11, Virtual, p. 50.
- Martinez, M., Valderrama, J. T., Alvarez, I., Vargas, J. L., and de la Torre, A. (2023). "Auditory brainstem responses obtained with randomised stimulation level," *Int. J. Audiol.* **62**(4), 368–375.
- MettingVanRijn, A. C., Dankers, A. P. K. T. E., and Grimbergen, C. A. (1996). "Low-cost active electrode improves the resolution in biopotential recordings," in *Proceedings of the 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, October 31–November 3, Amsterdam, The Netherlands, pp. 101–102.
- Metzger, B., Magnotti, J., Wang, Z., Nesbitt, E., Karas, P., Yoshor, D., and Beauchamp, M. (2020). "Responses to visual speech in human posterior superior temporal gyrus examined with iEEG deconvolution," *J. Neurosci.* **40**, 6938–6948.
- Ozdamar, O., and Bohorquez, J. (2006). "Signal-to-noise ratio and frequency analysis of continuous loop averaging deconvolution (CLAD) of overlapping evoked potentials," *J. Acoust. Soc. Am.* **119**, 429–438.
- Polonenko, M., and Maddox, R. (2019). "The parallel auditory brainstem response," *Trends Hear.* **23**, 233121651987139.
- Polonenko, M., and Maddox, R. (2022). "Optimizing parameters for using the parallel auditory brainstem response (PABR) to quickly estimate hearing thresholds," *Ear Hear.* **43**(2), 646–658.
- Press, W. H., Teutolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2002). *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, NY).
- Schimmel, H. (1967). "The (\pm) reference: Accuracy of estimated mean components in average response studies," *Science* **157**, 92–94.
- Stoll, T., and Maddox, R. (2023). "Enhanced place specificity of the parallel auditory brainstem response: A modelling study," *Trends Hear.* **27**, 1–11.
- Thornton, A. R. D. (2007). "Instrumentation and recording parameters," in *Auditory Evoked Potentials: Basic Principles and Clinical Application*, edited by R. Burkard, M. Don, and J. Eggermont (Lippincott Williams & Wilkins, Baltimore, MD), pp. 73–101.
- Thornton, A. R. D., and Coleman, A. (1975). "The adaptation of cochlear and brainstem auditory evoked potentials in humans," *Electroencephalogr. Clin. Neurophysiol.* **39**, 399–406.
- Thornton, A. R. D., and Slaven, A. (1993). "Auditory brainstem responses recorded at fast stimulation rates using maximum length sequences," *Br. J. Audiol.* **27**, 205–210.
- Valderrama, J. T., Alvarez, I., de la Torre, A., Segura, J. C., Sainz, M., and Vargas, J. L. (2012). "Recording of auditory brainstem response at high stimulation rates using randomized stimulation and averaging," *J. Acoust. Soc. Am.* **132**, 3856–3865.
- Valderrama, J. T., de la Torre, A., Alvarez, I., Segura, J. C., Thornton, A. R. D., Sainz, M., and Vargas, J. L. (2014a). "Auditory brainstem and middle latency responses recorded at fast rates with randomized stimulation," *J. Acoust. Soc. Am.* **136**, 3233–3248.
- Valderrama, J. T., de la Torre, A., Alvarez, I., Segura, J. C., Thornton, A. R. D., Sainz, M., and Vargas, J. L. (2014b). "A study of adaptation mechanisms based on abr recorded at high stimulation rate," *Clin. Neurophysiol.* **125**, 805–813.
- Valderrama, J. T., de la Torre, A., Dun, B. V., and Segura, J. C. (2019). "Towards the recording of brainstem and cortical evoked potentials from the fine structure of natural speech," in *Proceedings of the XVI International Evoked Response Audiometry Study Group (IERASG) Biennial Symposium*, June 30–July 4, Sydney, Australia.
- Valderrama, J. T., de la Torre, A., Medina, C., Segura, J. C., and Thornton, A. R. D. (2016). "Selective processing of auditory evoked responses with iterative-randomized stimulation and averaging: A strategy for evaluating the time-invariant assumption," *Hear. Res.* **333**, 66–76.
- Woldorff, M. G. (1993). "Distortion of ERP averages due to overlap from temporally adjacent ERPs: Analysis and correction," *Psychophysiology* **30**, 98–119.