



## PAPER

## A neural network predicting the amplitude of the N2pc in individual EEG datasets

Francesca Marturano<sup>1</sup> , Sabrina Brigadoi<sup>1,2</sup> , Mattia Doro<sup>2</sup> , Roberto Dell'Acqua<sup>2,3</sup> and Giovanni Sparacino<sup>1,\*</sup> <sup>1</sup> Department of Information Engineering—DEI, University of Padova, Padova, Italy<sup>2</sup> Department of Developmental Psychology—DPSS, University of Padova, Padova, Italy<sup>3</sup> Padova Neuroscience Center, University of Padova, Padova, Italy

\* Author to whom any correspondence should be addressed.

E-mail: [gianni@dei.unipd.it](mailto:gianni@dei.unipd.it)**Keywords:** EEG/ERP, N2pc, artificial neural network, machine-learning, time-frequency analysis**Abstract**

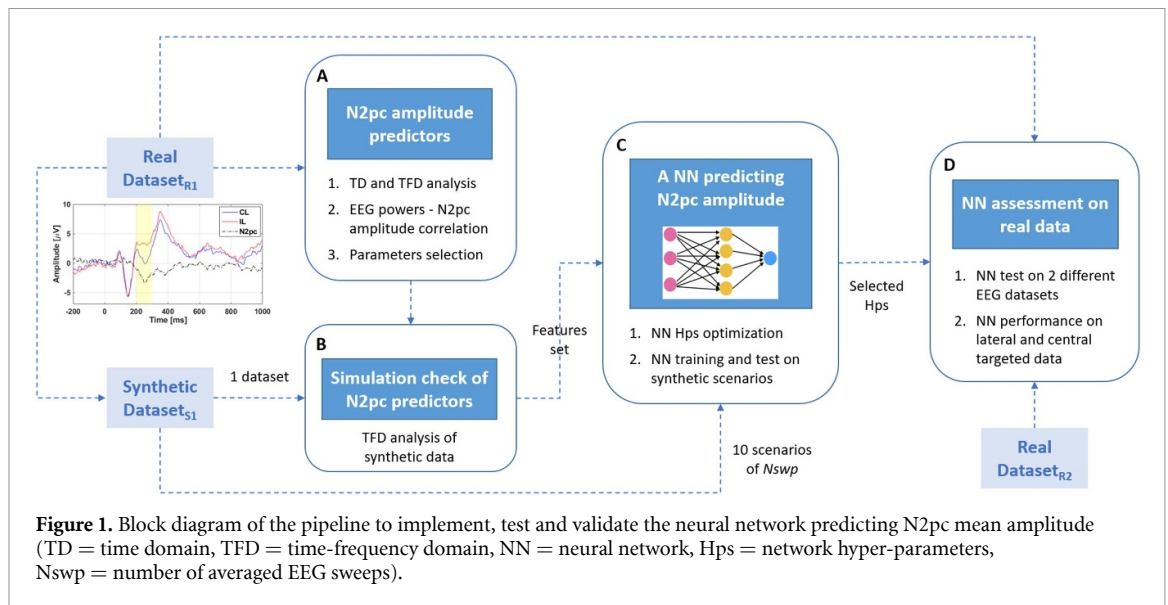
**Objective.** The N2pc is a small amplitude transient interhemispheric voltage asymmetry used in cognitive neuroscience to investigate subject's allocation of selective visuo-spatial attention. N2pc is typically estimated by averaging the sweeps of the electroencephalographic (EEG) signal but, in absence of explicit normative indications, the number of sweeps is often based on arbitrariness or personal experience. With the final aim of reducing duration and cost of experimental protocols, here we developed a new approach to reliably predict N2pc amplitude from a minimal EEG dataset. **Approach.** First, features predictive of N2pc amplitude were identified in the time-frequency domain. Then, an artificial neural network (NN) was trained to predict N2pc mean amplitude at the individual level. By resorting to simulated data, accuracy of the NN was assessed by computing the mean squared error (MSE) and the amplitude discretization error (ADE) and compared to the standard time averaging (TA) technique. The NN was then tested against two real datasets consisting of 14 and 12 subjects, respectively. **Main result.** In simulated scenarios entailing different number of sweeps (between 10 and 100), the MSE obtained with the proposed method resulted, on average, 1/5 of that obtained with the TA technique. Implementation on real EEG datasets showed that N2pc amplitude could be reliably predicted with as few as 40 EEG sweeps per cell of the experimental design. **Significance.** The developed approach allows to reduce duration and cost of experiments involving the N2pc, for instance in studies investigating attention deficits in pathological subjects.

**1. Introduction**

A popular practice to study mechanisms and time-course of covert visuo-spatial attention allocation in humans is to display multi-element visual arrays including a unique searched-for target in the left or right visual hemifield while observing event-related potentials (ERPs) recorded at posterior electrodes (usually P7/P8 [1, 2] or PO7/PO8 [3, 4]). A ubiquitous finding in so-called visual search tasks is that the voltage of ERPs recorded at electrodes contralateral (CL) to the visual hemifield occupied by the target is more negative than the voltage recorded at ipsilateral (IL) electrodes in a 200–300 ms time-window. This transient voltage asymmetry has been termed N2pc [5]. The amplitude of N2pc, classically

quantified by subtracting the IL from the CL activity, seldom exceeds 2  $\mu$ V [3, 6, 7]. This places N2pc closer to the lower end of the signal-to-noise ratio (SNR) scale, whereby the higher end is represented by other difference ERP components with amplitudes that can be almost one order of magnitude greater than N2pc, such as the frequency-related P3b (quantified by subtracting ERPs to frequent stimuli from ERPs to infrequent stimuli, e.g. [8, 9]), or the error-related negativity (ERN, quantified by subtracting ERPs recorded on trials associated with a correct response from ERPs recorded on trials associated with an incorrect response, e.g. [10]).

N2pc is commonly estimated by averaging a substantial number of sweeps of the electroencephalographic (EEG) signal, even hundreds [7], the only



apparent limit in data collection being human tolerance. In many neuroscientific studies, however, finding the right balance between number of sweeps and quality of the results is an issue, because long experiments with fast-pace event administration are mentally vexing and source of anxiety [11]. In a recent work relying on simulated data mimicking a common visual search task designed to evoke the N2pc, we demonstrated that there is no substantial improvement in detecting the presence of a reliable N2pc in group-average ERPs when more than 20–30 EEG sweeps for each CL and IL component of the N2pc are averaged, even in particularly noisy conditions [12]. This suggests that, at least for relatively simple experimental designs at the group-average level [13], reducing—rather than increasing—the number of EEG sweeps may result in an increment in data quality, by further mitigating the risk of deteriorating participants' performance because of the prolonged experimental duration.

In cognitive and clinical neuroscience, besides group studies, one may be also interested in the administration of quick tests exploring the residual attentive skills of a single subject with a unique or peculiar pattern of symptoms (e.g. [14, 15]). In this context, there is the need of a method able to reliably predict the individual N2pc amplitude from an EEG dataset acquired during short-duration tasks. In the present paper, an approach resorting to a neural network (NN) fed by features predictive of N2pc amplitude (identified from a time-frequency domain analysis of real EEG data) is developed and assessed against the standard time averaging (TA) technique.

Figure 1 describes the organization of the paper, the four blocks representing the four major steps entailed in the proposed methodology. The used datasets, including a simulated dataset, i.e. Dataset<sub>S1</sub>,

and two real datasets, i.e. Dataset<sub>R1</sub> and Dataset<sub>R2</sub> (consisting in 14 and 12 subjects, respectively), are described in section 2. Section 3 documents the first step of the procedure, which consisted in identifying which EEG time-frequency (TF) domain indicators could be predictive of N2pc amplitude in individual EEG datasets (block A in figure 1). More specifically, by using Dataset<sub>R1</sub>, we tested retrospectively the correlation stability between power in  $\delta$ ,  $\theta$ ,  $\alpha$  and  $\beta$  bands and N2pc amplitude while the number of individual EEG sweeps associated to laterally displayed visual targets and contributing to the CL and IL ERP components of N2pc was progressively reduced. The results of the analysis were clear-cut in showing that power at  $\delta$  and  $\theta$  bands were particularly stable predictors of N2pc amplitude in individual EEG datasets. Section 3 also describes how, as second step of the procedure, the synthetic Dataset<sub>S1</sub> was used to confirm that correlations between amplitude of the N2pcs and the TF parameters previously identified on the real dataset were stable also in simulated data (block B in figure 1). In section 4, the third step of the procedure is reported (block C). Specifically, the selected EEG features were used to train and test an artificial NN designed to predict the amplitude of synthetic N2pcs. Section 5 describes the fourth step (block D), which consists in assessing the NN on real data collected by administering to human adults common, albeit physically different, variants of a visual search design specifically devised to elicit an N2pc in response to lateralized targets embedded among distractors. After a preliminary check of the optimal NN proficiency in N2pc amplitude detection on Dataset<sub>R1</sub>, the final validation test was carried out on a previously unseen real EEG dataset labelled as Dataset<sub>R2</sub>. Finally, section 6 provides a discussion of methods and results, while section 7 ends the paper with some closing remarks.

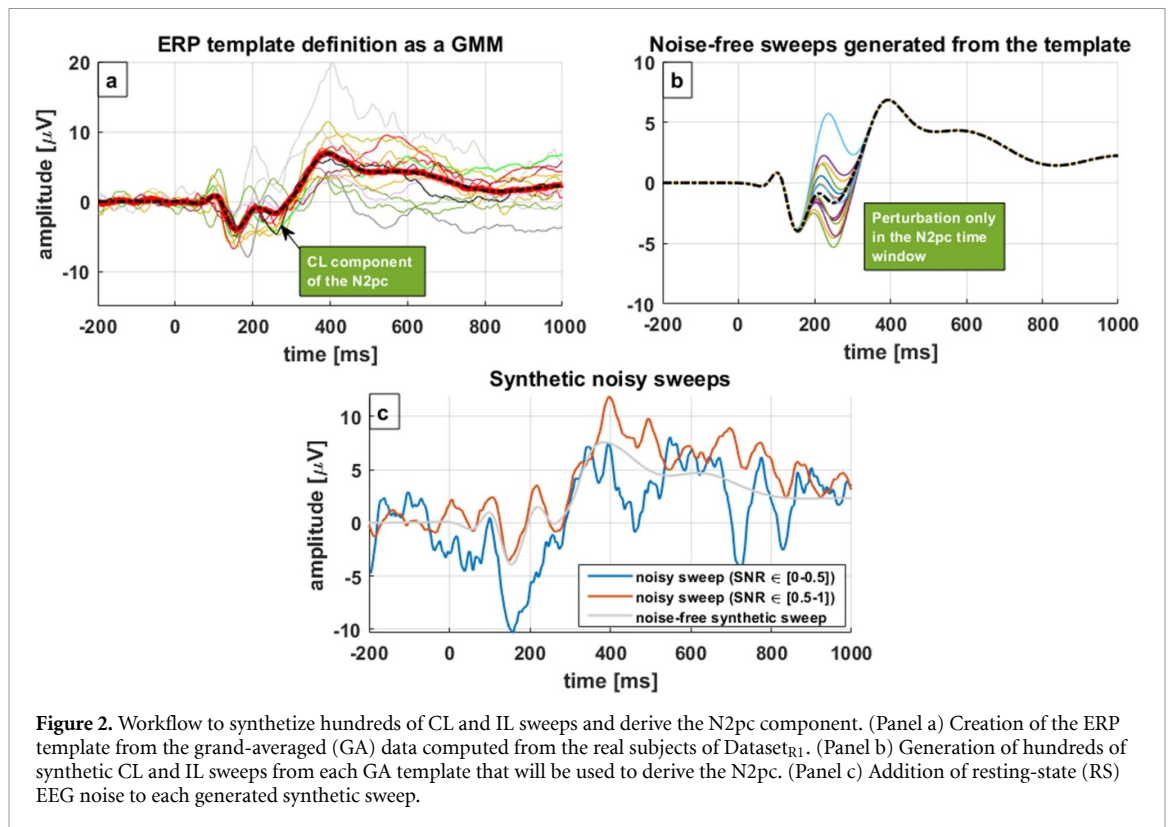
## 2. Datasets

Dataset<sub>R1</sub>, described in details in [12], included the data from 14 participants (mean age  $23.2 \pm 2.1$  years, five males) to a standard visual search task designed to elicit the N2pc. Participants were asked to report the orientation of a bar (either tilted or vertical) displayed in a colored circle, which was the visual target, surrounded by homogeneous distractors (all white circles). Each participant contributed 600 EEG sweeps, i.e. 200 EEG sweeps for each of CL and IL components of the N2pc (associated lateralized targets), and 200 EEG sweeps generated when targets were displayed along the vertical meridian. EEG data in [12] were registered at a sampling rate of 500 Hz from 28 electrodes positioned according to the international 10–20 system. Three additional electrodes were used to register eye movements [14, 15], and placed at the outer canthi and below the left eye. Following removal of EEG sweeps associated with incorrect responses and artifacts, EEG data were band-pass filtered (0.1–30 Hz). EEG activity recorded at P7 (left parietal hemisphere) and P8 (right parietal hemisphere) electrodes [2, 4] was segmented into 1600 ms sweeps, starting 600 ms before the stimulus onset. EEG sweeps were classified according to the target position in the visual field in right (R), left (L), or central (C)-targeted sweeps. EEG sweeps recorded in response to laterally displayed targets were further classified as CL and IL according to the relative position of electrodes and target side (i.e. P7 electrode was CL with respect to right targets and IL for left targets, and *viceversa* for P8). Six design-cell clusters of EEG sweeps resulted from this subdivision and were labelled P7-R, P7-L, P7-C, P8-R, P8-L and P8-C. Dataset<sub>R1</sub> was used both to develop the grand-average (GA) template for the generation of the synthetic dataset and as preliminary test of the NN performance on real data (figure 1, block D).

Dataset<sub>R2</sub> was described in detail in Doro *et al* [16]. Briefly, it was composed of the data from 12 participants (mean age  $31 \pm 6$  years, six males), each contributing 960 EEG sweeps, i.e. 320 EEG sweeps for each of CL and IL components for lateral targets and 320 EEG sweeps generated when targets were central (i.e. displayed along the vertical meridian). Participants were asked to identify and report the presence or absence of the target object (a coloured disk) among homogeneous distractor items (all grey disks). EEG data in [16] were registered at a sampling rate of 500 Hz from 27 electrodes positioned according to the international 10–10 system and were pre-processed offline using the same parameters as those used for Dataset<sub>R1</sub> with the exception that EEG was recorded at PO7 (left parieto-occipital hemisphere) and PO8 (right parieto-occipital hemisphere) electrodes [3, 4], rather than P7/P8 [1, 2], due to a different montage available. EEG activity was segmented into 1600 ms sweeps, starting 600 ms before the search array onset.

For the sake of labelling consistency across datasets, EEG sweeps were classified according to the target position in right (R), left (L), or central (C). EEG sweeps recorded in response to lateral targets were further classified as CL and IL according to the spatial arrangement of electrodes and targets' side (i.e. PO7 channel was CL for right targets and IL for left targets, and *viceversa* for PO8), resulting in six design-cell clusters of EEG sweeps, that were labelled PO7-R, PO7-L, PO7-C, PO8-R, PO8-L and PO8-C. Dataset<sub>R2</sub> was employed to further test the NN and validate it (figure 1, block D) using data measured in a different, albeit similar, experimental variant than the one of Dataset<sub>R1</sub>.

The synthetic Dataset<sub>S1</sub> was used to train and test in a controlled environment the NN on ten scenarios with a different numerosity in terms of CL and IL sweeps used to calculate the N2pc, with the aim to evaluate NN prediction performance as a function of the dataset dimension (figure 1, block C). In simulation it is possible to perform a wider and more precise quantitative evaluation of NN prediction error having a higher dataset dimension and a ground truth for the estimates, so that different scenarios can be tested. The synthetic dataset was generated using the simulator described in [12]. Briefly, the GA ERP template for each of the six design-cell clusters of Dataset<sub>R1</sub> was created by fitting a Gaussians mixture model (GMM) to the corresponding group-average ERP [17–20] (figure 2(a)). To create realistic synthetic data, we estimated the variability ranges for amplitude and latency on the CL and IL sweeps in Dataset<sub>R1</sub> and used these to constrain the variability of synthetic sweeps in the N2pc time range. For each design-cell cluster, synthetic sweeps were obtained by applying a random perturbation to each average CL and IL template simply by summing to the template two additional Gaussians with variable amplitude, variance, and latency in the N2pc time range only, so as to create arbitrary variable individual sweeps in (200–300) ms time window (figure 2(b)). Only synthetic sweeps with amplitude and latency values within the bounds previously estimated from Dataset<sub>R1</sub> were retained for analysis, to avoid misrepresenting synthetic data, and hence introducing possible biases in the results. Eventually, the noisy sweeps of Dataset<sub>S1</sub> were obtained by corrupting the generated noise-free synthetic traces of each design-cell cluster with real background EEG noise acquired from two additional participants (not included in Dataset<sub>R1</sub>) under resting-state (RS) conditions, to preserve the statistical properties of spontaneous EEG activity (figure 2(c)) and reduce the leakage between simulated and real data. In order to create EEG sweeps associated with distinct levels of SNR, the corrupting RS noise was modulated in amplitude and power by multiplying the original noise traces by a series of multiplicative factors (i.e. 0.25:0.25:1.25). Eventually, each of the six design-cell clusters of Dataset<sub>S1</sub>



contained 200 simulated EEG sweeps, 100 with  $\text{SNR} \in [0-0.5]$  and 100 with  $\text{SNR} \in [0.5-1]$ .

Synthetic N2pcs in Dataset<sub>S1</sub> were then derived through the standard procedure of subtracting an IL from a CL sweep, both selected at random from the pool of generated signals in the corresponding design-cell cluster. The described procedure was devised to limit the leakage of information between the parameters and the statistical properties of the generated N2pcs and those of Dataset<sub>R1</sub>, since only the CL and IL parameter bounds and the GA waveform needed to create the template—but then considered for the analyses in the restricted time range of the N2pc—were derived from the real dataset, whereas synthetic N2pcs were generated arbitrarily from the original templates.

*Remark.* A note is in order concerning EEG sweeps in design-cell clusters P7/8 C of Dataset<sub>R1</sub> and Dataset<sub>S1</sub> and PO7/8 C of Dataset<sub>R2</sub>, namely, EEG sweeps recorded on trials in which a searched-for target occupied a ‘central’ position in the search array (i.e. a position aligned to the vertical meridian of the screen used for stimuli presentation in the corresponding visual search experiment). These EEG sweeps, which can only arbitrarily be classified as CL or IL relative to electrode/scalp location for obvious reasons, are expected to give rise, when averaged and subtracted, to N2pc amplitude approximately equal to 0  $\mu\text{V}$ . As described more in detail below, ‘central’ EEG sweeps were useful for the present purposes to simulate individual cases of absence of N2pc,

which is one among the possible real cases that a properly trained NN designed to predict the N2pc amplitude ought to discriminate in individual EEG datasets.

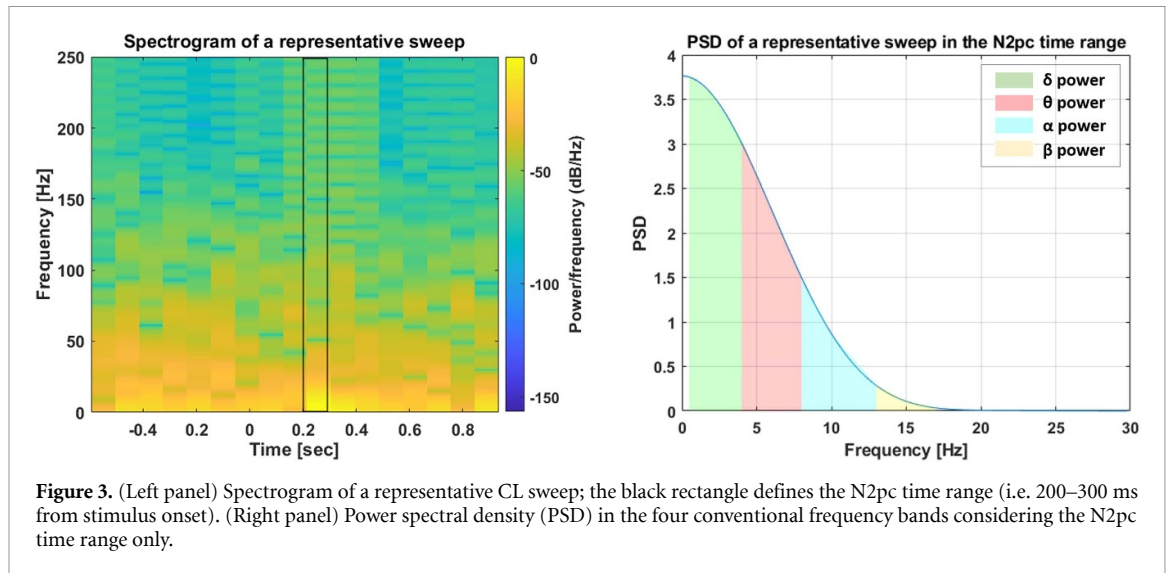
### 3. Time-frequency (TF) analysis for the identification of N2pc amplitude predictors

#### 3.1. Identification of TF predictors of N2pc amplitude using Dataset<sub>R1</sub>

We addressed our investigation toward TF domain indicators, given their well-established relationship with human cognition in general (e.g. [21]), and high-level visual processing in particular (e.g. [22, 23]). In particular, given the different amplitude scale of N2pc and EEG noise in the time domain, we extracted EEG markers predictive of the individual N2pc from the TF domain through an analysis of CL vs. IL variations in terms of power across the entire spectrum of frequency bands. Analyses were performed on Dataset<sub>R1</sub> (figure 1, block A) considering EEG sweeps recorded on trials with lateral targets only (i.e. disregarding EEG sweeps for central targets; see previous remark).

The spectrogram of each averaged signal was computed for every participant over windows of 50 samples with five overlapping samples and by applying the Hamming windowing to each averaged signal to reduce the discontinuities at the boundaries. Both the CL (obtained as the average of P7-R and P8-L),





**Table 1.** Correlation between N2pc power at each frequency band and N2pc mean amplitude obtained varying number and position (i.e. first, central or final) of the cluster of averaged EEG sweeps used to compute the CL and IL components of the N2pc. Bold-faced numbers are the significant correlations ( $p < .05$ ).

<i>f-band</i>	<i>All sweeps</i>	<i>First 40 sweeps</i>	<i>First 20 sweeps</i>	<i>First 10 sweeps</i>	<i>Central 40 sweeps</i>	<i>Central 20 sweeps</i>	<i>Final 40 sweeps</i>	<i>Final 20 sweeps</i>
$\alpha$	<b>−0.92</b>	<b>−0.90</b>	<b>−0.61</b>	<b>−0.58</b>	<b>−0.75</b>	<b>−0.82</b>	<b>−0.94</b>	<b>−0.81</b>
$\beta$	<b>−0.80</b>	<b>−0.78</b>	0.34	0.42	<b>−0.68</b>	<b>−0.59</b>	<b>−0.54</b>	−0.05
$\delta$	<b>−0.95</b>	<b>−0.94</b>	<b>−0.91</b>	<b>−0.92</b>	<b>−0.81</b>	<b>−0.87</b>	<b>−0.95</b>	<b>−0.92</b>
$\theta$	<b>−0.94</b>	<b>−0.93</b>	<b>−0.85</b>	<b>−0.88</b>	<b>−0.79</b>	<b>−0.86</b>	<b>−0.95</b>	<b>−0.91</b>

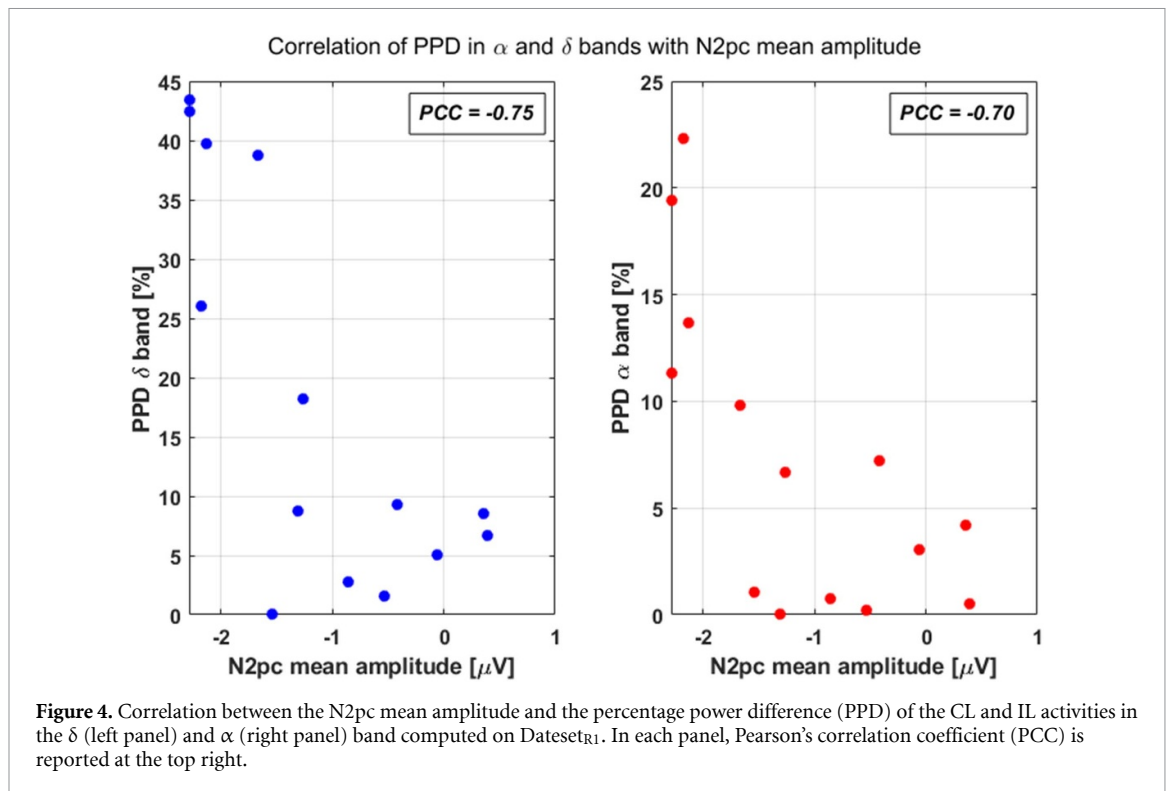
and the IL ERP (obtained as the average of P7-L and P8-R) as well as the corresponding N2pc of each subject were analyzed in the TF domain as described. We computed the absolute power at the standard frequency bands from each individual spectrogram within the N2pc time range (i.e. 200–300 ms) (figure 3) and correlated each power band with the mean amplitude of the corresponding N2pc.

Correlations between N2pc mean amplitude and EEG power bands were evaluated while progressively increasing the number of EEG sweeps considered in the average for each CL and IL component of the N2pc (i.e. by considering subsamples of 10, 20, 40 sweeps or all the sweeps available for each subject in the dataset), with the aim to evaluate whether and when correlations were stable. Furthermore, the position along the entire EEG recording of the sweeps included in the average was varied (i.e. sweeps were selected either during the initial, or middle or final phase of the visual search task) to check whether correlations were dependent on the temporal stint of the data acquisition.

The results of the correlation analysis are reported in table 1. When all EEG sweeps (mean  $\pm$  SD = 101  $\pm$  39 for IL data, 102  $\pm$  38 for CL, with a minimum of 41 sweeps) were considered in the average, a statistically significant correlation between N2pc amplitude and N2pc powers at all frequency bands was observed. Correlations

remained statistically significant even when using the minimum number of EEG sweeps available across all subjects and channel-target design-cell clusters, i.e. 40 sweeps. When the number of sweeps was reduced below 40, the correlation between  $\alpha$  and  $\beta$  bands and the N2pc amplitude started to decrease, with  $\beta$  band becoming uncorrelated when less than 20 sweeps were used, possibly implying that high-frequency bands started to be affected by noise. Moreover, for the low-frequency power bands, i.e. the  $\delta$  and  $\theta$  band, the observed correlations were statistically significant independently of the temporal location of the considered ensemble of sweeps inside the subject's recording (table 1). This result seems to suggest that low-frequency band powers in the N2pc time range correlate with N2pc amplitude stably and could therefore be good predictive features to be fed to an NN designed to predict the amplitude of N2pc in individual EEG datasets using a reduced number of EEG sweeps.

Similar results were obtained when evaluating the relationship between the powers at the different frequency bands for the CL and IL average activities and the N2pc mean amplitude. In particular, higher amplitude N2pcs showed a peculiar difference between power values for CL and IL activities, particularly in the  $\delta/\theta$  and  $\alpha$  band, whereas this difference decreased to nil as the N2pc amplitude approximated zero (see [24]).



In figure 4, for each subject, the absolute difference between the power of the CL and IL components for both  $\delta$  ( $\theta$  band was similar) and  $\alpha$  band, computed in percentage of the maximum power value between the corresponding CL and IL power distributions (defined as percentage power difference,  $PPD = \frac{\text{abs}(\text{CL} - \text{IL})}{\text{max}(\text{CL}, \text{IL})} \times 100$ ), is reported as a function of the individual N2pc mean amplitude. Both  $\alpha$  and  $\delta$  power and N2pc mean amplitude were calculated by considering the maximum number of available sweeps for each subject to reduce to the minimum the influence of EEG noise. A significant Pearson's correlation coefficient (PCC) resulted for both power bands (PCC =  $-0.75$  for the  $\delta/\theta$  bands and PCC =  $-0.7$  for the  $\alpha$  band). This result indicates that also CL/IL power bands could be new important predictors to be fed to the NN.

### 3.2. Evaluation of the stability of TF predictors on dataset<sub>S1</sub>

Before training and testing the NN with synthetic data fully resembling a real scenario, we had to control whether the correlation results observed on the real EEG data also held on data entirely simulated in the time domain, without accounting for subtle synchronization of phases that might run in the background (figure 1, block B).

To accomplish this goal, Dataset<sub>S1</sub> was used. From the ensemble of synthetic noisy traces, 100 random sweeps were selected for the CL and IL components of both P7 and P8 electrodes and averaged to compute the corresponding N2pc. The choice of averaging such a high number of EEG sweeps is due to the

need of reducing to the minimum the noise level for synthetic data. The procedure was repeated 10 times, thus obtaining 140 average N2pcs with random amplitude and latency values used for this second TF analysis. Synthetic N2pcs underwent the same procedure described for the real data.

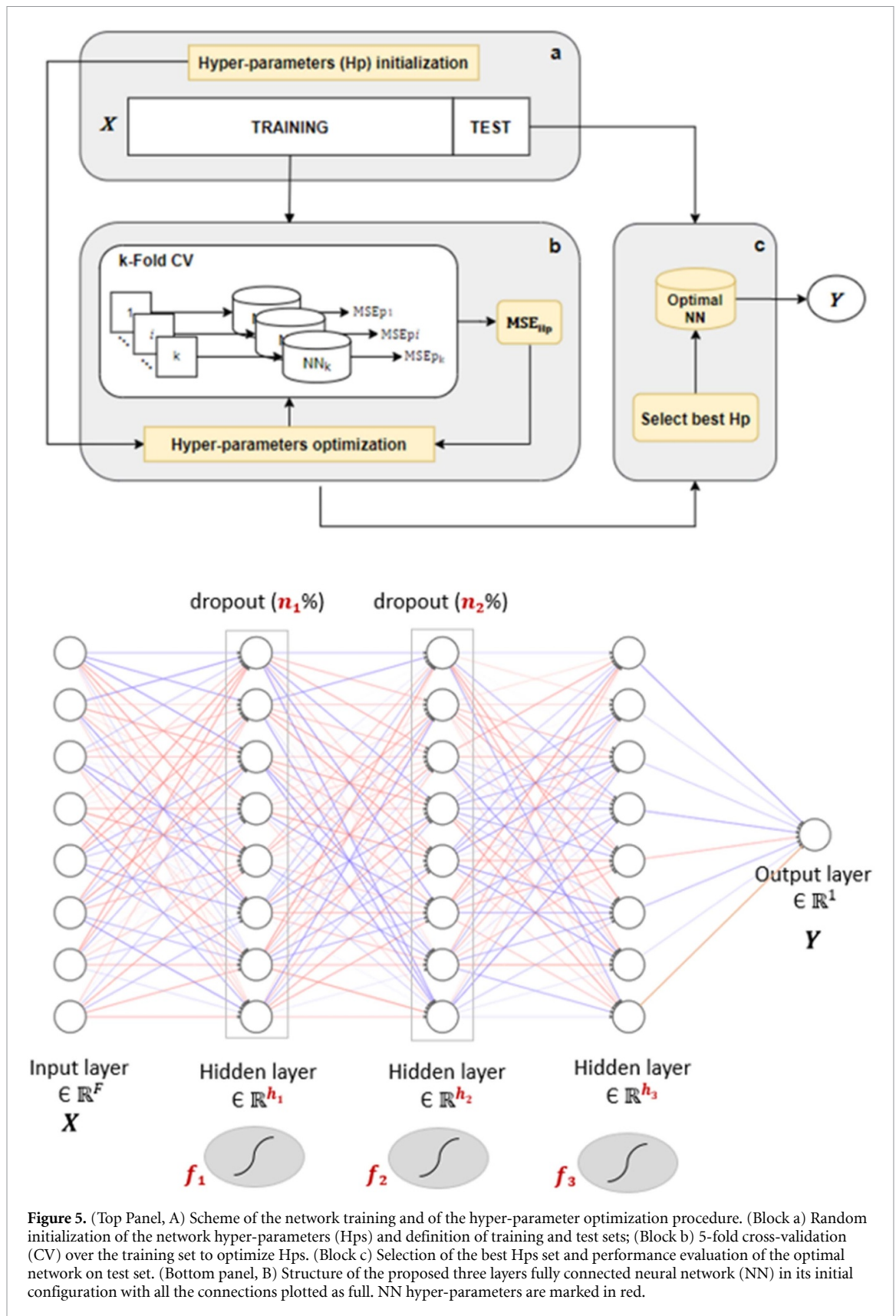
Correlation between low-frequency band powers and N2pc mean amplitude was statistically significant (PCC =  $-0.94$  for both the  $\delta$  and  $\theta$  band), mirroring the results of real data. By grouping powers according to N2pc mean amplitude ( $A_{\text{N2pc}}$ ), i.e. whether  $A_{\text{N2pc}} < 0.5$ ,  $0.5 \leq A_{\text{N2pc}} < 1$  or  $A_{\text{N2pc}} \geq 1$ , a clear segregation emerged, above all, between the first and last group.

## 4. Development of neural network (NN) to predict N2pc mean amplitude

In the present section, the different steps schematically summarized by block C of figure 1 are described in depth.

### 4.1. Overview of NN structure and testing

The rationale of NN implementation, as well as the estimate of the average amplitude of synthetic N2pcs under different scenarios varying sweeps numerosity, is described in detail with reference to the blocks (from 'a' to 'c') of the scheme illustrated in figure 5, panel A. Briefly, ahead of the NN training phase, each synthetic scenario was split into training (80% of the samples) and test (20% of the samples) set (figure 5(A), block a). For each scenario, the training set was used firstly to optimize the



**Figure 5.** (Top Panel, A) Scheme of the network training and of the hyper-parameter optimization procedure. (Block a) Random initialization of the network hyper-parameters (Hps) and definition of training and test sets; (Block b) 5-fold cross-validation (CV) over the training set to optimize Hps. (Block c) Selection of the best Hps set and performance evaluation of the optimal network on test set. (Bottom panel, B) Structure of the proposed three layers fully connected neural network (NN) in its initial configuration with all the connections plotted as full. NN hyper-parameters are marked in red.

NN hyper-parameters (Hps) during a cross validation (CV) procedure, secondly to define the best NN structure (figure 5(A), block b) and finally to train the optimal NN (figure 5(A), block c). Afterwards, N2pcs belonging to the test set of the corresponding scenario

were used to score NN performance in amplitude prediction. NN implementation, as well as training and test phases, were entirely performed in Python (Python Software Foundation. Python Language Reference, version 3.6. Available at [www.python.org](http://www.python.org)).

**Table 2.** Percentages of synthetic N2pc sweeps in each scenario of CL and IL sweeps numerosity and in each N2pc amplitude range, considering both training and test sets.

Amplitude ( $A_{N2pc}$ )	10 sweeps	20 sweeps	30 sweeps	40 sweeps	50 sweeps	60 sweeps	70 sweeps	80 sweeps	90 sweeps	100 sweeps
$A_{N2pc} < 0.5$	35%	35%	37%	37%	37%	38%	38%	38%	37%	37%
$0.5 \leq A_{N2pc} < 1$	19%	19%	18%	17%	17%	17%	18%	16%	18%	18%
$A_{N2pc} \geq 1$	46%	46%	45%	46%	46%	45%	44%	46%	45%	45%

#### 4.2. Hyper-parameter (Hp) optimization

A feedforward fully connected NN was used to estimate the amplitude of simulated N2pcs. The selected NN included an input layer of dimension  $F$ , whose neurons received one of the  $F$  final features that will be discussed in depth in section 4.4, three hidden layers (HLs), a one-dimensional output layer (with linear activation function) and two additional dropout layers. Dropout layers are typically used in NNs implementation for regularization purposes and to prevent overfitting. The rationale of dropout layers is to randomly drop out (i.e. set to zero), at each training iteration, some neurons (and the related connections) of the layer right before the dropout layer [25] to optimize the network structure starting from its initial configuration visible in figure 5, panel B. The fraction of neurons that are set to zero during the training phase is called dropout percentage [26]. In our setting, one dropout layer was interposed between the first two HLs, while the other was between the second and the last HL.

The size of the three HLs, the activation function of the input layer and of the three HLs, the dropout percentages and the mini-batch size were the hyper-parameters (Hps) of the network, necessary to determine the optimal structure. Hps were optimized on training set data during a 5-folds CV procedure, following the rationale of reaching a compromise between capability of fitting and ability to generalize to new datasets (figure 5(A), block b). Trees of Parzen estimators (TPE) technique was used to solve Hps optimization and predict the set of Hps with the best mean squared error (MSE). At each iteration, the algorithm analyzes a new Hps set and decides the following Hps to evaluate. During CV, the training set is split into five folds, four of which are used for training the NN and one is left for validation. The average MSE is computed for each set of Hps across folds and the one with lower  $MSE_{Hp}$  defines the optimal network structure, which is eventually ran on test data. A scheme of the described NN structure is illustrated in figure 5, panel B, in its fully connected configuration before starting the training. In the figure, the main network Hps are indicated in red.

#### 4.3. Synthetic datasets to train and test the NN

Ten different synthetic scenarios varying sweeps numerosity, i.e. from 10 to 100 with a step of ten sweeps, were created to train and test the artificial NN by randomly selecting sweeps from  $Dataset_{S1}$

described in section 2 (figure 1, block C). Each scenario counted  $N = 2720$  samples (i.e. averaged CL, IL and N2pc sweeps); 80% of these were used for the training and 20% for the testing of the NN. Moreover, each scenario comprised the same number of averaged CL and IL sweeps for lateralized targets, averaged central target sweeps (obtained by averaging P7-C with P8-C) and hence averaged N2pcs for both lateralized and central targets. For lateralized targets, N2pc sweeps in each scenario were computed conventionally by subtracting the IL activity from the CL activity. For central-target sweeps, fictitious N2pcs were computed by subtracting the activity at P8 (P7) from the one at P7 (P8), by randomly selecting each time the sign of the subtraction, that is, the CL and IL sweeps were selected from P7-C or P8-C ensembles at random and with equal probability. This procedure should produce an N2pc close to zero. Each scenario was created so that the N2pc sweeps comprised in the dataset could be partitioned in the three amplitude ranges discussed in the previous section (table 2).

The ten scenarios only differed in the number of EEG sweeps (labelled as  $N_{swp}$ ) used to calculate the average CL and IL components of the N2pc. The number of EEG sweeps in each scenario was incremented with a step of 10 sweeps per target side at a time (and 20 per CL and IL component), starting from a minimum of 10 for the first scenario up to a maximum of 100 sweeps for the last one. For example, the average CL component in the dataset corresponding to  $N_{swp} = 100$  was obtained by averaging 200 random sweeps in total, 100 taken from P7-R and 100 from P8-L (both groups including the CL signals of P7 and P8 electrodes). To enlarge the dataset this procedure was repeated many times.

A different NN was trained and tested considering each of the ten different simulated scenarios and their performances were compared to establish the minimum value of EEG sweeps numerosity achieving acceptable prediction error, hence the optimal NN to be finally tested on real experimental data.

#### 4.4. NN input features

The NN was fed with the selected TF-based features, i.e. the  $\alpha$ ,  $\beta$ ,  $\delta$  and  $\theta$  power bands of both the CL and IL ERP and of the resulting N2pc, with additional information related to the noise content of the signals into play derived from the time domain, represented by the SD of their pre-stimulus. Moreover, the variance of the N2pc within its time window



**Table 3.** List of the features (*F*) fed to the NN (SD = standard deviation, BL = baseline, VAR = variance, CL = contralateral component, IL = ipsilateral component).

Features ( <i>F</i> )	$\alpha, \beta, \delta, \theta$	$\alpha, \beta, \delta, \theta$		SD <sub>BL</sub> (N2pc)	SD <sub>BL</sub> (CL)	SD <sub>BL</sub> (IL)	VAR (N2pc time range)	CL-IL (N2pc time range)
	power (N2pc)	power (CL)	power (IL)					

(i.e. 200–300 ms) was also considered among the features. This parameter strengthened the set of features by carrying information about the noise level in the time window of interest, hence improving the final NN estimation performance, especially for N2pcs associated to midline targets. To account for inverted (namely, positive) N2pcs, whose behavior in the four power bands was very similar to negative N2pcs, we considered as a new input feature the difference between the average CL and IL signal in the N2pc time range. Indeed, on synthetic data, we observed that when the difference between CL and IL ERP was greater than  $0.5 \mu\text{V}$  in the N2pc time window, this metric succeeded in predicting the actual sign of the N2pc: the overall error was equal to 2.8% when averaging 100 EEG sweeps to compute the N2pc and to 7.1% with 40 EEG sweeps only, independently of the SNR range. The final number of features was  $F = 17$  (table 3).

#### 4.5. Assessment of the NN for N2pc mean amplitude estimation

##### 4.5.1. Metrics of performance

NN training was performed through a gradient descent RMSprop algorithm applied in a mini-batch mode, while NN performance on training and validation set (25% of the training samples) was evaluated through MSE and loss function. On test set, the percentage MSE, henceforth abbreviated as MSEp, was instead considered to score and compare network performances among the ten different synthetic scenarios so as to weight for the typical low amplitudes of the N2pc component.

As additional error metric, we computed the amplitude discretization error (ADE) of the NN estimates by discretizing the estimated amplitudes through a reasonable threshold at  $0.5 \mu\text{V}$ , considered as the minimum amplitude value distinguishable from the background EEG. Amplitudes were classified into two classes, that is, whether higher or lower than the selected threshold, indicating whether the individual N2pc was present (and well visible) or not in the considered data, respectively. Then, to compute ADE for each amplitude class, the number of misclassified samples was considered in percentage of the total number of samples of that class.

For comparison purposes, for each synthetic scenario of sweeps numerosity, we compared NN estimates on test set with the temporal averaging (TA)—in the N2pc time window—of the mean ERP computed from the corresponding scenario. TA is a very fast and old-fashioned approach conventionally used

to easily estimate the parameters of an ERP component from an averaged noisy signal. In the synthetic scenarios, by comparing NN results with those of TA we can quantitatively compute the estimation error of both approaches, while in real scenarios, we can only observe whether there is similarity/difference between them in the estimate of N2pc amplitude.

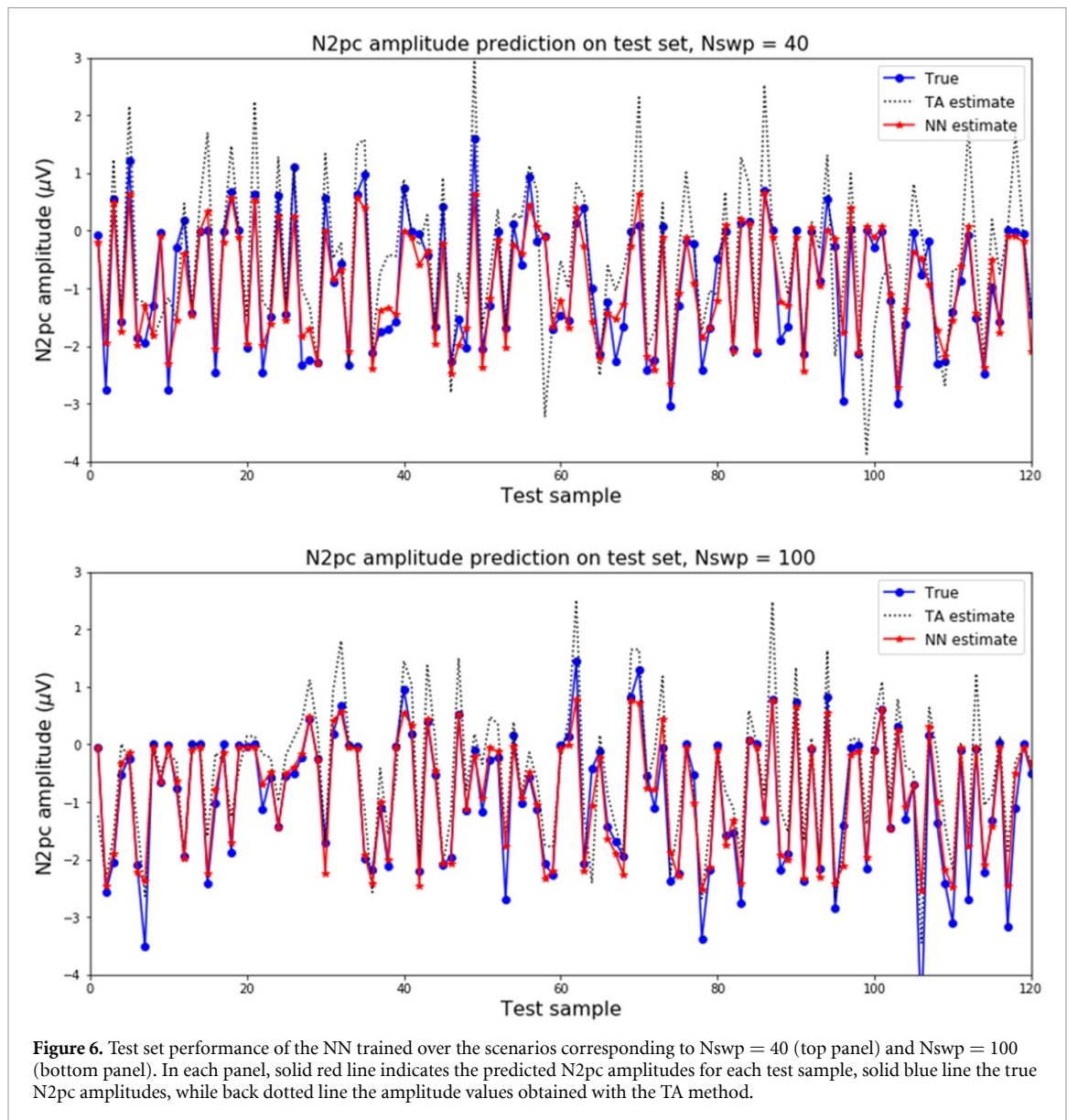
##### 4.5.2. Optimal NN structure and CV results

The optimal NN structure was the same in each synthetic scenario and included the input layer, three dense hidden layers, the output layer, and two dropout layers. The first layer was composed of 17 nodes corresponding to the 17 input features and had sigmoid activation function, the following three layers had (dimension, activation function) of (64, sigmoid), (128, sigmoid), and (32, linear), respectively. The last layer presented 1 node with linear activation function and was meant to perform a linear regression to estimate continuous N2pc amplitudes. The dropout percentages of the two interposed layers were 0.727 and 0.763, respectively, while the optimal mini-batch size was equal to 32. In total 13 939 parameters were trained in a 5-fold CV procedure to select the best NN structure.

The final loss of the selected NN was evaluated over 350 epochs in a CV procedure. The MSE of the NN corresponding to the scenario with  $N_{\text{swp}} = 40$ , that was selected to be finally tested on real data, was equal to 0.21 for the validation set—including the 25% of training samples selected at random—and to 0.28 for the training set.

##### 4.5.3. Assessment on simulated data

Figure 6 shows two representative examples of the NN amplitude estimates obtained in the synthetic scenarios corresponding to  $N_{\text{swp}} = 40$  and  $N_{\text{swp}} = 100$  EEG sweeps (top and bottom panel, respectively). The MSEp of the former scenario is equal to 10.7% for NN estimates vs. 58.4% for TA estimates; instead, the MSEp of the latter—considered as a gold-standard scenario in terms of noise level—is 4.3% for the NN vs. 26.3% for TA. Not surprisingly, by increasing the number of EEG sweeps considered in the average, the noise level is significantly reduced by averaging effect, so that NN estimates improve and the MSEp gets lower and lower. Conversely, the higher the number of EEG sweeps, the more the estimates get closer to those of TA (figure 6, bottom panel) and the gain of the NN becomes negligible. Despite the evident decrease in the error as  $N_{\text{swp}}$  approaches its maximum value of 100, it is necessary to establish



**Figure 6.** Test set performance of the NN trained over the scenarios corresponding to  $N_{swp} = 40$  (top panel) and  $N_{swp} = 100$  (bottom panel). In each panel, solid red line indicates the predicted N2pc amplitudes for each test sample, solid blue line the true N2pc amplitudes, while back dotted line the amplitude values obtained with the TA method.

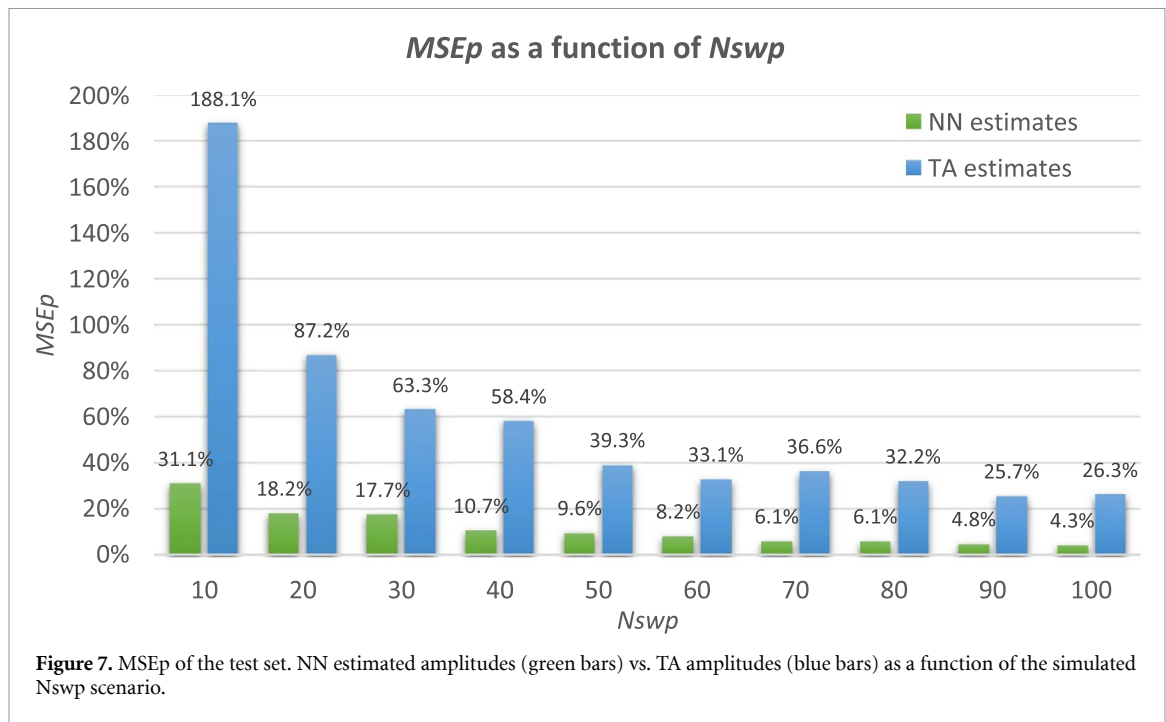
a trade-off between dimension of the dataset and estimation accuracy.

Figures 7 and 8 illustrate the overall network results on test set as a function of the synthetic  $N_{swp}$  scenario both in terms of MSEp and ADE, respectively. In figure 7, the MSEp obtained with the NN in each scenario (green bars) is compared to the corresponding MSEp of the TA method (blue bars). For each scenario, the maximum SD of the MSEp was of about 1%–2%.

Figure 8 instead reports, for each synthetic scenario, the values of ADE in two amplitude classes, separated by a fixed threshold on the N2pc amplitude, and the global network error, both for the NN and the TA method (panel A and B, respectively). Here we used  $0.5 \mu V$  as a reasonable threshold for discriminating datasets with a well visible N2pc response from datasets with a very little or absent component. Notably, the global error reported in the figures can be considered as the potential error achieved when

using the NN to discriminate presence or absence of the N2pc response with respect to a threshold on the estimated amplitude. Panel A of figure 8 shows that the global error for the NN is below 15% when averaging more than  $N_{swp} = 40$  EEG sweeps. By increasing  $N_{swp}$  up to 100 EEG sweeps, the error is lowered to a minimum of about 4%. Instead, the ADE for the TA method is worse and equal to 32% and 26% for  $N_{swp} = 40$  and 100, respectively (figure 8, panel B), thus implying a gain of the NN of around 20% for both scenarios. In terms of MSEp, the NN gain for the same scenarios is approximately 48% and 22%, respectively.

Table 4 shows, for each dataset, the true positive rate (TPR), i.e. the percentage of samples with an actual N2pc response which are correctly classified, and the false positive rate (FPR), i.e. the percentage of samples with a visible N2pc which are incorrectly classified, for both NN (top two table rows) and TA (bottom two table rows). Results highlight that,



despite the two methods perform similarly in terms of TPR (or sensitivity), they substantially differ for the FPR (or 1-specificity). FPR is considerably higher for TA than for NN estimates, suggesting that the former has a lower specificity. The presented outcomes show that the NN correctly predicted the presence/absence of N2pc in 85% of individual cases when fed with only 40 sweeps per each cell of the experimental design, with a significant improvement with respect to the TA method, both in terms of accuracy and specificity.

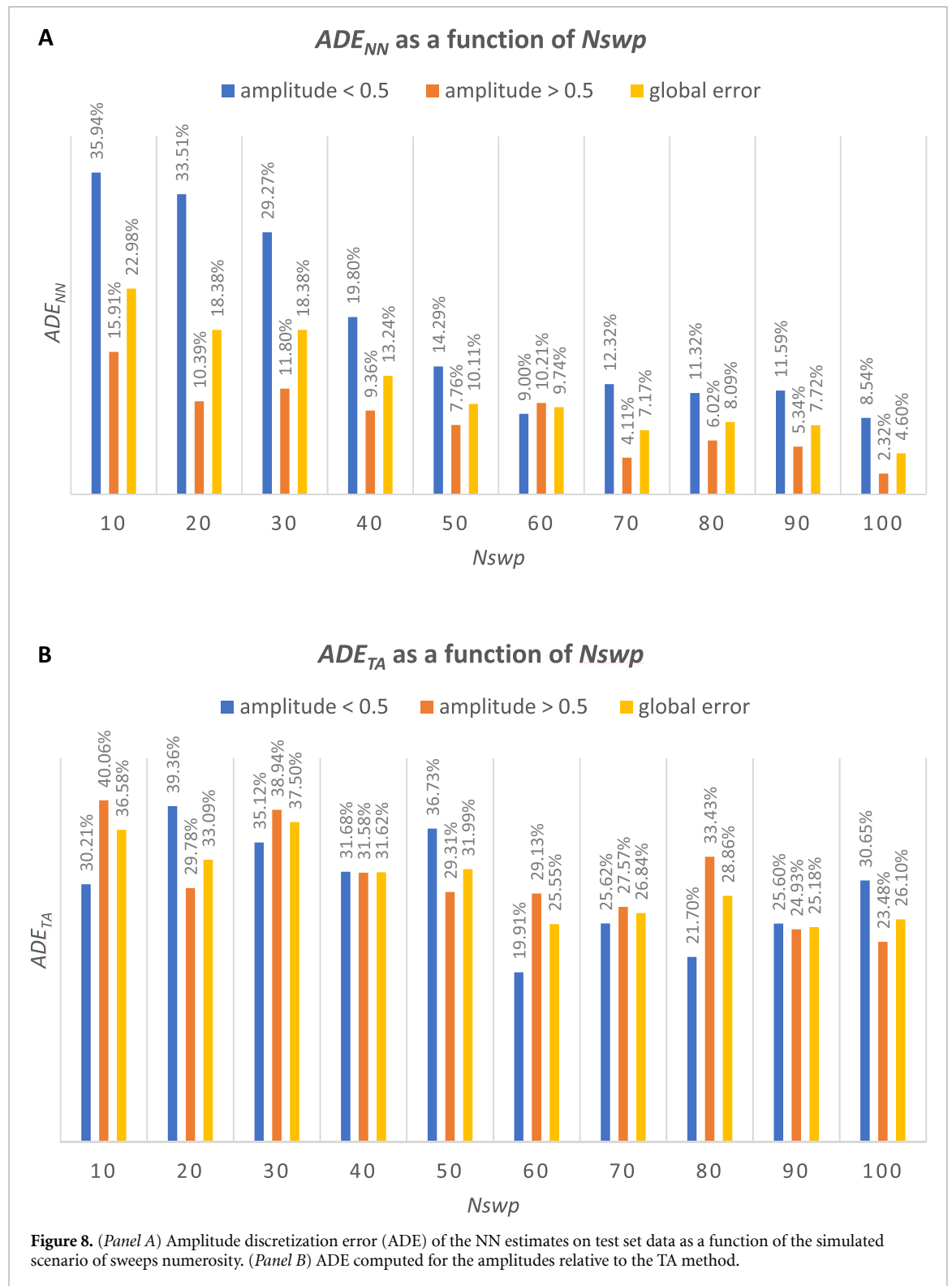
## 5. Assessment of the NN on real data

As depicted in block D of figure 1, the NN (trained over the synthetic scenarios with  $N_{swp} = 40$ ) was finally tested on the two real datasets, Dataset<sub>R1</sub> and Dataset<sub>R2</sub>, considering, for each subject, 40 EEG sweeps per channel-target design-cell cluster (except of one subject who had only 37 sweeps per cluster). In both real datasets, each NN was first evaluated on N2pcs associated to lateralized targets (i.e. right and left visual stimuli), where we expected to observe a significant response, and then on N2pcs computed from trials associated to central targets (i.e. placed on the vertical midline), to test for the NN capability of recognizing reduced or absent N2pc responses. For each subject in the real datasets, NN predictions were compared to the N2pc amplitudes estimated with TA method, averaging the same number of EEG sweeps used for the NN, to evaluate the two methods within the same noise scenario.

Amplitude predictions for both the selected NN and TA and for each real subject are shown in figure 9, left two panels for the lateral targets and right two

panels for the midline ones. Panels A and B are instead related to the specific dataset. Of note, on real data, given the absence of ground truth amplitudes, we could neither compute the exact NN and TA errors, nor we could make any quantitative comparison between the estimates obtained with two methods, to draw any conclusion regarding which method performs better. However, we can discuss the different performance of the two approaches also considering the estimates obtained for the N2pcs associated to midline targets, which we expect, theoretically, to be close to zero. These results are reported in the right two panels of figure 9 for the NN and TA; panel A is related to Dataset<sub>R1</sub>, while panel B to Dataset<sub>R2</sub>. These results suggest that the NN seems to outperform TA by more efficiently recognizing fake N2pcs associated to midline targets, even when fed with 40 sweeps only: amplitude estimates were close to zero for most of the real subjects, as expected, both for Dataset<sub>R1</sub> and (especially) for Dataset<sub>R2</sub>. TA amplitudes, estimated using the same number of sweeps fed to the NN, were instead more influenced by background noise compared to the NN, thus supporting the speculation about the superiority of the proposed method for the prediction of N2pc amplitude from datasets of small dimension.

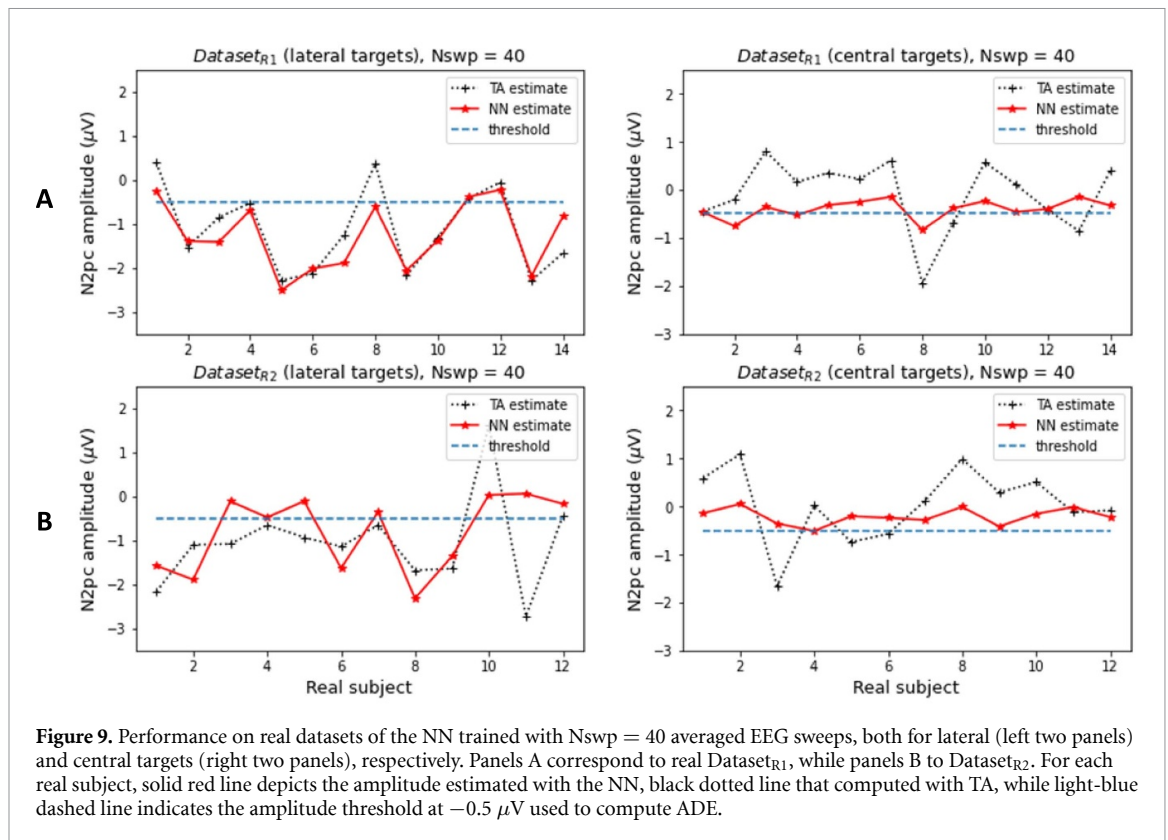
In conclusion, these results agree with simulation outcomes, supporting the general finding that as few as 40 EEG sweeps per cell of the experimental paradigm, that is a total of 80 sweeps for the CL and 80 for the IL component, could be fed to an NN to predict individual N2pcs with good reliability, at least for experimental variants and settings similar to those described here.



**Table 4.** True positive rate (TPR) and false positive rate (FPR) for the N2pc amplitudes estimated with NN and TA as a function of the synthetic  $N_{swp}$  scenario.

		10 sweeps	20 sweeps	30 sweeps	40 sweeps	50 sweeps	60 sweeps	70 sweeps	80 sweeps	90 sweeps	100 sweeps
NN	TPR	0.80	0.82	0.82	0.88	0.92	0.94	0.93	0.93	0.93	0.95
	FPR	0.29	0.20	0.20	0.16	0.14	0.16	0.07	0.09	0.09	0.04
TA	TPR	0.84	0.79	0.79	0.81	0.79	0.87	0.85	0.86	0.84	0.82
	FPR	0.73	0.56	0.64	0.53	0.52	0.46	0.46	0.52	0.41	0.41





**Figure 9.** Performance on real datasets of the NN trained with  $N_{swp} = 40$  averaged EEG sweeps, both for lateral (left two panels) and central targets (right two panels), respectively. Panels A correspond to real Dataset<sub>R1</sub>, while panels B to Dataset<sub>R2</sub>. For each real subject, solid red line depicts the amplitude estimated with the NN, black dotted line that computed with TA, while light-blue dashed line indicates the amplitude threshold at  $-0.5 \mu V$  used to compute ADE.

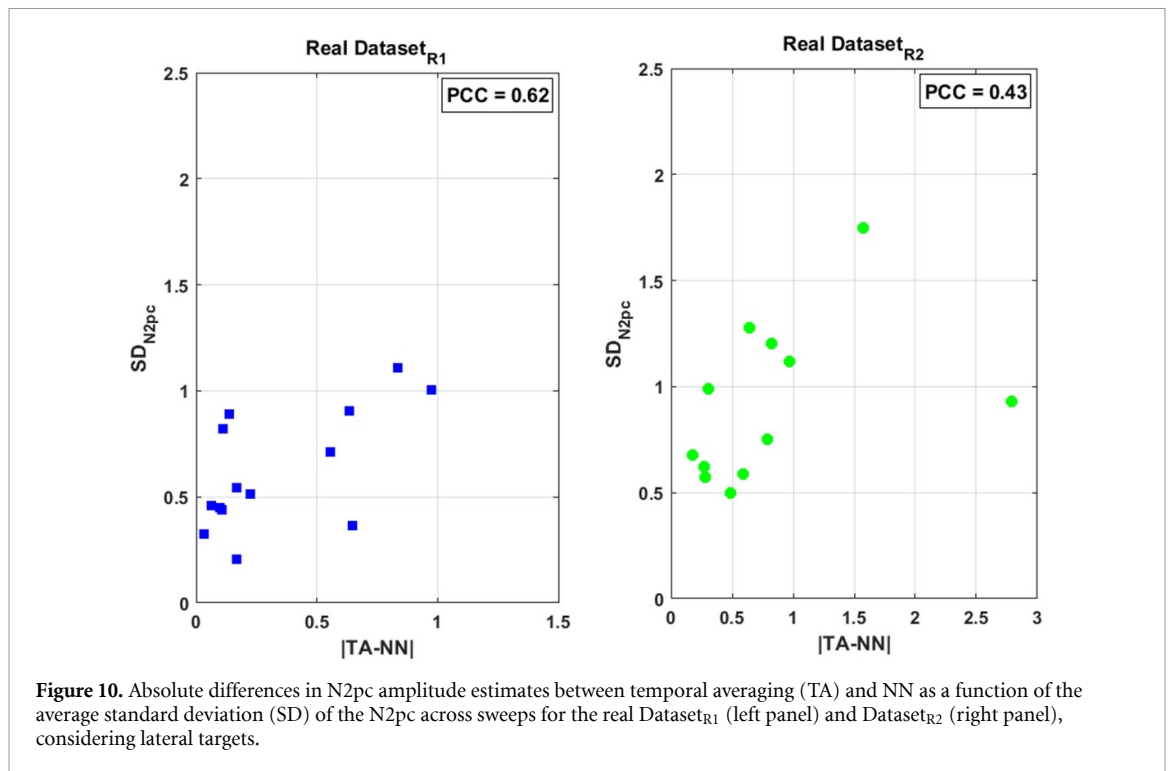
## 6. Discussion

In the present manuscript, we pursued the goal of predicting N2pc amplitude at the individual level, to be able to characterize patient's efficiency in the deployment of visual attention, and simultaneously reducing the number of EEG sweeps required for its estimation. To individuate parameters predictive of the investigated component, we first performed a preliminary analysis exploring time and TF domain of both N2pc and its CL and IL components. This analysis, carried out on the real experimental Dataset<sub>R1</sub>, revealed a high and stable correlation between the mean amplitude of the N2pc and the low-frequency EEG power bands. In addition, the pattern of the CL and IL ERP in the  $\alpha$  and  $\delta/\theta$  bands was predictive of the N2pc presence within the dataset. To confirm the stability of these features across datasets, we performed the same correlation analysis between N2pc mean amplitude and EEG powers also for the Dataset<sub>R2</sub>. The PCC resulted equal to  $-0.78$  for the  $\delta$  band and to  $-0.73$  for the  $\theta$  band, thereby confirming the promising use of TF-based features for predicting the N2pc on a different dataset.

We implemented an NN-based estimation tool driven by EEG power bands and other time-related features predictive of the noise content in the data to estimate N2pc mean amplitude and discriminate subjects with a well visible N2pc from those with a little to absent response, hence participants who efficiently allocate their attention to the cued target from those who do not. From Dataset<sub>S1</sub> we created

a realistic simulation scenario wherein to test the NN in the estimation of the N2pc mean amplitude while progressively increasing the number of EEG sweeps in order to define both the best NN and a lower bound for the EEG dataset dimension. Eventually, the selected NN was validated on real N2pc traces derived from two different datasets and considering signals associated both to lateral and central targets. This multiple validation strategy was intended to strengthen the validity domain of the developed NN tool and to generalize the present outcomes to different real contexts. Of note, the use of a different, albeit similar, experimental variant from that of Dataset<sub>R1</sub>, i.e. Dataset<sub>R2</sub>, for the final testing of the NN further reduced the risk of overfitting.

In the realistic synthetic scenario, the NN was able to predict N2pc mean amplitude with accuracy of around 85% from EEG datasets consisting of only 40 sweeps per each cell of the experimental design (i.e. right, left, and central target), hence obtaining good balance between dataset dimension and prediction accuracy. With our approach, the typical dimension of an EEG experiment designed to measure the N2pc would be reduced to more than 50%. Note, for comparison, that real Dataset<sub>R1</sub> and Dataset<sub>R2</sub> counted 200 and 160 sweeps per each cell of the experimental design, respectively. However, it is worth noting that our aim was not to establish the best number of EEG sweeps required to compute the N2pc, but to demonstrate that, with an NN approach, an N2pc could be reliably estimated using a minimal dataset.



The standard TA method was chosen to compare and evaluate the prediction performance of the NN. In simulation, the NN outperformed TA both when a high and a low number of sweeps was used, with a greater improvement in performance as the number of sweeps was reduced. The NN scored a decrease of about 50% for the MSE<sub>p</sub> and 15% for the global ADE compared to the TA method when trained with 40 sweeps and of about 20% for the MSE<sub>p</sub> and 20% for the global ADE when trained with 100 sweeps. On real data, since the ground truth is unknown, we cannot quantify the estimation error for either the NN or the TA method and results can be only qualitatively discussed. However, given the results in the simulated scenario—especially those for  $N_{\text{swp}} = 40$  (top panel of figure 6)—and those obtained for central targets on real data, where we expected a theoretical amplitude of zero for the N2pc, we can speculate that it is likely that the most correct results are those of the NN. The discrepancy between the methods observed for some subjects, especially in Dataset<sub>R2</sub> for lateral targets, could be attributable to other aspects that, for the sake of simplicity, have not been considered in this preliminary study but that will be investigated in the future. For example, we did not analyze the effect of noise level—which more likely affects TA estimates than NN ones, since the features fed to the NN include information about noise by means of the pre-stimulus SD—or the variability of the N2pc across sweeps. This speculation is supported by a further analysis we have run, which shows that the discrepancy between NN and TA increases as the average standard deviation across sweeps in the N2pc time

range increases (figure 10), pointing to a strong role of noise in the divergence of results.

A potentially critical aspect of the methodology illustrated in this paper concerns the possible leakage of information between synthetic Dataset<sub>S1</sub> and real Dataset<sub>R1</sub>. We presented in detail the strategy we used to simulate hundreds of CL and IL sweeps from each of the four average templates estimated from the design-cell clusters of Dataset<sub>R1</sub>. The random perturbation of the average models in the N2pc time range only, as well as the random coupling of CL and IL sweeps to derive the N2pc, created thousands of N2pc signals with a wider range of parameters and features to be fed to the NN than that of Dataset<sub>R1</sub> (and of Dataset<sub>R2</sub> as well), hence reducing to the minimum the leakage of information between the two datasets and allowing a better generalization capability of the NN. Nevertheless, we recognize as a limitation of the proposed approach that some statistical properties of the average models created from Dataset<sub>R1</sub> might influence the generation of Dataset<sub>S1</sub>, and some statistical properties of the real Dataset<sub>R1</sub> may have intruded the synthetic scenario. For example, we did not consider the influence of the electrode montage on the measured data or situations of high experimental noise, which may reduce the performance of our NN on datasets with very different experimental settings.

As far as the features set is concerned, we fed the NN with both EEG powers and other time-related features extracted from both the N2pc and the related CL and IL ERPs. However, we did not observe a significant pattern of any feature when considering

negative or positive (i.e. inverted) N2pcs. Therefore, to make the network able to discriminate inverted N2pcs, here we simply considered in the feature set the difference between the average CL and IL signal in the N2pc time window. In simulation, we observed that this feature was able to distinguish positive N2pcs obtained from the average of 100 EEG sweeps per design-cell cluster with an error below 3%. Further studies could investigate whether other more specific features, which could discriminate between positive and negative N2pcs, exist. The first step towards this aim would be an investigation of the neuropsychological origin of these inverted N2pcs.

The methodology implemented in this work moves towards the development of novel strategies for the reliable estimate of small ERP components. Prior attempts sharing some similarity with the present approach employed machine-learning algorithms for the diagnosis of pathologies like schizophrenia [27, 28], epilepsy [29], autism [30], and online decoding of motor imagery movements [31–33]. Machine-learning methods were applied also to predict single-trial individual brain responses. Si *et al* proposed a supervised learning strategy to extract the discriminative spatial network pattern to predict the individual decision-making response (i.e. acceptance rate) from single-trial EEG data [34]. Subject's acceptance rate was also predicted with multiple regression models [35]. TF domain, instead, was largely used in the EEG field for selecting particular features to diagnose primary insomnia [36] or for automatically identifying sleep stages [37]. The main innovation of our contribution consists in the successful combination of TF-derived indices and machine-learning tools to predict the amplitude of a small subtractive ERP response using a minimal EEG dataset.

## 7. Conclusion

The proposed NN yielded accurate individual N2pc amplitude estimates, without loss of accuracy, from a reduced number of sweeps. Results showed that the amplitude of individual N2pcs could be reliably predicted with as few as 40 EEG sweeps per each cell of the experimental design (i.e. 80 in total for both the CL and IL components of the N2pc). These results provide evidence of the potentiality of merging TF domain and machine-learning tools in the reliable prediction of ERP components. From the point of view of the cognitive neuroscientists interested in studying the N2pc, the possibility of reliably detecting the N2pc using a relatively small number of EEG sweeps opens a whole new range of possibilities. Reduced time and costs for experimentation apart, researchers endeavoring to design paradigms using unprecedented visual search variants may find it useful to carry out quick pilot tests to ascertain that an N2pc for lateral stimuli can indeed be found prior

to turn to longer and more articulated experimental sessions. Studies of selective visuo-spatial attention processes in patients, piloting of new visual search designs and brain-computer interface (BCI) applications might highly benefit from the use of this NN with short-duration tests.

Margins of improvement and possible future developments of this work include the identification of other features based on novel neurophysiological findings for inverted N2pcs and a further reduction of the number of EEG trials necessary to probe the individual N2pc, up to possibly reach its single-trial detection, which would obviously be an important breakthrough in N2pc-based BCI research [2, 38, 39].

## Data availability statement

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

## Acknowledgments

Part of this work was supported by MIUR (Italian Minister for Education) under the initiative 'Departments of Excellence' (Law 232/2016).

## ORCID iDs

Francesca Marturano  <https://orcid.org/0000-0002-5430-1234>

Sabrina Brigadoi  <https://orcid.org/0000-0003-3032-7381>

Mattia Doro  <https://orcid.org/0000-0003-2574-4308>

Roberto Dell'Acqua  <https://orcid.org/0000-0002-3393-1907>

Giovanni Sparacino  <https://orcid.org/0000-0002-3248-1393>

## References

- [1] Matran-Fernandez A and Poli R 2017 Towards the automated localisation of targets in rapid image-sifting by collaborative braincomputer interfaces *PLoS One* **12** e0178498
- [2] Awni H, Norton J J S, Umunna S, Federmeier K D and Bretl T 2013 Towards a brain computer interface based on the N2pc event-related potential *Int. IEEE/EMBS Conf. on Neural Engineering (NER)* (San Diego, CA, USA, November 2013) pp 1021–4
- [3] Eimer M and Kiss M 2007 Attentional capture by task-irrelevant fearful faces is revealed by the N2pc component *Biol. Psychol.* **74** 108–12
- [4] Jolicœur P, Brisson B and Robitaille N 2008 Dissociation of the N2pc and sustained posterior contralateral negativity in a choice response task *Brain Res.* **1215** 160–72
- [5] Luck S J and Hillyard S A 1994 Spatial filtering during visual search: evidence from human electrophysiology *J. Exp. Psychol. Hum. Percept. Perform.* **20** 1000–14
- [6] Kiss M, Van Velzen J and Eimer M 2008 The N2pc component and its links to attention shifts and spatially selective visual processing *Psychophysiology* **45** 240–9

- [7] Woodman G F, Arita J T and Luck S J 2009 A cuing study of the N2pc component: an index of attentional deployment to objects rather than spatial locations *Brain Res.* **1297** 101–11
- [8] McCarthy G and Donchin E 1981 A metric for thought: a comparison of P300 latency and reaction time *Science* **211** 77–80
- [9] Verleger R 1997 On the utility of P3 latency as an index of mental chronometry *Psychophysiology* **34** 131–56
- [10] Scheffers M K, Coles M G H, Bernstein P, Gehring W J and Donchin E 1996 Event-related brain potentials and error-related processing: an analysis of incorrect responses to go and no-go stimuli *Psychophysiology* **33** 42–53
- [11] Brigadoi S, Basso Moro S, Falchi R, Cutini S and Dell'Acqua R 2018 On pacing trials while scanning brain hemodynamics: the case of the SNARC effect *Psychon. Bull. Rev.* **25** 2267–73
- [12] Marturano F, Brigadoi S, Brigadoi S, Doro M, Dell'Acqua R, Sparacino G, Dell'Acqua R and Sparacino G 2020 Computer data simulator to assess the accuracy of estimates of visual N2/N2pc event-related potential components *J. Neural Eng.* **17** 036024
- [13] Boudewyn M A, Luck S J, Farrens J L and Kappenman E S 2018 How many trials does it take to get a significant ERP effect? It depends *Psychophysiology* **55** 1–16
- [14] Tandle A, Jog N, D'cunha P and Chheta M 2016 Classification of artefacts in EEG signal recordings and EOG artefact removal using EOG subtraction *Commun. Appl. Electron.* **4** 12–19
- [15] Ai G, Sato N, Singh B and Wagatsuma H 2016 Direction and viewing area-sensitive influence of EOG artifacts revealed in the EEG topographic pattern analysis *Cogn. Neurodyn.* **10** 301–14
- [16] Doro M, Bellini F, Brigadoi S, Eimer M and Dell'Acqua R 2020 A bilateral N2pc (N2pcb) component is elicited by search targets displayed on the vertical midline *Psychophysiology* **57** e13512
- [17] D'Avanzo C, Schiff S, Amodio P and Sparacino G 2011 A Bayesian method to estimate single-trial event-related potentials with application to the study of the P300 variability *J. Neurosci. Methods* **198** 114–24
- [18] D'Avanzo C, Goljahani A, Pilonetto G, De Nicolao G and Sparacino G 2013 A multi-task learning approach for the extraction of single-trial evoked potentials *Comput. Methods Programs Biomed.* **110** 125–36
- [19] Mortaheb S, Rostami F, Shahin S and Amirfattahi R 2016 Wavelet based single trial event related potential extraction in very low SNR conditions *2016 6th Int. Conf. on Computer and Knowledge Engineering (ICCKE) (October 2016)* pp 82–87
- [20] Ting C-M, Salleh S-H, Zainuddin Z M and Bahar A 2015 Modeling and estimation of single-trial event-related potentials using partially observed diffusion processes *Digit. Signal Process.* **36** 128–43
- [21] Engel A K, Fries P and Singer W 2001 Dynamic predictions: oscillations and synchrony in top-down processing *Nat. Rev. Neurosci.* **2** 704–16
- [22] Bacigalupo F and Luck S J 2019 Lateralized suppression of alpha-band EEG activity as a mechanism of target processing *J. Neurosci.* **39** 900–17
- [23] McCusker M C, Wiesman A I, Schantell M D, Eastman J A and Wilson T W 2020 Multi-spectral oscillatory dynamics serving directed and divided attention *Neuroimage* **217** 116927
- [24] Marturano F, Brigadoi S, Doro M, Dell'Acqua R and Sparacino G 2020 A time-frequency analysis for the online detection of the N2pc event-related potential (ERP) component in individual EEG datasets *42nd Annual Int. Conf. IEEE Engineering in Medicine & Biology Society (EMBC) (July 2020)* (IEEE) pp 1019–22
- [25] Srivastava N, Hinton G, Krizhevsky A, Sutskever I and Salakhutdinov R 2014 Dropout: a simple way to prevent neural networks from overfitting *J. Mach. Learn. Res.* **15** 1929–58
- [26] Chollet F 2018 *Deep Learning with Python* (Shelter Island, NY: Manning Publications Co.)
- [27] Zhang L 2019 EEG signals classification using machine learning for the identification and diagnosis of schizophrenia *Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBS)* pp 4521–4
- [28] Shim M, Hwang H J, Kim D W, Lee S H and Im C H 2016 Machine-learning-based diagnosis of schizophrenia using combined sensor-level and source-level EEG features *Schizophr. Res.* **176** 314–9
- [29] Abbasi B and Goldenholz D M 2019 Machine learning applications in epilepsy *Epilepsia* **60** 2037–47
- [30] Stahl D, Pickles A, Elsabbagh M and Johnson M H 2012 Novel machine learning methods for ERP analysis: a validation from research on infants at risk for autism *Dev. Neuropsychol.* **37** 274–98
- [31] Tayeb Z, Fedjaev J, Ghaboosi N, Richter C, Everding L, Qu X, Wu Y, Cheng G and Conradt J 2019 Validating deep neural networks for online decoding of motor imagery movements from EEG signals *Sensors* **19** 210
- [32] Tabar Y R and Halici U 2017 A novel deep learning approach for classification of EEG motor imagery signals *J. Neural Eng.* **14** 016003
- [33] Zhang R, Yao D, Valdés-Sosa P A, Li F, Li P, Zhang T, Ma T, Li Y and Xu P 2015 Efficient resting-state EEG network facilitates motor imagery performance *J. Neural Eng.* **12** 066024
- [34] Si Y et al 2020 Predicting individual decision-making responses based on single-trial EEG *Neuroimage* **206** 116333
- [35] Si Y et al 2019 Predicting individual decision-making responses based on the functional connectivity of resting-state EEG *J. Neural Eng.* **16** 066025
- [36] Ben Hamida S T, Penzel T and Ahmed B 2015 EEG time and frequency domain analyses of primary insomnia *Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBS)* pp 6206–9
- [37] Mahvash Mohammadi S, Kouchaki S, Ghavami M and Sanei S 2016 Improving time–frequency domain sleep EEG classification via singular spectrum analysis *J. Neurosci. Methods* **273** 96–106
- [38] Tian Y, Zhang H, Li P and Li Y 2019 Multiple correlated component analysis for identifying the bilateral location of target in visual search tasks *IEEE Access* **7** 98486–94
- [39] Matran-Fernandez A and Poli R 2017 Brain-computer interfaces for detection and localization of targets in aerial images *IEEE Trans. Biomed. Eng.* **64** 959–69