



HAL
open science

A unified weighted minimum norm solution for the reference inverse problem in EEG

Ricardo Salido-Ruiz, Radu Ranta, Gundars Korats, Steven Le cam, Laurent Koessler, Valérie Louis-Dorr

► **To cite this version:**

Ricardo Salido-Ruiz, Radu Ranta, Gundars Korats, Steven Le cam, Laurent Koessler, et al.. A unified weighted minimum norm solution for the reference inverse problem in EEG. *Computers in Biology and Medicine*, 2019, 115, pp.103510. 10.1016/j.combiomed.2019.103510 . hal-02374112

HAL Id: hal-02374112

<https://hal.science/hal-02374112>

Submitted on 24 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A unified weighted minimum norm solution for the reference inverse problem in EEG

Ricardo A. Salido-Ruiz^a, Radu Ranta^b, Gundars Korats^c, Steven Le Cam^b,
Laurent Koessler^{b,d}, Valerie Louis-Dorr^b

^aUniversity of Guadalajara, Department of Computer Science in the University Center for Exact Sciences and Engineering (CUCEI), Guadalajara, Jalisco, Mexico

^bUniversité de Lorraine, CNRS, CRAN, F-54000 Nancy, France

^cVentspils University of Applied Sciences, Ventspils Smart Technology Research Centre, Ventspils, Latvia

^dCHRU Nancy, Neurology Service, F-54000 Nancy, France

Abstract

A well known problem in EEG recordings deals with the unknown potential of the reference electrode. In the last years several authors presented comparisons among the most popular solutions, the global conclusion being that the traditional Average Reference (AR) and the Reference Standardization Technique (REST) are the best approximations [25, 17, 20, 4]. In this work we do not aim to further compare these techniques but to enlighten the fact that both solutions can be derived from a general inverse problem formalism for reference estimation. In particular, we show that the AR is the minimum norm solution, while REST is a weighted minimum norm including some approximate propagation model. AR is thus a particular case of REST, which itself uses a particular formulation of the source estimation inverse problem.

Keywords: EEG, Reference potential, Inverse problems

1. Introduction

The EEG signals measure the potential difference between an electrode, placed somewhere on the head surface, and a reference electrode, placed somewhere else on the body. Ideally, the reference electrode should be placed in the most electrically inactive position, far from the region of interest. In practice, the reference is contaminated with unknown local and/or propagated electrical activity. Still, most of the studies need reference free potentials for better results [8, 6, 36, 17, 27]. For example, cortical connections are estimated using coherence and phase delays among EEG channels, and the reference potential

Email addresses: ricardo.salido@cucei.udg.mx (Ricardo A. Salido-Ruiz),
radu.ranta@univ-lorraine.fr (Radu Ranta)

might strongly affect these methods¹ [27, 7, 26, 13, 12]. A previous study showed that even Blind Source Separation techniques might be affected by the reference problem [33].

This paper focuses on EEG recorded using a scalp reference. In this case, the reference electrode captures the same kind of activity as any other electrode, that is a mixture of brain sources. In this context, the most popular solution for cancelling the reference potential is to subtract the spatial average of the recorded signals from all measurements. This solution, known as average reference montage (AR) was challenged by the Reference Standardization Technique (REST) [36]. Several authors compared the two solutions (or other re-referencing techniques) [38, 37, 25, 17, 32, 20, 11], the overall conclusion being that REST is deemed to yield better results, although this is not always the case, especially when the head model is prone to inaccuracies. The goal of this paper is to theoretically show that both AR and REST belong to the same family of inverse problems widely encountered in EEG analysis, *i.e.*, weighted minimum norms. If the same idea was recently put forward by [16, 15], our paper proposes an alternative simple algebraic proof (another algebraic proof can be found also in [32], in French). To avoid redundancies with the previously cited papers, we do not propose an extensive simulation protocol in order to assess if or when one of the solutions should be used, but we present an evaluation of the difference between them depending on the used head model (necessary for REST).

2. EEG measurement model

Consider the classical EEG linear model given by:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^{m \times 1}$ is a vector with unknown real potentials under each electrode with respect to infinity, from here defined as absolute potentials for the m electrodes, $\mathbf{A} \in \mathbb{R}^{m \times p}$ is the mixing lead-field matrix and $\mathbf{s} \in \mathbb{R}^{p \times 1}$ is the source vector. Note that the reference potential is included among the m ; for convenience, it will be assumed in the rest of the paper that this potential is the m -th one. In EEG, the matrix \mathbf{A} represents the head model, obtained either by analytical approximations (e.g. spherical head models) or numerically, for realistic head models obtained after MRI segmentation of head structures (brain, CSF, skull, scalp). The sources vector \mathbf{s} contains the amplitudes of the dipoles that model the neural sources [2]. For these realistic models, possible source emplacements depend on the brain volume discretization but, regardless of the used head model, the number of sources p is far bigger than the number of electrodes m ($p \gg m$). Equation (1) introduces the so-called instantaneous

¹Although some of these studies were focused on intracranial EEG, the effect of the reference potential is the same. Note though that for iEEG, the reference estimation techniques are different, as they assume that the reference is sufficiently far from the measuring electrodes to be considered independent/uncorrelated [14, 21, 31].

mixture model, with the matrix \mathbf{A} time-invariant and of rank m . For real signals recorded in time, *i.e.*, having $n \gg 1$ time samples, the vectors \mathbf{x} (\mathbf{s}) become matrices $m \times n$ ($p \times n$), and we consider during all the paper that \mathbf{x} is full-row rank m .

The actually measured potentials are given by the common reference montage (CR), with potentials \mathbf{x}_{CR} modeled by subtracting the potential of the chosen reference electrode from the other electrodes. This can be seen as a matrix transform of the absolute potentials from (1):

$$\mathbf{x}_{CR} = \mathbf{T}_{CR}\mathbf{x} = \mathbf{T}_{CR}\mathbf{A}\mathbf{s} = \mathbf{A}_{CR}\mathbf{s} \quad (2)$$

with \mathbf{T}_{CR} the $m - 1$ rank matrix:

$$\mathbf{T}_{CR} = \begin{bmatrix} \mathbf{I}_{m-1} & -\mathbf{1}_{m-1} \end{bmatrix} \quad (3)$$

where \mathbf{I}_{m-1} is the $(m - 1) \times (m - 1)$ identity matrix and $\mathbf{1}_{m-1} \in \mathbb{R}^{m-1 \times 1}$ is a vector of 1's. As mentioned above and with no loss of generality, we assume the reference electrode potential as the last absolute potential in EEG linear model (1). **Note that, unlike \mathbf{x} in (1), the dimension of \mathbf{x}_{CR} is $m - 1$ (number of available signals).**

The most employed solutions propose to cancel the reference by transforming the CR into other EEG montages (average reference - AR, bipolar - BM, Laplacian - LM)[14, 38, 8, 19, 10]. A more elaborated solution, based on head modeling, is the reference standardization (REST) [36]. Among these, only AR and REST propose absolute potential estimations, the BM and the LM being local estimators (along with the reference potential, they also cancel propagated activities originating in far situated regions). We will only focus here on the former.

2.1. Average Reference (AR)

The classical rationale behind average reference montages (AR) is that, under the hypothesis of equidistributed electrodes covering the whole head, the sum of all electrode potentials on the scalp should be zero (see [3] for a proof for the spherical head model).

In practice, as the only available signals are \mathbf{x}_{CR} , AR signals are obtained by calculating the sum of the $m - 1$ \mathbf{x}_{CR} signals, dividing it by the total number of electrodes m and subtracting it from each measurement [28, 29, 6]. More formally, the AR is obtained by averaging over all electrodes (reference electrode included with a null potential (*i.e.*, its potential with respect to itself):

$$\mathbf{x}_{AR} = \left(\mathbf{I}_m - \frac{1}{m} \mathbf{1}_m \mathbf{1}_m^T \right) \begin{bmatrix} \mathbf{x}_{CR} \\ 0 \end{bmatrix} = \mathbf{T}_{AR} \mathbf{x}_{CR} \quad (4)$$

with

$$\mathbf{T}_{AR} = \begin{bmatrix} \mathbf{I}_{m-1} - \frac{1}{m} \mathbf{1}_{m-1} \mathbf{1}_{m-1}^T \\ \frac{1}{m} \mathbf{1}_m^T \end{bmatrix} \quad (5)$$

As we can see, \mathbf{x}_{AR} has a dimension of m and rank of $m - 1$, *i.e.*, it preserves the rank of \mathbf{x}_{CR} (one less than \mathbf{x}). In practice, the average signal is included as the last row $(x_{m,AR})^2$.

2.2. Reference Electrode Standardization Technique (REST)

The REST method [36, 38] allows to estimate absolute potentials at any point on the scalp by solving a forward problem with “equivalent sources” $\tilde{\mathbf{s}}$ constrained to a dipolar layer completely surrounding the actual sources, *i.e.*, the brain. Several versions can be theoretically proposed, depending on the configuration chosen for the dipolar layer (*e.g.*, a sphere circumscribing the whole brain, or the numerically approximated cortical surface). The chosen dipolar layer will yield a specific forward model between the dipoles situated on this layer and the actual electrodes placed on the head surface. Let this model be $\tilde{\mathbf{A}}$. The equivalent dipolar sources on this layer are estimated from scalp EEG recordings (equivalently in CR or in AR) by using a simple inverse problem formalism. We give below only the CR based approach:

$$\hat{\mathbf{s}} = \tilde{\mathbf{A}}_{CR}^+ \mathbf{x}_{CR} \quad (6)$$

with $+$ designating the classical Moore-Penrose pseudo-inverse. Of course, $\tilde{\mathbf{A}}_{CR}$ depends on a mixing model (see Eq. (3)), here assumed $\tilde{\mathbf{A}}$. A forward model using the assumed $\tilde{\mathbf{A}}$ and the estimated equivalent sources $\hat{\mathbf{s}}$ yield the REST estimations of the EEG absolute potentials:

$$\mathbf{x}_{REST} = \tilde{\mathbf{A}} \hat{\mathbf{s}} = \tilde{\mathbf{A}} \tilde{\mathbf{A}}_{CR}^+ \mathbf{x}_{CR} = \mathbf{T}_{REST} \mathbf{x}_{CR} \quad (7)$$

It is important to recall that different REST solutions can be obtained by choosing different dipolar layer configurations and thus different $\tilde{\mathbf{A}}$ models.

3. Unified Inverse Problem Modeling

As seen above, two main solutions exist for estimating the absolute potentials: average montage and REST. Three relatively recent comparative studies [17, 25, 20] conclude that both solutions present valid theoretical arguments and that both are acceptable. Our aim is to deepen the analysis, to show the connections between the two techniques and their integration in a larger inverse problem framework.

Consider the case of a EEG recording with a cephalic reference as given in (Eq. 2). The estimation of absolute potentials \mathbf{x} from the measured \mathbf{x}_{CR} and matrix transformation \mathbf{T}_{CR} , is an ill-posed inverse problem somehow similar to the classical EEG source estimation with a known mixing model \mathbf{T}_{CR} .

²Note that in [11], the formulation is done differently (*i.e.*, they use a \mathbf{T}_R of dimension $Ne \times Ne$, $m = Ne$ in our case), but they are equivalent (*i.e.*, one of the rows of \mathbf{T}_R has only zeros). Our rationale was to start, as in classical inverse problems, from the actual measurements \mathbf{x}_{CR} and not from the absolute potentials. Still, the two formulations yield the same conclusions.

Thus, **the absolute potentials** inverse problem writes as follows:

$$\hat{\mathbf{x}} = \mathbf{T}\mathbf{x}_{CR} = \mathbf{T}\mathbf{T}_{CR}\mathbf{x} \quad (8)$$

where the unknown matrix $\mathbf{T} \in \mathbb{R}^{m \times m-1}$ is a generalized inverse of the common reference transformation matrix \mathbf{T}_{CR} :

$$\mathbf{T} = \mathbf{W}^{-1}\mathbf{T}_{CR}^T [\mathbf{T}_{CR}\mathbf{W}^{-1}\mathbf{T}_{CR}^T]^{-1} \quad (9)$$

with \mathbf{W} a weighting matrix allowing infinite solutions. In the next paragraphs we describe several particular Weighted Minimum Norm (WMN) solutions.

3.1. WMN Optimal Solution: Oracle1

The underdetermined system (8) has an infinite number of solutions. In theory, the best inverse transform \mathbf{T} can be estimated by minimizing the least-squares error between absolute and measured potentials. This optimal transform \mathbf{T}_{O1} can be obtained by multiplying absolute potentials \mathbf{x} by the Moore-Penrose pseudo inverse of measured potentials \mathbf{x}_{CR} :

$$\mathbf{T}_{O1} = \mathbf{x}\mathbf{x}_{CR}^T [\mathbf{x}_{CR}\mathbf{x}_{CR}^T]^{-1} = \mathbf{A}\mathbf{\Sigma}_s\mathbf{A}^T\mathbf{T}_{CR}^T [\mathbf{T}_{CR}\mathbf{A}\mathbf{\Sigma}_s\mathbf{A}^T\mathbf{T}_{CR}^T]^{-1} \quad (10)$$

where $\mathbf{\Sigma}_x$ is the absolute potentials covariance matrix, \mathbf{A} the head model and $\mathbf{\Sigma}_s$ the source covariance matrix (for completeness, the proof is given in the appendix). This optimal solution is equivalent to (9) for $\mathbf{W}^{-1} = \mathbf{\Sigma}_x = \mathbf{A}\mathbf{\Sigma}_s\mathbf{A}^T$. Of course, it remains theoretical, because neither the absolute potentials covariance matrix nor, equivalently, the propagation coefficients between the actual sources \mathbf{s} and the sensors (depending on the sources positions and orientations and on the head model) and the source covariance matrix, are known in practice.

3.2. WMN Sub-Optimal Solutions: Oracle2 and REST

If, in practice, $\mathbf{\Sigma}_s$ is not known, rather accurate head models can be obtained for \mathbf{A} for all possible source locations inside the brain volume.

Ignoring the source covariance matrix but supposing that we have some *a priori* knowledge on their positions and orientations, as well as a good head model (*i.e.*, making $\mathbf{\Sigma}_s = \mathbf{I}_p$ but assuming a known \mathbf{A}), one obtains:

$$\begin{aligned} \mathbf{T}_{O2} &= \mathbf{A}\mathbf{A}^T\mathbf{T}_{CR}^T [\mathbf{T}_{CR}\mathbf{A}\mathbf{A}^T\mathbf{T}_{CR}^T]^{-1} \\ &= \mathbf{A} [\mathbf{T}_{CR}\mathbf{A}]^+ = \mathbf{A}\mathbf{A}_{CR}^+ \end{aligned} \quad (11)$$

A step further towards more realistic situations is to ignore the source positions and orientations, but to still assume a complete head model, *i.e.*, a mixing matrix \mathbf{A}_C corresponding to a complete lead-field matrix for all possible source configurations (of course, in practice it will be computed for a grid of points inside the brain volume). The equations are exactly the same as (11), with a complete mixing matrix \mathbf{A}_C replacing \mathbf{A} (which is an incomplete lead-field

matrix, in the sense that it accounts only for the *a priori* known source positions and orientations). Finally, as pointed out in [36], distributing equivalent sources on a layer surrounding the actual brain sources, theoretically yields another mixing model \mathbf{A}_R used in the REST solution.

If Oracle1 and Oracle2 cannot be used in practice (because they are based on unknown information on the actual sources), it is noteworthy that REST-like solutions (either based on a complete volume model or an equivalent layer) could be in principle used, because they only need a propagation model between some chosen sources and the surface electrodes, and these models can be estimated with more or less accuracy from imaging techniques and physical considerations (see e.g. [9] for a review of the forward problem in EEG).

Finally, it is easy to see that all these weighted minimum norm solutions share the same equation (11), which enlightens the fact that any full-rank matrix $\tilde{\mathbf{A}}$ can be used to construct a generalized inverse of \mathbf{T}_{CR} . Indeed,

$$\mathbf{T}_{CR}\tilde{\mathbf{A}}\tilde{\mathbf{A}}^T\mathbf{T}_{CR}^T\left[\mathbf{T}_{CR}\tilde{\mathbf{A}}\tilde{\mathbf{A}}^T\mathbf{T}_{CR}^T\right]^{-1}=\mathbf{I}_{m-1}$$

regardless of the accuracy of the model $\tilde{\mathbf{A}}$. In this sense, a completely false or random model $\tilde{\mathbf{A}}$ will yield false estimates of the absolute potentials $\hat{\mathbf{x}}$, but they will still verify $\mathbf{x}_{CR}=\mathbf{T}_{CR}\hat{\mathbf{x}}$.

3.3. Minimum Norm Solution, MN

Within this context, the most evident solution of equation (9), without any *a priori* information about the mixture or the sources (*i.e.*, neither on \mathbf{A} nor on the source covariance), is the Minimum Norm Solution (MN) obtained when $\mathbf{W}=\mathbf{I}_m$. Then, an estimation of absolute potentials is given by:

$$\begin{aligned}\hat{\mathbf{x}} &= \mathbf{T}_{CR}^T(\mathbf{T}_{CR}\mathbf{T}_{CR}^T)^{-1}\mathbf{x}_{CR} \\ &= \mathbf{T}_{CR}^+\mathbf{x}_{CR}\end{aligned}\tag{12}$$

with \mathbf{T}_{CR}^+ the Moore-Penrose pseudo-inverse of the common reference transformation matrix \mathbf{T}_{CR} .

Proposition. The minimum norm solution to the inverse reference problem is the AR solution from (5):

$$\mathbf{T}_{CR}^+=\mathbf{T}_{AR}$$

Proof. By Sherman-Morrison formula and using the definition of the Moore-

Penrose pseudo-inverse and the expression of \mathbf{T}_{CR} (3), one can write:

$$\begin{aligned}
\mathbf{T}_{CR}^+ &= \mathbf{T}_{CR}^T (\mathbf{T}_{CR} \mathbf{T}_{CR}^T)^{-1} \\
&= \begin{bmatrix} \mathbf{I}_{m-1} \\ -\mathbf{1}_{m-1}^T \end{bmatrix} \left(\begin{bmatrix} \mathbf{I}_{m-1} & -\mathbf{1}_{m-1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{m-1} \\ -\mathbf{1}_{m-1}^T \end{bmatrix} \right)^{-1} \\
&= \begin{bmatrix} \mathbf{I}_{m-1} \\ -\mathbf{1}_{m-1}^T \end{bmatrix} [\mathbf{I}_{m-1} + \mathbf{1}_{m-1} \mathbf{1}_{m-1}^T]^{-1} \\
&= \begin{bmatrix} \mathbf{I}_{m-1} \\ -\mathbf{1}_{m-1}^T \end{bmatrix} \left(\mathbf{I}_{m-1} - \mathbf{1}_{m-1} \frac{\mathbf{1}_{m-1}^T}{1 + \mathbf{1}_{m-1}^T \mathbf{1}_{m-1}} \right) \\
&= \begin{bmatrix} \mathbf{I}_{m-1} \\ -\mathbf{1}_{m-1}^T \end{bmatrix} \left(\mathbf{I}_{m-1} - \mathbf{1}_{m-1} \frac{1}{m} \mathbf{1}_{m-1}^T \right) \\
&= \mathbf{T}_{AR}
\end{aligned} \tag{13}$$

In the first row, we use block-matrix multiplication, while in the second we use the Sherman-Morrison formula (or more generally the Woodbury identity), see for example [30]). Readers interested by a second algebraic proof are directed to [32].

Note that the AR montage obtained by pseudo-inverting \mathbf{T}_{CR} can be seen as another particular case of REST, obtained for a dipolar layer having the head geometry and situated infinitely close below the electrodes (closely mimicking an $\hat{\mathbf{A}} = \mathbf{I}_m$).

4. Results

The aim of this section is to briefly present numerical results, both on simulated and real data, supporting the previous analysis.

4.1. Simulation

The simulations presented here illustrate the fact that the accuracy of different estimations of the absolute potentials, obtained using the previously described approaches. These estimated potentials are compared with simulated ground truth potentials, computed as follows: a three shell mesh model (Colin 27) was extracted from Brainstorm [1] toolbox in order to have a realistic geometry. The cortical layer mesh (inner shell) had 642 nodes with 13mm mean distance between neighbouring points, while the scalp mesh (outer shell) had 1922 nodes.

A regularly spaced grid was constructed inside the inner shell (brain), with a 13mm distance between neighbouring points, which yielded 1774 points. Several source configurations were tested, namely using $p = \{10, 100, 500\}$ dipolar brain sources randomly chosen among the inner shell grid points. The orientations were random, and the time courses were simulated as spatially and temporally correlated Gaussians of length $N_s = (\frac{m}{8} + 1) \times f_s$ seconds (rule derived from [18]), f_s being sampling frequency and m the number of scalp electrodes. Three

sensor setups were tested using $m = \{64, 128, 256\}$ scalp electrodes placed on the head surface (outer shell) according to the 10-10 system, one of them being the reference electrode. Sensor coordinates were also extracted from the Brainstorm toolbox and *snapped* to the mesh vertices by nearest-neighbour rule. The absolute EEG scalp potentials were simulated by projecting the sources of interest to the sensors using the BEM model implemented in the Helsinki toolbox [34], with conductivity ratios of 40:1 between the brain and the skull and 1:1 between brain and scalp. Simulated absolute potentials \mathbf{x} were transformed into measured potentials \mathbf{x}_{CR} by (2). One hundred simulations were performed for every configuration (position, orientations and time courses of the dipoles).

In order to test the influence of the **model accuracy** on the results, we tested the different weighted inverse solutions, from the fully specified Oracle (10) to the model-free average reference (4), passing through REST and volume REST_v (11). For REST_v, the lead-field \mathbf{A}_C was computed for the complete grid of 1774 volume points inside the brain shell and all three orientations ($m \times 5322$). For REST, the lead-field \mathbf{A} was computed for the cortical layer of 642 points and for dipoles orthogonal to the scalp ($m \times 642$). For completeness, we also implemented a REST-like solution, in which we used a random \mathbf{A} mixing model ($m \times 642$, generated as a spatially correlated uniform random variables), with no physical significance. We re-emphasize here that all models (including the random one) perfectly explain the CR measurements. The evaluation criterion was the relative error introduced in [36]:

$$RE = \|\mathbf{X} - \hat{\mathbf{X}}\| / \|\mathbf{X}\| \quad (14)$$

where \mathbf{X} contains the absolute potentials, $\hat{\mathbf{X}}$ are the estimates and $\|\cdot\|$ is the Frobenius norm.

The simulation results confirm the hypothesis that the accuracy of the estimated zero-referenced potentials is proportional to the amount of additional correct information injected in the solution. Completely wrong models (random \mathbf{A}) yield the worst solutions, far below the average reference AR (they are not figured here in order to ease the comparisons among the other methods). In principle, one should use as complete models as possible (Figs 2). If we put aside the Oracles, very accurate but impossible to use in practice, the complete full-volume REST_v model generally outperforms the cortical REST, at least for the same spatial density of the lead-field matrix (not shown here, a denser cortical layer REST also improves the estimations)³.

4.2. Real signals

According to our previous argumentation, REST and AR solutions should converge to the same solution when the dipolar layer used for REST approaches the head surface and thus the sensors. We have evaluated their relationship using real data, without any previous knowledge on the sources. The EEG signals

³Note that Oracle solutions cannot function if the number of sources is below the number of sensors because of the matrix inversion in (10) and (11).

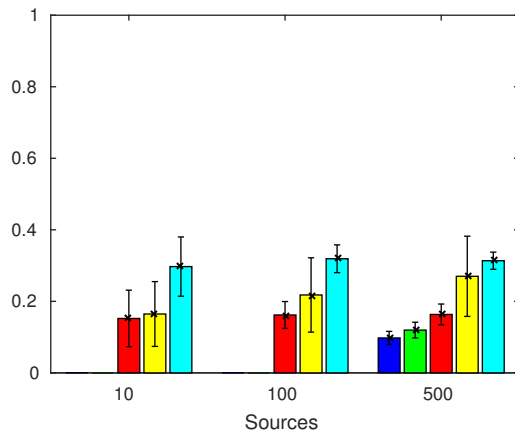


Figure 1: Relative errors RE . The bars represent mean errors over 100 simulations, the standard deviations being given by the vertical whiskers on the top of each bar. Only the 128 sensors setup is shown here, but the results are similar for the other tested montages

Depth	10mm	8mm	6mm	4mm
RE	0.136	0.101	0.082	0.078

Table 1: Relative error between the AR and REST estimations, depending on the depth of the dipolar layer

(resting state) were recorded at the Nancy University Hospital (CHU Nancy) during routine clinical evaluation of an epileptic patient (XXXXXXXXXX). Seventy-six electrodes were placed according to the 10-10 system, the reference electrode being chosen, for clinical reasons at FPz. The head (scalp, outer skull, inner skull) were computed from the MRI of the patient using the BrainStorm pipeline (1922 nodes and 3840 faces for each surface).

Several REST solutions were computed for the absolute potentials, for dipolar layers placed at different depths with respect to the head surface, but outside the brain (inner skull) mesh (see figure 2. The geometry of the layer was the same as the one of the scalp, in order to be able to keep a constant distance between the sensors and the layer, except in the lower part of the brain, where we considered a flat surface 10 mm outside the inner skull. The number of dipoles on the layer was constant and the same as the number of points on the scalp mesh (1922). We used the same criterion RE for evaluating the difference between REST and AR montages, REST being considered the ground truth (\mathbf{X} in (14)) and AR the estimate $\hat{\mathbf{X}}$. The results are given in table 1, for different depths of the dipolar layer. As it can be seen, the closer the dipolar layer to the surface, the smaller the “distance” (the relative difference, in Forbenius norm) between the AR and REST estimates (for numerical reasons, dipolar layers closer to the surface than 3mm yield worse results, in the sense that the REST solution is less accurate than the simple average montage AR).

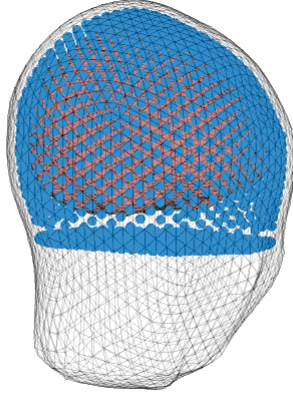


Figure 2: Partial head mesh for the studied epileptic patient. To ease the visualization, only the scalp and the inner skull are represented, along with the dipolar layer used for REST (here at 8mm depth below the scalp surface)

5. Discussion

This work presents a unified inverse problem framework for the reference problem in surface EEG, which is an algebraic alternative to the approach recently proposed by [11, 15]. The main difference is that our approach is purely deterministic and based only on linear algebra (i.e., matrix pseudo-inverse estimation, the minimum norm solution and weighted minimum norm in the least-squares (LS) sense [5, 35, 24] while Hu *et al.* approaches are semi [15] or completely non deterministic [11]. In other words, our approach belongs to the family of LS estimators and does not imply any distributional assumptions as in frequentists statistics (maximum likelihood estimates, ML) or Bayesian inference (maximum a posteriori, MAP) [23]. To the best of our knowledge, this algebraic approach was not proposed elsewhere, even if the same unified framework paradigm for the reference problem was discussed in [11, 15].

The LS solution is estimated when an inverse problem has no exact solution and it can be extended to the general linear inverse problem as presented in [22]. The solution to general linear inverse problems from the point of view of LS yields to the Moore-Penrose pseudo-inverse. The absolute potentials estimation by pseudo-inversion is an underdetermined inverse problem, as the number of absolute potentials to be estimated is greater than the number of available signals (m with respect to $m - 1$ in the notations used in this paper). Consequently, the problem has an infinite number of solutions, parametrized by the *a priori* information that can be injected. The different parametrizations lead to different so-called *Weighted MN* solutions, having different accuracies depending on the quality of the injected *a priori* information, as we have shown in the previous section.

WXWXWXWXWWX do we need all that follows, till the Conclusion ?
 WXWXWXWX

In the other hand ML solutions of the linear inverse problem assumes data with multivariate Gaussian distribution, unknown model parameters and if we know the data covariance matrix the model parameters can be estimated [22]. Here, is the main difference ML solution seeks to estimate the model parameters in order to find the inverse problem solution while WMN seek to estimate the inverse problem solution from a known model. ML obtains the optimum values for the model parameters by maximizing the probability that the observed data are in fact observed [22]. This means that the ML solution of the model parameters is the weighted least squares (WLS) solution with the weighting matrix as the inverse of the covariance matrix of the observed data. Now, in the case of the underdetermined problem the LS inverse does not exist. This, is because the distribution of the observed data has no well-defined maximum with respect to variations of the model parameters. Thus, in order to solve this underdetermined problem ML needs to add *a priori* information that allows the distribution to have a well-defined peak. The way of doing this is explained in [22], in resume the ML problem for an exact theory translates into finding the maximum of the joint distribution of the *a priori* distribution of data and the *a priori* probability distribution of model parameters. Note that if the *a priori* probability distribution for the model parameters is much more certain than that of the observed data, then the estimate of model parameters (the maximum likelihood point) tends to be close to the *a priori* model parameters. On the other hand, if the data are far more certain than the model parameters, then the estimates of the model parameters primarily reflect information contained in the data [22]. In this context the ML solution strongly depends of the probability distribution of the model parameters.

MAP solution comes from Bayesian inference statistics and it works on a posterior distribution not only the likelihood (i.e., the probability of the data given the model parameters assuming true population of model parameters). The posterior distribution in change, is the probability of the model parameters given the data.

The algebraic proof that we propose allows to uncover a family of WMN solutions, with in principle different performances from the methods for inverse problem solutions based on inference statistics such as ML and MAP solutions. Our presented solution enlightens the fact that a good model yields the best solution, a false model yields a wrong solution and the model-free AR is somewhere in the middle, as REST (which outperforms AR if the model is right). As we have discussed here, ML can be used to estimate a model that can be used in the WMN solution in order to obtain solutions comparable with REST or even better. MAP solutions can be used in the same way.

A second major difference with other similar works such as [11, 15] is how we tackle the inverse problem of the reference. First, we start estimating the generalized inverse of the common reference transformation matrix (i.e., the matrix which is implicitly multiplying the unknown zero-reference potentials), formally is the pre-multiplication of the reference transformation matrix with the inaccessible zero-reference EEG potentials. The difference between our reference transformation matrix \mathbf{T}_{CR} is that is a non-square matrix. In the work

presented by Hu in [11] they present their equivalent matrix \mathbf{H} or Tr with the difference that is a square matrix pre-multiplying the and is a square matrix (i.e., $\mathbf{I}_{N_e} - \mathbf{1}\mathbf{f}^t$, where \mathbf{f} is a vector of the same length as vector $\mathbf{1}$, vector of ones). Our solution was obtained in one step because it is well known that the generalized inverse has a known form, thus the inverse of our transformation matrix is presented in eq.(9) with \mathbf{W} a weighting matrix allowing infinite solutions. If \mathbf{W} is the identity we obtain the minimum norm solution in the least squares sense which yields to the solution given by the average reference. This approach also allows to replace \mathbf{W} for any other matrix which describes the variance of the signals to be estimated, that we need to know *a priori* and can be obtained from the head model.

WXWXWXWXWXWXWXWXWXWX

6. Conclusions

As mentioned, our aim is not to present new simulation results, for extensive simulation results and comparisons (including the effects of model errors, geometry or noise errors), the reader is referred to [38, 37, 25, 17, 20, 4]. The main contribution of the work presented in this paper is the reformulation of different reference estimation methods. Indeed, all methods, from the best possible model based **Oracle1** to the model-free average reference, can be seen as ill-posed inverse problems aiming to estimate absolute potentials from the measurements. In particular, we have shown that the AR is the minimum-norm solution to this problem, while the other methods are weighted norms, more or less accurate depending on the amount and on the precision of the injected *a priori* information.

Acknowledgements

The authors would like to thank for the funding to the National Council of Science and Technology of Mexico (CONACyT) and the Latvian National Research Programme “The next generation of information and communication technologies” (NexIT).

References

- [1] Baillet, S., Mosher, J. C., Leahy, R. M. and Shattuck, D. W. [1999]. Brainstorm: a matlab toolbox for the processing of meg and eeg signals, *NeuroImage* **9**: S246–S246.
- [2] Baillet, S., Mosher, J. and Leahy, R. [2001]. Electromagnetic brain mapping, *IEEE Signal Process. Mag.* **18**(6): 14–30.
- [3] Bertrand, O., Perrin, F. and Pernier, J. [1985]. A theoretical justification of the average reference in topographic evoked potential studies, *Electroencephalography and clinical neurophysiology* **62**: 462–464.

- [4] Chella, F., Pizzella, V., Zappasodi, F. and Marzetti, L. [2016]. Impact of the reference choice on scalp eeg connectivity estimation, *Journal of Neural Engineering* **13**(3): 036016.
URL: <http://stacks.iop.org/1741-2552/13/i=3/a=036016>
- [5] Demmel, J. W. [1997]. *Applied numerical linear algebra*, Vol. 56, Siam.
- [6] Dien, J. [1998]. Issues in the application of the average reference: review, critiques and recommendations, *Behavior Research Methods, Instruments, & Computers* **30**(1): 34–43.
- [7] Essl, M. and Rappelsberger, P. [1998]. EEG coherence and reference signals: experimental results and mathematical explanations, *Medical and Biological Engineering and Computing* **39**: 399–406.
- [8] Geselowitz, D. B. [1998]. The zero of potential, *IEEE Engineering in Medicine and Biology Magazine* **17**(1): 128–132.
- [9] Hallez, H., Vanrumste, B., Grech, R., Muscat, J., De Clercq, W., Vergult, A., D’Asseler, Y., Camilleri, K., Fabri, S., Van Huffel, S. and Lemahieu, I. [2007]. Review on solving the forward problem in EEG source analysis, *J Neuroeng Rehabil.* pp. 4–46.
- [10] Hjorth, B. [1975]. An on-line transformation of EEG scalp potentials into orthogonal source derivations, *Electroencephalography and Clinical Neurophysiology* **39**(5): 526 – 530.
- [11] Hu, S., Lai, Y., Valdes-Sosa, P. A., Bringas-Vega, M. L. and Yao, D. [2018]. How do reference montage and electrodes setup affect the measured scalp eeg potentials?, *Journal of neural engineering* **15**(2): 026013.
- [12] Hu, S., Stead, M., Dai, Q. and Worrell, G. [2010]. On the recording reference contribution to EEG correlation, phase synchrony, and coherence, *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* **40**(5): 1294 –1304.
- [13] Hu, S., Stead, M., Gardner, A. B. and Worrell, G. A. [2007]. *Lecture Notes in Computer Science: Advances in Neural Networks*, Springer-Verlag, chapter The Effect of Recording Reference on EEG: Phase Synchrony and Coherence, pp. 1273–1280.
- [14] Hu, S., Stead, M. and Worrel, G. [2007]. Automatic identification and removal of scalp reference signal for intracranial EEGs based on Independent Component Analysis, *IEEE Trans. Biomed. Eng.* **54**(9): 1560–1572.
- [15] Hu, S., Yao, D., Bringas-Vega, M. L., Qin, Y. and Valdes-Sosa, P. A. [2019]. The statistics of eegunipolar references: Derivations and properties, *Brain Topography* .

- [16] Hu, S., Yao, D. and Valdes-Sosa, P. A. [2018]. Unified Bayesian estimator of EEG reference at infinity: rREST (Regularized Reference Electrode Standardization Technique), *Frontiers in Neuroscience* **12**: 297.
- [17] Kayser, J. and Tenke, C. E. [2010]. In search of the Rosetta Stone for scalp EEG: Converging on reference-free techniques., *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* **121**: 1973–1975.
- [18] Korats, G., Cam, S., Ranta, R. and Hamid, M. [2013]. Applying ICA in EEG: Choice of the window length and of the decorrelation method, *in* J. Gabriel, J. Schier, S. van Huffel, E. Conchon, C. Correia, A. Fred and H. Gamboa (eds), *Biomedical Engineering Systems and Technologies*, Vol. 357 of *Communications in Computer and Information Science*, Springer Berlin Heidelberg, pp. 269–286.
- [19] Libenson, M. [2012]. *Practical approach to electroencephalography*, Elsevier Health Sciences.
- [20] Liu, Q., Balsters, J., Baechinger, M., van der Groen, O., Wenderoth, N. and Martini, D. [2015]. Estimating a neutral reference for electroencephalographic recordings: The importance of using a high-density montage and a realistic head model, *Journal of Neural Engineering* **12**(5): 056012.
- [21] Madhu, N., Ranta, R., Maillard, L. and Koessler, L. [2012]. A unified treatment of the reference estimation problem in depth EEG recordings, *Med. Biol. Eng. Comput.* **50**(10): 1003–1015.
- [22] Menke, W. [2018]. *Geophysical data analysis: Discrete inverse theory*, Academic press.
- [23] Myung, I. J. [2003]. Tutorial on maximum likelihood estimation, *Journal of mathematical Psychology* **47**(1): 90–100.
- [24] Noble, B. [1988]. Applied linear algebra, *Technical report*.
- [25] Nunez, P. [2010]. REST: A Good Idea but Not the Gold Standard, *Clinical Neurophysiology* **121**(12): 2177–2180.
- [26] Nunez, P. L., Silberstein, R. B., Shi, Z., Carpenter, M. R., Srinivasan, R., Tucker, D. M., Doran, S. M., Cadusch, P. J. and Wijesinghe, R. S. [1999]. EEG coherency II: experimental comparisons of multiple measures., *Clinical Neurophysiology* **110**(3): 469–86.
- [27] Nunez, P. L., Srinivasan, R., Westdorp, a. F., Wijesinghe, R. S., Tucker, D. M., Silberstein, R. B. and Cadusch, P. J. [1997]. EEG coherency. I: Statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales., *Electroencephalography and clinical neurophysiology* **103**(5): 499–515.

- [28] Offner, F. F. [1950]. The EEG as potential mapping: The value of the average monopolar reference, *Electroencephalography and Clinical Neurophysiology* **2**(1-4): 213 – 214.
- [29] Osselton, J. W. [1965]. Acquisition of EEG data by bipolar, unipolar and average reference methods: a theoretical comparison, *Electroencephalography and clinical neurophysiology* **19**(5): 527–8.
- [30] Petersen, K. B. and Pedersen, M. S. [2012]. *The matrix cookbook*, Technical University of Denmark.
- [31] Ranta, R. and Madhu, N. [2012]. Reference Estimation in EEG: Analysis of Equivalent Approaches, *IEEE Signal Processing Letters* **19**: 12–15.
- [32] Salido-Ruiz, R. [2012]. *Problmes inverses contraintes en EEG : applications aux potentiels absolus et l'influence du signal de rfrence dans l'analyse de l'EEG*, PhD thesis, Universit de Lorraine.
- [33] Salido-Ruiz, R., Ranta, R. and Louis-Dorr, V. [2011]. EEG montage analysis in the blind source separation framework, *Biomedical Signal Processing and Control* **6**(1): 77 – 84.
- [34] Stenroos, M., Mäntynen, V. and Nenonen, J. [2007]. A matlab library for solving quasi-static volume conduction problems using the boundary element method, *Computer methods and programs in biomedicine* **88**(3): 256–263.
- [35] Trefethen, L. N. and Bau III, D. [1997]. *Numerical linear algebra*, Vol. 50, Siam.
- [36] Yao, D. [2001]. A method to standardize a reference of scalp EEG recordings to a point at infinity, *Physiological Measurement* **22**: 693–711.
- [37] Yao, J. and Dewald, J. P. [2005]. Evaluation of different cortical source localization methods using simulated and experimental EEG data, *NeuroImage* **25**: 369–382.
- [38] Zhai, Y. and Yao, D. [2004]. A study on the reference electrode standardization technique for a realistic head model, *Computer Methods and Programs in Biomedicine* **76**(3): 229–238.

7. Appendix

Demonstration of \mathbf{T}_O

$$\begin{aligned}\mathbf{T}_O &= \mathbf{x}\mathbf{x}_{CR}^T [\mathbf{x}_{CR}\mathbf{x}_{CR}^T]^{-1} \\ &= \mathbf{x}\mathbf{x}^T \mathbf{T}_{CR}^T [\mathbf{T}_{CR}\mathbf{x}\mathbf{x}^T \mathbf{T}_{CR}]^{-1} \\ &= \mathbf{R}_x \mathbf{T}_{CR}^T [\mathbf{T}_{CR}\mathbf{R}_x \mathbf{T}_{CR}]^{-1} \\ &= \mathbf{A}\mathbf{s}(\mathbf{A}\mathbf{s})^T \mathbf{T}_{CR}^T [\mathbf{T}_{CR}\mathbf{A}\mathbf{s}(\mathbf{A}\mathbf{s})^T \mathbf{T}_{CR}]^{-1} \\ &= \mathbf{A}\mathbf{R}_s \mathbf{A}^T \mathbf{T}_{CR}^T [\mathbf{T}_{CR}\mathbf{A}\mathbf{R}_s \mathbf{A}^T \mathbf{T}_{CR}]^{-1}\end{aligned}\tag{15}$$