

Unsupervised Self-Adaptive Auditory Attention Decoding: Supplementary Material

Simon Geirnaert, Tom Francart, and Alexander Bertrand, *Senior Member, IEEE*

In this supplementary material corresponding to the paper ‘Unsupervised Self-Adaptive Auditory Attention Decoding’ by Geirnaert et al. [1], we show convergence to a unique fixed point of the fixed-point iteration on the updating model (Equation (18) in the original paper [1])¹. We hypothesize that under three reasonable conditions on the accuracies of the attended and unattended decoder, there exists a unique fixed point p^* to which the fixed-point iteration $p_{i+1} = \phi(p_i)$ converges, starting from any (possibly random) decoder. In Section I, we first show that there always exists such a fixed point, while in Section II we check the uniqueness of and convergence to this fixed point under the hypothesized conditions.

I. EXISTENCE

Consider the following fixed-point theorem, also known as Brouwer’s fixed-point theorem [2]:

Theorem 1 (Brouwer’s fixed point theorem [2]). *Any continuous self map of a nonempty compact convex subset of a Euclidean space has a fixed point.*

As the function $\phi(p_i) : [0, 100]\% \rightarrow [0, 100]\%$ in (18) [1] is a continuous function that maps its domain onto itself and $[0, 1]$ is a closed (thus, compact) convex subset of \mathbb{R} , Brouwer’s fixed point theorem assures that there exists at least one fixed point.

II. UNIQUENESS AND CONVERGENCE

We evaluate the model in (18) [1] in a relevant range of the parameters μ_1, μ_2 , and σ , obeying three reasonable conditions, to show the convergence to a unique fixed point.

A. Three conditions for convergence

Consider the supervised subject-specific attended decoder \hat{d}_a with accuracy p_a (on the attended labels) and supervised subject-specific unattended decoder \hat{d}_u with accuracy p_u (on the unattended labels). We then a priori postulate the following three intuitive and reasonable conditions on the accuracies p_a and p_u (which will turn out to be satisfied for all subjects in both datasets in [1]):

- $p_a - p_u > 5\%$, i.e., the attended decoder needs to perform 5% better (on the attended labels) than the unattended decoder (on the unattended labels). Given that the attended speech envelope is typically better represented in the EEG, we indeed expect a difference in performance

between both decoders. Moreover, this condition can be linked to the expectation that the cross-correlation between the EEG and attended speech envelope is on average larger than with the unattended speech envelope, serving as a possible explanation for the self-leveraging effect (see Section IV-B in the original paper [1]).

- $p_u < 85\%$, i.e., the *unattended* decoder may not perform better than 85% (on the unattended labels). If the unattended decoder performs too well, then, again, the self-leveraging effect may not be present for the same reason as mentioned in the previous condition.
- $p_a > 100\% - p_u$, i.e., the attended decoder is better at predicting attended labels than the unattended decoder. This assures that the starting point of the model curve $\phi(0\%) = 100\% - p_u$ (e.g., see Figure 3 in the original paper [1]) is below the end point $\phi(100\%) = p_a$.

In the following sections, we will use the model in (18) [1] to show that there is convergence to a unique fixed point when these three conditions are satisfied. However, it is noted that these postulated conditions are conservative in the mathematical sense, i.e., they are ‘sufficient’ but *not* ‘necessary’ conditions. When they are not satisfied, there can still be convergence to a unique fixed point.

Moreover, the three conditions are also intuitive and very reasonable from a practical point of view, as they are satisfied for all subjects in both datasets [1] on 50/60s decision windows (i.e., the length of the segments on which the updating is performed): the minimum across all subjects of $p_a - p_u = 8.3\% > 5\%$, the maximum across all subjects of $p_u = 76.7\% < 85\%$, and the minimum across all subjects of $p_a + p_u = 124\% > 100\%$.

B. Convergence to a unique fixed point

Consider the following fixed-point theorem that provides sufficient conditions for convergence to a unique fixed point of the fixed-point iteration $p_{i+1} = \phi(p_i)$ [3]:

Theorem 2. *Let ϕ be a continuous function on $[a, b]$, such that $\phi(p_i) \in [a, b], \forall p_i \in [a, b]$, and suppose that ϕ' exists $\forall p_i \in [a, b]$ and that a constant $0 < \alpha < 1$ exists such that:*

$$|\phi'(p_i)| \leq \alpha, \forall p_i \in [a, b],$$

then there is exactly one fixed point $p^ \in [a, b]$ and the fixed-point iteration $p_{i+1} = \phi(p_i)$ will converge to this unique fixed point in $[a, b]$.*

We now evaluate the model $\phi(p_i)$ in (18) [1] and its derivative $\phi'(p_i)$ to show convergence to a unique fixed point

¹This supplementary material has also been peer-reviewed together with the original article in [1].

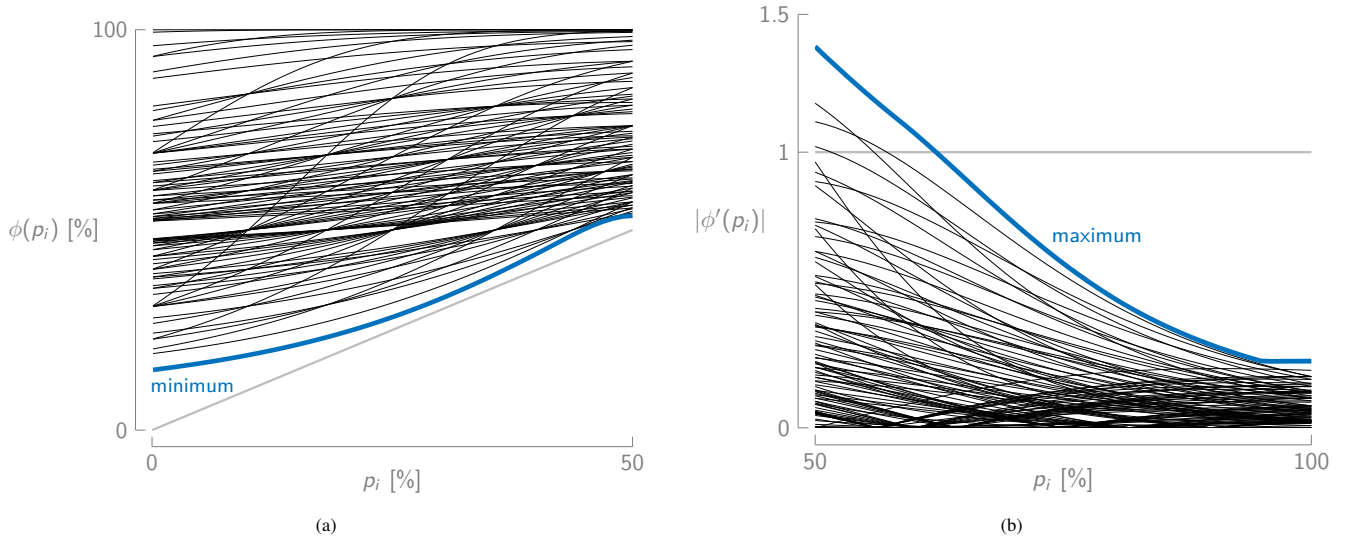


Figure 1: (a) A subset of the evaluated $\phi(p_i)$ for $p_i \in [0, 50]\%$ and the minimum over all evaluated (μ_1, μ_2, σ) that obey the conditions are all above the identity line, where $\phi(p_i) = p_i$, which shows that $\phi(p_i) > p_i, \forall p_i \in [0, 50]\%$. (b) A subset of the evaluated $|\phi'(p_i)|$ for $p_i \in [50, 100]\%$, together with the maximum over all evaluated (μ_1, μ_2, σ) that obey the conditions.

based on Theorem 2 for the case where the conditions in Section II-A are satisfied.

The derivative $\phi'(p_i)$ of the model in (18) can be computed by hand or by using any symbolic math software and is equal to:

$$\phi'(p_i) = \frac{p_i \sigma_z(p_i)^2 \mu_2 + (1 - p_i) \sigma^2 \mu_z(p_i)}{\sqrt{2\pi} p_i^3 \sigma_z(p_i)^3} e^{-\frac{1}{2} \left(\frac{\mu_z(p_i)}{\sigma_z(p_i)} \right)^2}. \quad (1)$$

To evaluate (18) [1] and its derivative (1), we take 300 equidistant samples of $\mu_1 \in [-2, 2]$, 300 equidistant samples of $\mu_2 \in [-2, 2]$, and 100 equidistant samples of $\sigma \in]0, 4]$. These intervals contain the complete range of parameters concerning the difference in correlation coefficients R_1 and R_2 . From this parameter range, we select all combinations of (μ_1, μ_2, σ) for which the three conditions of Section II-A are satisfied. The connection between p_a and p_u (as used in the three conditions) and the model parameters (μ_1, μ_2, σ) is given by:

$$p_a = P(R_1 > 0) = \frac{1}{\sigma \sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{1}{2} \left(\frac{x - \mu_1}{\sigma} \right)^2} dx \text{ and}$$

$$p_u = P(R_2 > 0) = \frac{1}{\sigma \sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{1}{2} \left(\frac{x - \mu_2}{\sigma} \right)^2} dx,$$

using the assumptions in Section IV-A in the original paper [1]. These connections can be derived from the updating model in Equation (18) from the original paper [1] by setting $p_i = 100\%$, resp. $p_i = 0\%$, resulting in the decoder accuracy of the supervised attended, resp. unattended decoder.

Figure 1a now shows a subset of $\phi(p_i)$ for $p_i \in [0, 50]\%$, for all evaluated (μ_1, μ_2, σ) that obey the three conditions, together with the minimum over all these $\phi(p_i)$. Similarly, Figure 1b shows a subset of $|\phi'(p_i)|$ for $p_i \in [50, 100]\%$, for all evaluated (μ_1, μ_2, σ) that obey the three conditions, together with the maximum over all these $|\phi'(p_i)|$. Both results are

required to show convergence to a unique fixed point using Theorem 2:

- **Result 1:** From Figure 1a, it can be seen that $\phi(p_i) > p_i, \forall p_i \in [0, 50]\%$. This implies that there is no fixed point within this interval and that the fixed-point iteration will always diverge to the $p_i \in [50, 100]\%$ interval. This is because $\forall p_i \in [0, 50]\% : p_{i+1} = \phi(p_i) > p_i$, i.e., the new accuracy in the fixed-point iteration is always larger than the previous one, such that, inevitably, at a certain iteration, $p_{i+1} > 50\%$. It thus suffices to show that there is convergence to a unique fixed point for $p_i \in [50, 100]\%$, which is shown in the next result.
- **Result 2:** From Figure 1b, there are two possible cases, which both individually can be shown to guarantee convergence to a unique fixed point:
 - 1) $|\phi'(p_i)| < 1, \forall p_i \in [50, 100]\%$. For all these cases, we then numerically confirmed that $\phi(p_i) \in [50, 100]\%, \forall p_i \in [50, 100]\%$ such that all conditions of Theorem 2 are fulfilled to show convergence to a unique point.
 - 2) $\exists x \in [50, 100]\% : \phi'(p_i) \geq 1, \forall p_i \in [50, x]\%$ and $|\phi'(p_i)| < 1, \forall p_i \in [x, 100]\%$. Since $\phi(50\%) > 50\%$ (see Result 1) and since the derivative is positive, it is guaranteed that $\phi(p_i) > p_i, \forall p_i \in [50, x]\%$, i.e., there is no fixed point and the fixed-point iteration diverges to the $p_i \in [x, 100]\%$ interval (using a similar reasoning as in Result 1). Furthermore, it can again be numerically checked that $\phi(p_i) \in [x, 100]\%, \forall p_i \in [x, 100]\%$ to show that there is a unique point to which there is convergence in this interval (see Theorem 2).

REFERENCES

- [1] S. Geirnaert, T. Francart, and A. Bertrand, "Unsupervised Self-Adaptive Auditory Attention Decoding," *IEEE Journal on Biomedical and Health Informatics*, vol. 25, no. 10, pp. 3955–3966, 2021.

- [2] J. M. Borwein and A. S. Lewis, *Convex Analysis and Nonlinear Optimization: Theory and Examples*, 2nd ed. Springer-Verlag New York, 2006.
- [3] Walter Gautschi, *Numerical Analysis*. Birkhäuser Basel, 2012.