Identification of epileptogenic networks in drug-resistant partial epilepsy

Paul Berraute

UNIVERSITÉ DE RENNES 1

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE RENNES 1

ÉCOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : *Signal, Image, Vision*

Par

Paul BERRAUTE

Identification of epileptogenic network in drug-resistant partial epilepsy

Thèse présentée et soutenue à Rennes, le 19/11/2021
Unité de recherche : Laboratoire Traitement du Signal et de l'Image (LTSI), UMR Inserm 1099

Rapporteurs avant soutenance :

Grova Christophe	Professeur	Université Concordia, Montreal
Michel Christoph	Professeur	Université de Genève, Genève

Composition du Jury :

Président(e) :	Meste Olivier	Professeur	Université Côte d'Azur, Côte d'Azur
Examinateurs :	Grova Christophe	Professeur	Université Concordia, Montreal
	Michel Christoph	Professeur	Université de Genève, Genève
Dir. de thèse :	Albera Laurent	Maitre de conférence	Université de Rennes 1, Rennes
Co-dir. de thèse :	Merlet Isabelle	Chargé de Recherche INSERM	Université de Rennes 1, Rennes
	Rochette Michel	Directeur de Recherche Ansys	ANSYS France, Lyon

Invité(s) :

Kachenoura Amar Ingénieur de recherche, Université de Rennes 1, Rennes

To my mom
To my dad
To my family and friends ...

Acknowledgements

I want to thank Mr. Grova Christophe, Associate Professor at the University of Concordia, as well as Mr. Christoph Michel, Professor at the Institute of Geneva, for the honor they have given me by accepting to review this thesis. I also wish to thank Mr. Meste Olivier for having accepted to chair my thesis jury. I am grateful for their time, comments, and suggestions that allowed me to complete this manuscript and develop new perspectives on the proposed work. Finally, I thank them for coming to Rennes to make this moment unique during the crisis of Covid-19.

I want to express my sincere gratitude and thanks to Mr. Albera Laurent, Professor, Ms. Merlet Isabelle, INSERM research fellow, and M. Kachenoura Amar, Research Engineer at the University of Rennes 1, who supervised me throughout this thesis and shared their brilliant intuitions with me. I would also like to thank them for the numerous encouragements they gave me, and their taste for rigor and precision proper to each of their specialties. On a personal level, I want to thank them for their kindness and the extraordinary moments we have shared, such as burger and correction evenings, running in the woods, wellness talks and many other memorable events that made this thesis an unforgettable experience.

This project would not have been possible without the help and financial support of the company Ansys. In particular, I would like to thank Michel Rochette, head of the Ansys research team in Lyon, for his management and investment in this project and the opportunity to discover the international business world. I would also like to thank him for his invitations to annual events, which have been a source of wonder and fruitful exchanges with other Ph.D. students.

I want to thank Mr. Lotfi Senhadji, director of the LTSI and Mr. Fabrice Wendling, SESAME team leader, for having welcomed me in their team. I also thank the permanent members of the LTSI who helped me in many ways. They were my teachers in engineering school and master of research. They were a great help in choosing my orientation from the beginning to the end during those years and their dynamism

and enthusiasm create a unique working atmosphere that is great to be a part of. I also thank them for the opportunities they offer me to continue my work with them and hope our collaboration will remain.

I thank my colleagues who supported me during these four years both new and old faces are always present in my heart and on social networks. It was a great pleasure to see so many of you on the defense day, and I hope we will remain close beyond our adventure at LTSI. A special thanks to the members of the 404 office for their exceptional help and motivation both early in the morning and late at night.

I thank my friends and my dear girlfriend for their moral support and for all those role-playing and movie nights that allowed me to surpass myself in my work.

Finally, I want to thank my family for their unconditional love during those years, especially my father, who never stopped fighting so that his family would not lack anything and support me in every way at every step of my academic journey.

Contents

*

List of Figures

List of Tables

Glossary

$2q$-ExSo-MUSIC $2q$-Extended Source-Multiple Signal Classification.

ADMM Alternating Direction Method of Multipliers.

AGCMN-RUG Array-Gain Constraint Minimum Norm with Recursively Updating Gram matrix.

AOLS Asymptomatically Optimal Linear Shrinkage.

AUC Area Under Curve.

BCI Brain-Computer Interface.

BEM Boundary Element Methods.

BEst Brain Entropy in space and time.

BLUE Best Linear Unbiased Estimator.

BSI-EEG Brain Source Imaging based on EEG.

BSS Blind Source Separation.

cLORETA constrained LORETA.

cMEM coherent Maximum Entropy on the Mean.

CSP Common Spatial Pattern.

DICS Dynamic Imaging of Coherent Sources.

DLE Dipole Localisation Error.

dSPM dynamic Statistical Parametric Maps.

EEG ElectroEncephaloGraphy.

ERP Event-Related Potential.

EZ Epileptogenic Zone.

FIR Finite Impulse Response.

fMRI functional Magnetic Resonance Imaging.

FN False Negative.

FP False Positive.

FPF False Positive Fraction.

GABA Gamma-AminoButyric Acid.

GSC Global Sidelobe Canceller.

HFO High-Frequency Oscillations.

hr-EEG high resolution-EEG.

ICA Independent Component Analysis.

$L_{1,2}$-SISSY $L_{1,2}$-Source Imaging based on Structured Sparsity.

LASSO Least Absolute Shrinkage and Selection Operator.

LCMV Linear Constraints Minimum Variance.

LMMSE Linear Minimum Mean Squared Error.

LORETA Low-Resolution Electromagnetic Tomography.

M-P Marchenko-Pastur.

MAI Multiple Activity Index.

Max SNR Max Signal-Noise-Ratio.

MEG MagnetoEncephaloGraphy.

MER Multiple Event-Related.

MIA Multiple-step Iterative Approach.

MMSE Minimum Mean Square Error.

MNE Minimum Norm Estimate.

MP Minimum Power.

MPZ Multiple Pseudo Z.

MRI Magnetic Resonance Imaging.

MUSIC Multiple-Signal Classification.

MV Minimum Variance.

MV-PURE Minimum Variance - Pseudo Unbiased low-Rank Estimator.

NAI Neural Activity Index.

OAS Oracle Approximating Shrinkage.

PCA Principal Component Analysis.

PLS Partial Least-Squares.

pSPM power Statistic Parametric Map.

PZ Pseudo-Z.

RAP-MUSIC Recursively Applied and Projected MUSIC.

rMER reduced Multiple Event-Related.

RMT Random Matrix Theory.

ROC Receiver Operating Characteristic.

ROI(s) Region(s) Of Interest.

RSA Ratio of distant Spurious Activity.

rZER reduced localizer to the Event-Related.

SABLE Sparsity And Beamforming for brain source Localization and Estimation.

SAM Synthetic Aperture Magnetometry.

SCM Sample Covariance Matrix.

SD Spatial Dispersion.

SEEG Stereotactic EEG.

SIA Single-step Iterative Approach.

SISSY Source Imaging based on Structured Sparsity.

sLORETA standardized LORETA.

SNR Signal to Noise Ratio.

SOI(s) Source(s) Of Interest.

SOSE Second-Order wide-sense Stationary Ergodic.

SPA Single-step Peak Approach.

SSP Signal Space Projection.

SSS Signal Space Separation.

Stdr Standardized minimum variance beamformer.

SVB-SCCD Sparse Variation-Based Sparse Cortical Current Distribution.

SVD Singular Value Decomposition.

tDCS transcranial Direct-Current Stimulation.

TP True Positive.

TPF True Positive Fraction.

TSVD Truncated Singular Value Decomposition.

UG Unit Gain.

UNG Unit Noise Gain.

VB-SCCD Variation-Based Sparse Cortical Current Distribution.

wMNE weighted MNE.

ZER Localizer to the Event-Related.

Nomenclature

Scalars

α	SISSY hyper-parameters
$\alpha(\theta_\ell)$	Lagrangian multiplier value for the position θ_ℓ
α_i	Shrinkage trade-off variable
δ_{thr}	Threshold value for the EEG similarity metric
λ	Generic hyper-parameters
μ	Regularization parameter of the diagonal loading preprocessing method
τ	Time lag
$\tau(\theta_\ell)$	Scalar factor for the position θ_ℓ
$\widehat{s}(\theta_\ell, t)$	Estimated source activity at the position θ_ℓ and time t
a_ℓ^+	Factor used in AOLS definition
a_ℓ^-	Factor used in AOLS definition
c	Ratio of the number of electrodes over the number of time samples
f	Frequency variable
f_0	Specific frequency f_0
K	Number of SOI
L	Number of positions (vertices of the triangular mesh)
L_k	Number of dipole for the k-th SOI
$L_{\mathcal{M}}$	Number of original active dipole
$L_{\widehat{\mathcal{M}}}$	Number of estimated active dipole
N	Number of sensors
T	Number of time samples
t	Time variable

Vectors

$\boldsymbol{\alpha}(\theta_\ell)$	Lagrangian vector for the AOLS
$\boldsymbol{\nu}$	Noise vector
$\boldsymbol{\nu}(t)$	Realization of the noise random process at time t
$\boldsymbol{\phi}_k$	Spatial support of the k-th SOI
$\boldsymbol{s}$	Brain source activity vector
$\boldsymbol{s}^{(b)}$	Source vector for the the background
$\boldsymbol{s}_k^{(e)}$	Source vector for the k-th SOI
$\boldsymbol{\theta}$	Vector regrouping multiple dipole position
$\boldsymbol{\theta}_{\text{ref}}$	Vector regrouping the locations of the previously identified SOIs
$\boldsymbol{x}$	EEG vector
$\hat{\boldsymbol{s}}_{\text{wMNE}}$	Source activity estimated by wMNE
$\bar{s}_k^{(e)}$	Mean electrical activity of the k-th SOI
$\boldsymbol{h}_f(t)$	Impulse response for frequency f
$\boldsymbol{r}$	Desired beamformer response
$\boldsymbol{y}(t)$	Independant components of interest
$\boldsymbol{z}(f)$	Fourier transform of $\boldsymbol{x}(t)$

$\boldsymbol{s}(t)$ Realization of the source random process at time t

$\boldsymbol{w}(\theta_\ell)$ Beamforming band filtering weights for position θ_ℓ

$\boldsymbol{x}(n)$ EEG time course for electrode n

$\boldsymbol{x}(t)$ Realization of the EEG signals random process at time t

$\boldsymbol{x}_{\overline{n}}(t)$ Averaged data across epochs

Matrices

$\boldsymbol{A}$ ICA mixing matrix

$\boldsymbol{\Gamma}(\theta_\ell)$ Filter basis for the AOLS

$\boldsymbol{K}$ Diagonal weight matrix

$\boldsymbol{T}$ Space gradient operator

$\boldsymbol{W}_{\text{wMNE}}$ Kernel of wMNE

$\boldsymbol{G}$ Leadfield matrix

$\boldsymbol{G}^{(b)}$ Leadfield matrix for the background dipoles

$\boldsymbol{G}_k^{(e)}$ Leadfield matrix for the k-th SOI

$\boldsymbol{W}$ Matrix regrouping the beamforming filter from all position

$\Sigma_k^{\mathbf{B}}$ Matrix of singular values or eigenvalues of matrix $\mathbf{B}$ thresholded at rank k

$\boldsymbol{B}$ Any given matrix

$\boldsymbol{D}$ Temporal-Smoothing operator

$\boldsymbol{L}$ Surface Laplacian matrix

$\boldsymbol{P}(\boldsymbol{\theta}_{\text{ref}})$ Projector for the RAP beamforming

$\boldsymbol{U}_k^{\mathbf{B}}$ Matrix of left subspace vectors of matrix $\mathbf{B}$ thresholded at rank k

$\boldsymbol{V}_k^{\mathbf{B}}$ Matrix of right subspace vectors of matrix $\mathbf{B}$ thresholded at rank k

$\boldsymbol{E}$ Matrix used in MV-PURE beamforming definition

$\widehat{\boldsymbol{C_\nu}}$ Noise Sample Covariance Matrix

$\widehat{\boldsymbol{C_\nu}}^{-\frac{1}{2}}$ Whitening operator

$\widehat{\boldsymbol{C_{\overline{x}}}}$ Averaged data Sample Covariance Matrix

$\widehat{\boldsymbol{C_x}}$ Data Sample Covariance Matrix

I_d Identity matrix of dimension d

Sets

$\boldsymbol{C_x}$ Unknown theoretical data covariance matrix

Γ Set of frequencies

$\Omega_{\mathcal{M}}$ Original set of indexes of all dipoles of an active patch

$\Omega_{\boldsymbol{G}}$ Set of constraints applied to the leadfield matrix

$\Omega_{\boldsymbol{w}_{\mathcal{L}}}$ Set of linear constraints applied to the beamforming filters

$\Omega_{\boldsymbol{w}_{\mathcal{Q}}}$ Set of quadratic constraints applied to the beamforming filters

$\Omega_{\boldsymbol{w}_{\mathcal{RR}}}$ Set of reduced rank constraints applied to the beamforming filters

$\Omega_{\boldsymbol{w}_{\mathcal{S}}}$ Set of shrinkage constraints applied to the beamforming filters

$\Omega_{\boldsymbol{w}}$ Set of constraints applied to the beamforming filters

$\Omega_{\widehat{\mathcal{M}}}$ Estimated set of indexes of all dipoles of an active patch

Ω_k Set of the dipoles indexes for the k-th SOI

$\{\boldsymbol{\nu}(t)\}$ Second order wide-sense stationary ergodic noise random process

$\{\boldsymbol{s}(t)\}$ Second order wide-sense stationary ergodic source random process

$\{\boldsymbol{x}(t)\}$ Second order wide-sense stationary ergodic EEG signals random process

Norms and modifiers

$\mathbb{E}[...]$ Mathematical expectation

$||...||_2$ Euclidian norm

$||...||_{\mathsf{F}}$ Frobenius norm

$||...||_\infty$ Uniform norm or supremum norm

$||...||_1$ L_1 norm

Exponents and indices

$\hat{a}$ Estimate of a

$\overline{a} = \mathbb{E}[a]$ Mathematical expectation of a

$\widetilde{a}$ Reformulation of a

$a^{\dagger}$ Pseudo-inverse of a

$a^{-\frac{1}{2}}$ Inverse of the square root of a

$a^{\intercal}$ Transposition of a

a_n Iteration step n of a

a_{n+1} Iteration step $n+1$ of a

Functions and Metrics

$*$ Convolution operator

δ_p Dirac distribution of the parcel p

$\gamma(..., \theta_{\text{ref}})$ Coherence metric

$\gamma_1(..., n_{\text{ref}})$ EEG similarity metric - step 1

$\gamma_2(..., n_{\text{ref}})$ EEG similarity metric - step 2

$\kappa(...)$ Kurtosis metric

$\mathcal{L}(...)$ Lagrangian

$\mathcal{N}(\mu_p, \boldsymbol{C}_p)$ Gaussian distribution with mean and variance of the parcel p

ρ_p Probability of activation of the parcel p

$\sigma(...)^2$ Variance metric

$d_{...}^{Euc}$ Euclidian distance

$dp(\boldsymbol{s})$ Source probability distribution

$dv_p(\boldsymbol{s}_p)$ Reference distribution for the parcel p

$f(...)$ Any given function

$m(...)$ Neural Activity Index metric

$m_{\text{M}}(...)$ Multiples Activity Index metric and associate

$m_{\text{RAP}}(...)$ Iterative RAP beamforming metric

m_{sLORETA} sLORETA metric

$p(..., f)^2$ Power statistic parametric map metric for a frequency f

$SNR(...)$ SNR metric

French extended summary of the thesis

Résumé de la thèse

Contexte et motivation

L'épilepsie est l'une des maladies neurologiques les plus courantes qui affectent environ 2,4 millions de personnes dans le monde. Il s'agit d'une entité clinique hétérogène et multifactorielle caractérisée par l'apparition soudaine d'une activité électrique hyper-synchronisée dans une ou plusieurs régions corticales. Cette activité anormale entraîne des dysfonctionnements cérébraux qui ont un impact sur l'intégrité physique, la santé mentale et la vie sociale du patient. Dans la plupart des cas, les patients épileptiques peuvent être traités avec succès par des médicaments, qui empêchent les crises d'épilepsie ou permettent d'en réduire la fréquence. Toutefois, 20 à 30% de ces patients sont dit pharmaco-résistant, ce qui signifie qu'ils ne montrent aucun signe d'amélioration après la prise de médicaments. Pour ces patients, une intervention chirurgicale ou une stimulation transcrânienne peuvent être envisagées pour arrêter ou atténuer les crises. Cependant, ces alternatives thérapeutiques nécessitent une délimitation précise de la zone épileptogène.

Avant de subir une intervention chirurgicale ou une stimulation, les patients doivent généralement se soumettre à une évaluation pré-chirurgicale approfondie, comprenant des séances d'électroencéphalogramme (EEG) de longue durée et des enregistrements d'EEG stéréotaxique (SEEG). L'EEG est une modalité plutôt bon marché directement liée à l'activité électrique du cerveau car elle enregistre les potentiels de champs locaux produits par des milliers de neurones qui se sont propagés jusqu'au cuir chevelu. Cette diffusion électrique entre la surface corticale et les capteurs EEG du scalp peut être modélisée avec précision en fournissant un modèle de tête réaliste et un modèle de source puis en résolvant l'approximation quasi-statique des équations de Maxwell. Ce processus appelé le problème direct crée la matrice dite de leadfield qui réduit l'EEG à un système linéaire. Le processus inverse, cependant, est plus compliqué. En effet, la localisation de la zone épileptogène et la reconstruction de son activité sont possibles en résolvant le système linéaire EEG. Malheureusement, comme nous ne disposons pas d'autant d'électrodes que de sources, ce système est

un problème inverse mal posé au sens de Hadamard et nécessite des méthodes indirectes pour être résolu.

L'imagerie des sources cérébrales basée sur l'EEG (BSI-EEG), est la discipline qui regroupe l'ensemble des méthodes permettant de résoudre le problème inverse en EEG. C'est un domaine en constante évolution grâce aux avancées en neurophysiologie, en mathématiques et en algorithmique. Ces découvertes sont à l'origine des aprioris qui sont des connaissances nécessaires pour développer une méthode adéquate. La classification de toutes les approches d'imagerie des sources cérébrales peut donc se faire selon trois aspects :

1. D'un point de vue physique, selon l apriori formulé sur l'espace source (paramétrique/non paramétrique) ou la distribution (distribution lisse/parcimonieuse) ;
2. D'un point de vue mathématique, selon la formulation du problème d'optimisation (moindres carrés régularisés/tenseur) ou la forme de la solution (image estimée/noyau d'inversion) ;
3. D'un point de vue technique, selon le niveau de développement et d'accessibilité de la méthode (mature/expérimentale).

Pour l'épilepsie, le premier problème est de choisir une méthode qui soit compatible avec les caractéristiques de la maladie. En l'état actuel de la littérature, il est bien admis que l'épilepsie est physiologiquement mieux représentée par un ensemble de zones corticales étendues plus ou moins connectées avec une activité est hypersynchronisée. Ce modèle, appelé modèle de source distribuée, implique l'utilisation de méthodes non paramétriques. Cependant, il n'est pas rare de voir des méthodes paramétriques utilisées en pratique car elles peuvent néanmoins fournir une solution de départ à d'autres algorithmes. De plus, certaines méthodes paramétriques exploitent des hypothèses spéciales qui ne sont pas encore prises en compte pour le modèle de source distribuée. Malheureusement, elles ont parfois été mal utilisées pour coller à ce modèle sans considérer les inconvénients de l'approche, comme dans le cas du beamforming.

Le deuxième problème se pose après le choix de la méthode : il s'agit de sa mise en pratique. En effet, selon la méthode choisie, le degré de développement peut varier d'un simple article original comme support (méthode expérimentale) à une solution logicielle prête à l'emploi (méthode mature). Cette dernière peut également bénéficier de l'aide d'une communauté et d'une documentation solide qui répond aux questions de réglage de l'algorithme. En revanche, les méthodes expérimentales peuvent connaître un destin aléatoire allant de l'oubli à la valorisation par un petit nombre d'utilisateurs. Pour éviter cela, il est de bonne pratique de fournir un pipeline détaillé en complément de la nouvelle méthode. En effet, le pipeline de la méthode joue un rôle crucial dans la localisation et la reconstruction des sources cérébrales et

assure la reproductibilité des expériences pour les utilisateurs externes.

Objectifs

L'objectif principal de cette thèse est d'étendre une famille de méthodes paramétriques appelée beamforming pour résoudre le problème inverse en BSI-EEG pour l'épilepsie. Afin d'atteindre cet objectif, nous i) faisons abstraction du modèle de sources adapté à l'épilepsie pour classifier, implémenter, tester et améliorer les méthodes de beamforming dans un cadre optimal et ii) proposons une version nouvelle et bien adaptée du beamforming pour le modèle de source distribuée. Des évaluations quantitatives et qualitatives sur des données simulées et réelles seront fournies à chaque étape ainsi que des pipelines détaillés pour faciliter la diffusion des méthodes proposées en dehors de la phase expérimentale.

Contributions

La première partie de cette thèse est consacrée à l'étude et à l'évaluation des différentes pratiques de beamforming utilisées en BSI. En effet, les méthodes de beamforming sont employées depuis de nombreuses années en magnétoencéphalographie et récemment en EEG. Malheureusement, le domaine du beamforming étant large, des raccourcis méthodologiques peuvent être observés dans certains articles, comme l'utilisation de noms inappropriés pour une méthode ou la mauvaise utilisation dans le modèle des sources distribuées. De plus, les étapes de prétraitement et de post-traitement constituant le pipeline sont souvent négligées, alors qu'elles sont la clé de la précision des résultats. Cette partie a pour but d'offrir un résumé synthétique et complet de l'état de l'art du beamforming pour tous les utilisateurs. Une fois toutes les pratiques identifiées, une étude comparative est réalisée entre les différents pipelines les mieux adaptés à une application réelle. Une partie de cette étude a déjà été partagée lors de la conférence Virtual Physiological Human (VPH) 2020 [2]. Cette partie est complétée par la proposition d'un nouveau terme de régularisation pour les méthodes de beamforming qui améliore de manière significative la reconstruction des sources réelles. Cette addition s'applique à de nombreuses méthodes de beamforming que nous avons rassemblées sous le nom de beamforming basée sur le lissage temporel.

La deuxième partie de cette thèse est consacrée à la proposition d'une nouvelle méthode de beamforming nommée SABLE (*Sparsity And Beamforming for brain source Localization and Estimation*) pour le modèle de source distribuée adapté à l'épilepsie. Cette méthode sera présentée depuis les hypothèses mathématiques jusqu'au pipeline utilisé pour la rendre mature et la préparer à une diffusion plus large. Une étude préliminaire sur SABLE a été publiée lors de la conférence IEEE Statistical Signal Pro-

cessing (SSP) [3]. Ce travail est conclu par une comparaison qualitative entre SABLE et trois méthodes matures implémentées dans le logiciel Brainstorm pour démontrer ses capacités à localiser trois sources étendues présentant une activité corrélée.

Contenu du manuscrit de thèse

Dans le chapitre 1, nous présentons le contexte de la localisation de source a partir de signaux EEG. À cette fin, nous avons présenté successivement l'anatomie, l'histologie et la physiologie du cerveau jusqu'à une brève description de l'épilepsie. Nous avons également mis l'accent sur le modèle EEG et les signaux caractéristiques de l'épilepsie. Ces signaux sont les éléments clés du diagnostic et sont indispensables pour localiser la zone épileptogène. En conclusion, un bref état de l'art sur le problème inverse a été présenté afin d'introduire les méthodes de beamforming.

Dans le chapitre 2, nous avons recensé toutes les méthodes de beamforming existantes et analysé les pratiques dans la littérature. Le résultat de ce travail a donné lieu à trois tableaux dont le contenu combiné permet de couvrir l'ensemble des pratiques en BSI-EEG. De nouvelles approches ont également été introduites pour compléter celles existantes. L'analyse comparative réalisée sur des données simulées EEG épileptiques réalistes a montré que la méthode de rétrécissement de Ledoit (*Ledoit's shrinkage method*), issue de la théorie des matrices aléatoires, est le meilleur prétraitement pour le beamforming. Le beamforming AOLS (*Asymptomatically Optimal Linear Shrinkage*), qui utilise une autre forme de rétrécissement, améliore également les résultats pour les sources non corrélées. Enfin, la métrique de similarité EEG nouvellement introduite s'avère être la meilleure méthode de post-traitement. Appliqué à des données réelles, ce pipeline a permis de localiser efficacement la zone épileptogène. Néanmoins, les signaux reconstruits présentent quelques imperfections car ils sont contaminés par l'activité d'autres sources. Par conséquent, nous avons proposé une nouvelle méthode régularisée appelée beamforming basée sur le lissage temporel (*Temporal-Smoothing-based beamforming*) pour améliorer la reconstruction de la source sur des données réelles. Le développement mathématique montre que cette méthode peut être facilement mise en œuvre car elle n'a d'impact que sur la matrice de covariance des données. Cette contrainte améliore le conditionnement de la matrice de covariance en plus d'obliger la distribution des sources estimées à être lisse. Les résultats de l'estimation des sources des méthodes de beamforming basées sur le lissage temporel présentent une amélioration significative sur les données simulées et réelles.

Dans le chapitre 3, nous avons présenté une nouvelle méthode qui étend les méthodes traditionnelles de beamforming au modèle de source distribuée adapté à l'épilepsie. En effet, les méthodes classiques de beamforming sont plus généralement appliquées au modèle de source équivalent pour localiser et reconstruire des activités de source

non corrélées. Au-delà des utilisations abusives, notre revue recense les tentatives sérieuses précédentes d'utilisation de la beamforming au modèle de source étendu. Cependant, elles nécessitent de connaître la position des sources à l'avance, ce qui est une hypothèse trop forte en pratique. La nouvelle méthode SABLE surmonte ce problème en utilisant une variable latente et en exploitant les méthodes UG (*Unit Gain*) et SISSY (*Source Imaging based on Structured Sparsity*). Les résultats préliminaires des simulations montrent que SABLE est plus efficace que ces deux méthodes pour localiser une source distribuée pour des valeurs de ratio signal sur bruit (SNR) élevées. Cependant, elle nécessite un réglage fin de l'hyper-paramètre entre chaque fonction de coût pour des valeurs de SNR faibles. Afin d'améliorer ces résultats préliminaires, nous présentons un pipeline pour SABLE et évaluons ses capacités par rapport à trois méthodes matures de la littérature, à savoir wMNE (*weighted Minimum Norm Estimate*), sLORETA (*standardized Low-Resolution Electromagnetic Tomography*) et cMEM (*coherent Maximum Entropy on the Mean*). Bien que ce pipeline soit fortement inspiré de ceux implémentés dans Brainstorm afin de fournir une comparaison équitable, il contient également de nouvelles pratiques telles que l'usage du fondu enchainé, la suppression du corps calleux et l'usage d'un nouveau critère de sélection du rang. Les résultats quantitatifs sur simulations montrent que SABLE est, d'un point de vue général, aussi efficace que SISSY et surpasse les trois méthodes traditionnelles wMNE, sLORETA et cMEM, pour localiser trois patchs étendus avec une activité synchronisée. Plus précisément, en utilisant le nouveau critère nommé le ratio d'amplitude par patch (*Patch Amplitude Ratio*), nous montrons que la partie beamforming de SABLE permet une meilleure distribution de l'énergie entre les patchs, ce que SISSY ne permet pas. Cette caractéristique permet de mettre en évidence des zones épileptogènes potentielles cachées par un défaut de l'algorithme, comme nous l'avons constaté sur des données réelles.

Perspectives

Cette thèse se situe à l'interface de trois domaines : physique, mathématique et algorithmique. Le travail présenté a été réalisé en tenant compte de ces trois dimensions avec leurs contraintes uniques. Les perspectives peuvent être distinguées entre celles concernant les méthodes et celles qui concernent les pipelines. En ce qui concerne la perspective des méthodes, la première qui pourrait résulter de l'ensemble du travail de thèse est celle de la proposition d'une fusion entre la beamforming basée sur le lissage temporel et SABLE. Cependant, comme l'ont montré les résultats du cas réel, il n'y a aucun intérêt à les fusionner puisque les résultats de reconstruction de SABLE sont déjà lisses. De plus, cela nécessiterait de régler encore plus d'hyper paramètres, ce qui pour le moment reste à régler à la main. Une deuxième perspective pour SABLE serait de tester son effet en combinaison avec la méthode de beamforming avec normalisation de la leadfield (*array gain*). En effet, si l'on prend les méthodes wMNE

et sLORETA, toutes deux ont introduit un a priori sur la profondeur de la source au moyen d'une matrice de poids. Cette hypothèse peut être à l'origine des résultats observés sur des données réelles et est similaire à la méthode de normalisation de la leadfield. Comme l'ajout d'une telle méthode de prétraitement n'est pas simple en raison de problèmes d'échelle, il serait intéressant de proposer et d'évaluer une nouvelle forme de la méthode SABLE. Enfin, nous devons étendre notre étude qualitative à un ensemble plus important de données réelles pour terminer le travail de valorisation de SABLE. Il ne nous restera plus qu'à renforcer le pipeline pour garantir la stabilité des résultats, ce qui fait l'objet du paragraphe suivant.

En ce qui concerne les perspectives du pipeline, un travail supplémentaire est justifié pour étudier de nouvelles stratégies de réglage des hyper paramètres. En effet, la méthodologie proposée dans la méthode SISSY est assez particulière et la fonction de l'hyper paramètre est différente des méthodes traditionnelles [1]. En raison de la substitution des variables dans les normes L_1, l'hyper paramètre entre les contraintes et la fonction de coût joue maintenant un rôle de seuillage dans l'opérateur proximal. Bien que ce nouveau rôle soit intéressant, il ne compense pas la perte de la capacité à mettre l'accent sur la fidélité des données ou sur les contraintes en fonction du niveau de bruit. Ainsi, nous devons revoir la stratégie des hyper paramètres pour SISSY et par extension pour SABLE. De plus, nous aimerions parfaire la méthode de sélection des rangs proposée dans le pipeline pour catégoriser plus finement les valeurs propres appartenant aux sources d'intérêt, aux bruits structurés et aux bruits Gaussien. Ces hypothèses sont déjà utilisées dans le domaine du traitement d'antennes et sont prometteuses dans le domaine de l'imagerie de sources cérébrales, notamment sur les questions de débruitage [2]. Une fois ces deux sujets abordés, nous serons en mesure de proposer un pipeline entièrement automatisé pour SABLE.

Introduction

Context and motivation

Epilepsy is one of the most common neurological diseases that affects about 2.4 million people worldwide. It is a heterogeneous and multifactorial clinical entity characterized by the sudden occurrence of a hyper-synchronized electrical activity in one or more cortical regions. This abnormal activity leads to brain dysfunctions that impact the physical integrity and the patient's mental health and social life. In most cases, epileptic patients can be successfully treated with drugs, which prevents epileptic seizures or allows to reduce their frequency. However, 20-30% of these patients are drug-resistant, meaning they show no sign of improvement after medication. These cases are among the most severe and often originate from one brain region or a delimited network of connected areas. Surgical intervention or transcranial Direct-Current Stimulation (tDCS) can be considered to stop or attenuate seizures. However, those therapeutic alternatives require the precise delimitation of the epileptogenic zone.

Before surgery or transcranial stimulation, patients must usually undergo extensive pre-surgical evaluation, including long-term ElectroEncephaloGraphy (EEG) sessions and Stereotactic EEG (SEEG) recordings. EEG is a cheap modality directly related to the brain's electrical activity as it records the local field potentials produced by thousands of neurons that have spread to the scalp. This electrical diffusion between the cortical surface and the scalp EEG sensors can be accurately modeled by providing a realistic head model and a source model and solving the quasi-static approximation of Maxwell equations. This process called the forward problem (see figure 1) creates the so-called leadfield matrix that reduces the EEG mathematically as a linear system of equations. The other way around, however, is more complicated. Indeed, the localization of the epileptogenic zone and the reconstruction of its activity are possible by solving the EEG linear system. Unfortunately, since we do not have as many electrodes as sources, this system is an ill-posed inverse problem in the Hadamard sense and requires indirect methods to be solved.

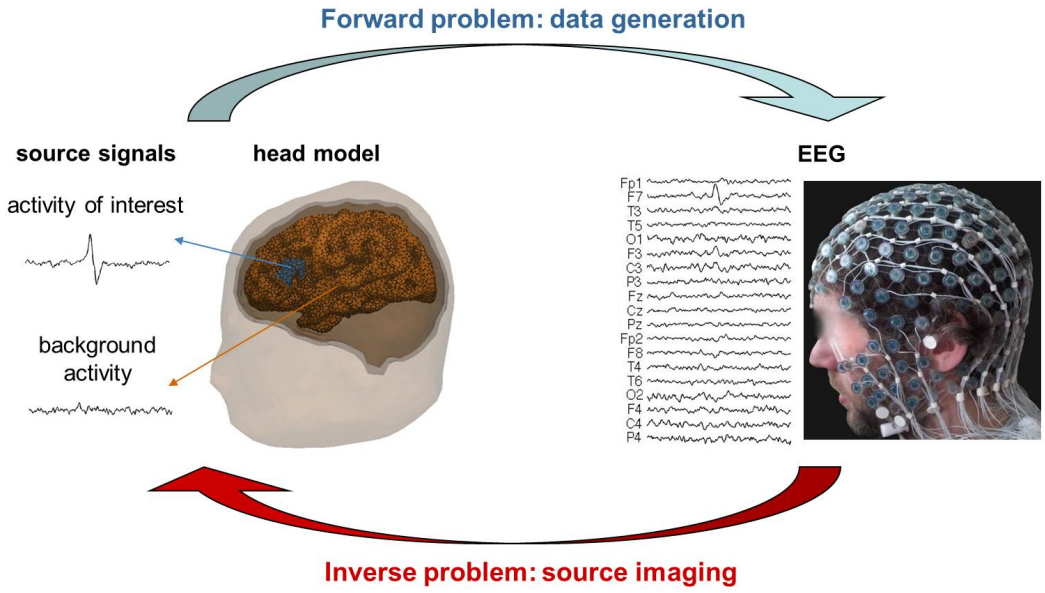


Figure 1: Illustration of the EEG forward and inverse problems.

Brain Source Imaging based on EEG (BSI-EEG), is the discipline that regroups all the methods that solve the inverse problem in EEG. It is a constantly evolving field through advances in neurophysiology, mathematics and algorithmic. These discoveries are at the origin of priors which are necessary knowledge for developing an adequate method. The classification of all BSI-EEG approaches can therefore be done according to three aspects:

1. From a physical point of view, according to the prior formulated on the source space (parametric/non-parametric) or distribution (smoothness/sparsity);
2. From a mathematical point of view, according to the optimization problem formulation (regularized least-squares/tensor) or the solution form (estimated image/inversion kernel);
3. From an engineering point of view, according to the method's level of development and accessibility (mature/experimental).

For epilepsy, the first problem is to choose a method that is compatible with the disease behaviors. As the literature stands, it is well accepted that epilepsy is physiologically better represented by a set of connected cortical areas with a hyper-synchronized activity. This model, called distributed source model, implies using non-parametric methods. However, it is not uncommon to see parametric methods used in practice as they can provide a good first solution to the inverse problem. In addition, some of them exploit specific assumptions which are not yet considered for the distributed source model. Unfortunately, they have sometimes been misused to fit the opposite model without considering the approach drawbacks, like in the case of beamforming.

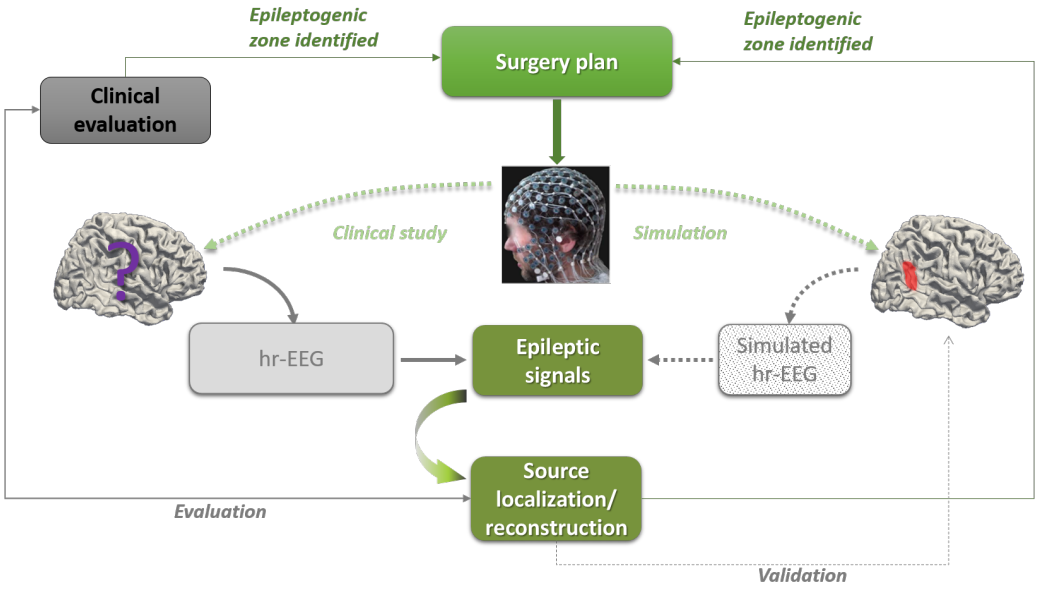


Figure 2: Overall scheme of the work carried out during the thesis.

The second problem arises after the method has been chosen: it is its practical implementation. Indeed, depending on the selected method, the degree of the development may vary from a mare original paper as support (experimental method) to a ready-made software solution (mature method). The latter may also have the help of a community and solid documentation that answers algorithm tuning issues. In contrast, the former may have a random fate ranging from being forgotten to being valued by a small number of users. To prevent this, it is good practice to provide a detailed pipeline to complement the new method. Indeed, the method pipeline plays a crucial role in the localization and reconstruction of brain sources and ensures the reproducibility of experiments for external users.

Objectives

The main objective of this thesis is to extend a family of parametric methods called beamforming to solve the inverse problem in BSI-EEG for epilepsy. In order to achieve this goal, we proceed to i) disregard the source model suitable for epilepsy to classify, implement, test and improve beamforming methods in an optimal setting and ii) suggest a novel and well-adapted version of beamforming for the distributed source model. Quantitative and qualitative evaluations on simulated and real data will be provided on each step (see figure 2) as well as detailed pipelines to facilitate the release of the proposed methods outside the experimental phase.

Contributions

The first part of this thesis is dedicated to study and assess the different beamforming practices used in BSI. Indeed, beamforming methods have been employed for many years in MagnetoEncephaloGraphy (MEG) and recently in EEG. Unfortunately, since the field of beamforming is broad, methodological shortcuts can be observed in some papers, such as the use of inappropriate names for a method or the misuse in distributed sources model. Moreover, the preprocessing and postprocessing steps constituting the pipeline are often overlooked, although these are the key to accurate results. This part aims to offer a synthetic and comprehensive summary of the beamforming state-of-the-art for all users. Once all practices have been identified, a comparative study is carried out between different pipelines best suited for a real application. Part of this study has already been shared in the Virtual Physiological Human 2020 conference [3]. This part is completed by proposing a novel regularization term for beamforming methods that significantly improve the reconstruction of real sources. This fine addition applies to many beamforming methods that we have gathered under the name Temporal-Smoothing-based beamforming.

The second part of this thesis is dedicated to proposing a new beamforming method named Sparsity And Beamforming for brain source Localization and Estimation (SABLE) for the distributed source model suitable for epilepsy. It will be presented from the mathematical priors to the pipeline used to mature it and prepare it for wider distribution. A preliminary study on SABLE was published in the IEEE Statistical Signal Processing Workshop [4]. This work is concluded with a qualitative comparison between SABLE and three mature methods implemented in the Brainstorm software to demonstrate its capacities to locate three extended correlated sources.

Thesis outline

Chapter 1 presents the background knowledge necessary for understanding this thesis, which includes presenting the different brain structures, electrogenesis, epilepsy, EEG modeling and the various methods of solving the inverse problem. **Chapter 2** presents the detailed state-of-the-art of beamforming methods. It is completed by a comparative study of the approaches requiring the least prior on real data and then on simulated data. It is concluded by presenting a simple improvement that can be applied to almost all beamforming methods. **Chapter 3** presents a new beamforming method named SABLE designed for a suitable source model for epilepsy. The chapter concludes with the presentation of a pipeline applicable to simulated and real data and a comparative study between mature inverse problem methods and SABLE.

Chapter 1

Physiological origin and modeling of EEG signals

This first chapter provides general information necessary to understand the context of this thesis. In section 1.1, basic aspects of brain anatomy, histology and functions are presented. The origin of electromagnetic brain signals and the mechanisms by which they are generated are described. In section 1.2, we give an introduction to epilepsy and its effect on brain signals. In section 1.3, the EEG is presented alongside the headsets configurations used in this thesis. Then, we formulate the mathematical model for the EEG in the specific context of epileptic signals in section 1.4. This model will be the cornerstone for designing the simulations as well as for processing the real data. Finally, section 1.5 provides a short state-of-the-art review of the various methods used to reconstruct and localize brain electrical signals from surface EEG.

1.1 Short review on the brain structure and function

1.1.1 Brain anatomy

The human brain is composed of several structures generally heavily protected by bones [5]. From an anatomical viewpoint (figure 1.1), they are organized from caudal to cranial (neck to the top of the head) [6, 7] as:

— The medulla

— The pons

— The cerebellum

— The mesencephalon

— The diencephalon

— The telencephalon (also called the cerebrum)

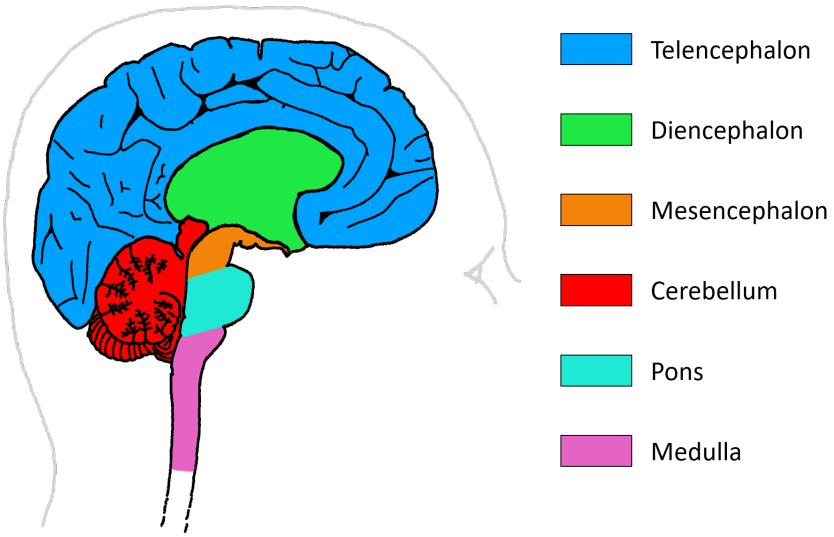


Figure 1.1: Representation of the human brain structures.

The telencephalon of the human species is the most highly developed of all vertebrates [8] due to its evolutionary needs. It provides the capacity for a higher level of information processing and a wider range of complex social emotions [9]. As a result, it takes up the most space inside the head. Since it is also the most external, its activity would be more likely to be recorded by EEG or MEG external sensors than other structures. The telencephalon is, therefore, the structure of interest with regard to EEG in this thesis (section 1.3).

The telencephalon comprises two hemispheres linked by a set of nerve fibers called the corpus callosum. Each hemisphere corresponds to a thick layer of neuronal tissues, folded into ridges (gyri) and grooves (sulci). From an external perspective, the central sulcus, the lateral sulcus, and the preoccipital notch divide each hemisphere into the frontal, parietal, temporal, and occipital lobes (figure 1.2a). From an internal perspective, the limbic lobe (1.2b) is a hidden arc-shaped region organized above and below the corpus callosum. It roughly follows the cingulate sulcus.

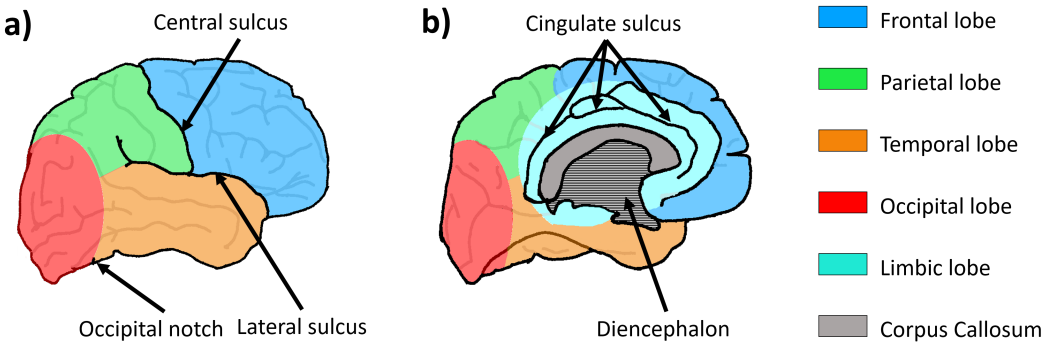


Figure 1.2: Representation of the human brain lobes from an outside viewpoint in a) and an inside viewpoint in b).

From a practical perspective, when one wishes to refer to a specific cortical area, it uses a purely anatomical mapping or atlas defined by each gyrus and sulcus name [10]. Additionally, brain atlases can also rely on the cellular organization of a cortical region [11, 12] or on a physiological function [13] which are of clinical and functional interest. It is noteworthy that all atlases do not necessarily overlap with each other but are generally identical from one person to another. In this thesis, visual support will be provided with the name of the targeted brain area when necessary.

The coronal view in figure 1.3 shows the internal organization of the telencephalon. It reveals that the cerebrum comprises an outer $2-5\ mm$ layer of grey matter called the cortex and an inner layer of white matter. Various grey matter nuclei forming the basal ganglia are located within the white matter. The coloration white/grey corresponds to different cell organization as visible in the box of Figure 1.3. As we will see in the next section, these cells structured in networks are responsible for brain activity.

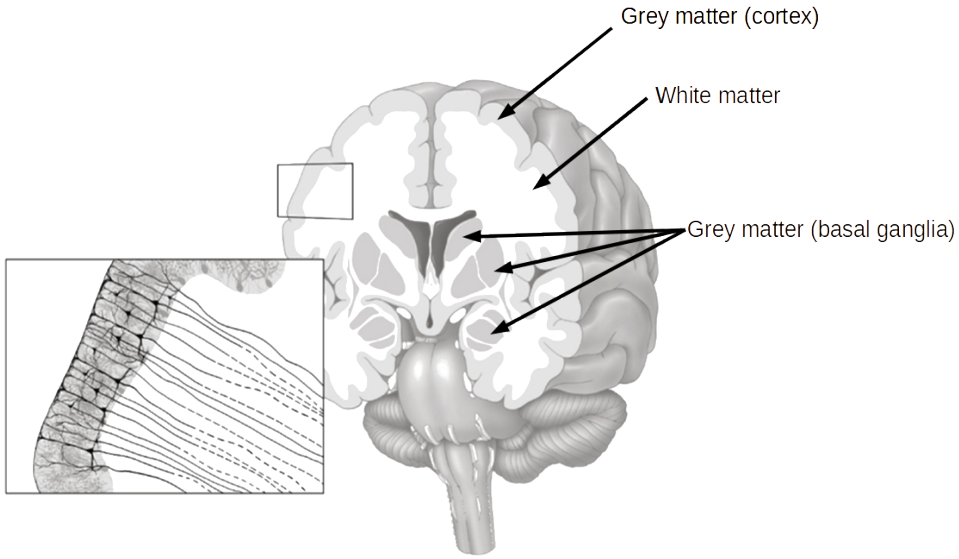


Figure 1.3: Connection between cortex and brain structure. [1]

1.1.2 Brain histology

The brain is made up of biological tissue unique to the nervous system called nervous tissue. The nervous tissue comprises two types of cells: neurons and glial cells, which contribute to creating and transmitting the electrochemical activity responsible for brain functions. Neurons generate electricity either automatically or in response to stimulation. They are present in large quantities (of the billion order [15]) in the cortex, basal ganglia and other brain structures. They present a wide variety of morphology (unipolar, bipolar, pyramidal), function (excitatory, inhibitory) and

1. Public domain picture available at [14]

electrical properties, which make them difficult to classify [16]. Nevertheless, despite this diversity, it is well accepted that almost all neurons present the same element of architecture with four distinctive parts (figure 1.4):

— The dendrite, which receives information from other cells;

— The cell body, also called soma, ensures the survival and functioning of the cell;

— The axon, which transmits electrical impulses to other cells;

— The axon terminal contains multiple connection points between numerous neurons.

Those connections are called synapses and involve complex chemical mechanisms to transfer the electrical activity from one neuron to another.

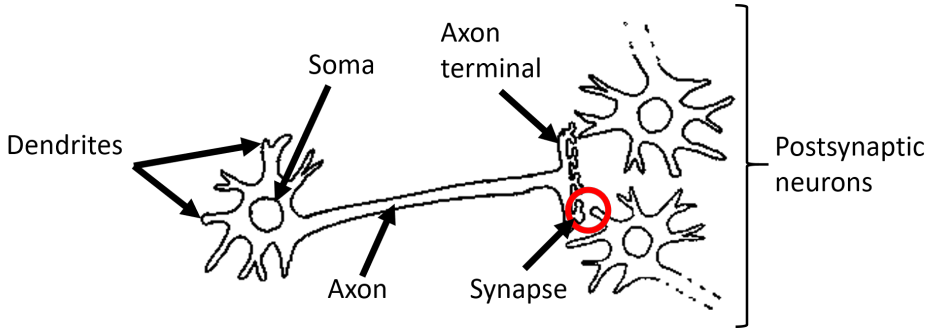


Figure 1.4: Schematic representation of the different parts of a neuron and its connection with two postsynaptic cells.

Glial cells provide support, nutrition, and protection to neurons. Among them, oligodendrocytes play a significant structural role by enveloping multiple axons around their membrane. It may also provide multiple layers of fat tissue, called myelin sheath, that accelerates the transmission of the electrical activity by isolating the axon.

Back to the anatomic organization (figure 1.3), the white matter is mainly composed of myelinated axons. These axons connect the deepest layer of the cortex with other brain structures (box in figure 1.3), thus forming a macroscopic network between brain structures. This network allows more complex brain tasks to be performed.

The cortical grey matter comprises six layers (figure 1.5) of neurons linked together with unmyelinated axons. As depicted by figures 1.5A and B, the microscopic network formed by those neurons is dense, and the number of connections for one neuron is of the order of a thousand. In layer III of the cortex (figure 1.5B), pyramidal neurons have an orientation perpendicular to the cortical surface. Due to their spatial organization, it is commonly admitted that the electrical activity recorded at the scalp's surface originates mainly from these pyramidal cells (see in section 1.3). The following section explains the mechanism allowing the creation of this electrical activity.

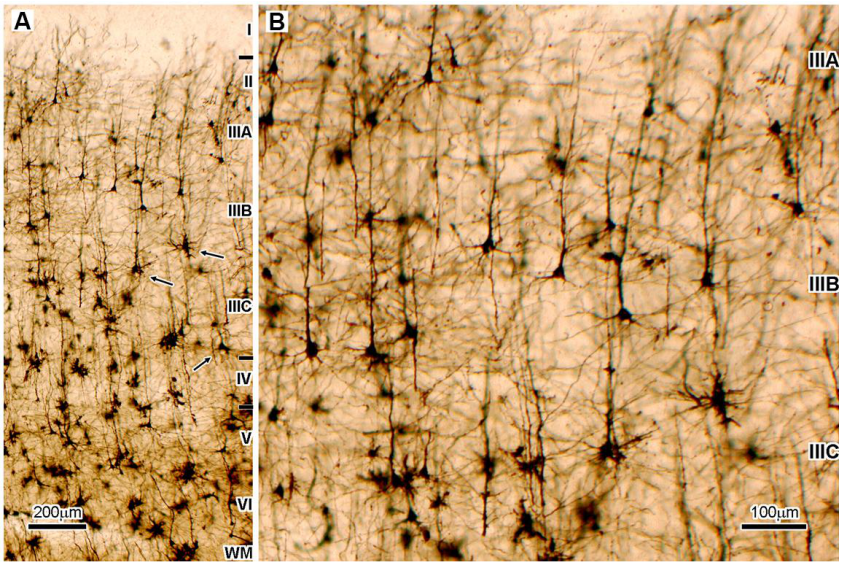


Figure 1.5: Microphotography of the Golgi Cox impregnated neuron layers of the associative cortex region in the human cortex[2]. Figure A represents the six layers of the cortex. Figure B zooms in on layer III to highlight the pyramidal cells. It should be noted that pyramidal cells are also represented in other layers (V and VI).

1.1.3 Neurons electrophysiology

Neurons are excitable cells which means they produce an electrical response to stimulation. This responsive behavior is due to a natural electric potential difference formed between both sides of the cytoplasmic membrane caused by a difference in the concentration of positive and negative ions [17]. In short, these concentrations will cause the extracellular surface of the neuron to be positively polarized instead of the intracellular surface. These differences in concentration and polarity on either side of the membrane will have the effect of creating an electrochemical gradient: the ions will naturally seek to diffuse towards the less concentrated or the inversely polarized milieu. However, ions cannot pass directly through the cytoplasmic membrane and must use a voltage-gated ion channel that is closed during the neuron resting state. If the basis of the axon is sufficiently depolarized (i.e., if the stimulation is sufficient), the channels open. The resulting ion flux will create a typical electrical response which, when reaching a sufficient threshold, can trigger an action potential. This response will be sent along the axon by means of a chain reaction on its way to other neurons. The electrical impulse ends its course in the axon terminal and provokes the release of unique molecules, the neurotransmitters, in the synapse cleft (see section 1.1.2). Neurotransmitters will interact with proteins of the dendrites of neighboring neurons (called postsynaptic neurons), opening ion channels and changing their inner electric potentials. In other words, neurotransmitters will act as stimuli for the postsynaptic cells. If the neurotransmitters are excitatory, the postsynaptic potential increases, thus promoting the continuation of the stimulation. If the neuro-

2. Free to share and use licensed pictures published in [18]

transmitters are inhibitory, the postsynaptic potential decreases. Since a postsynaptic neuron possesses multiple dendrites, all postsynaptic potential will be summed up. The results will determine if the fore all stimulus-response is passed along, stopped or even amplified.

The ionic movement produced by this phenomenon generates current flows inside and outside of each postsynaptic cell. On the inside, this current is called the primary or intracellular current. It reflects the direct activity of the neuron. If the current goes towards the interior of the cell (depolarization by positive ions entry = excitatory stimulus), the extracellular milieu at the synapse becomes more negative and forms a current sink. Conversely, if it is inhibitory, we have a current source. This can be modeled by a current dipole, which is oriented along the cell's dendrite and constitutes the basis for mathematical models of brain activity. On the outside, the extracellular (or secondary) current will generate a local field potential [19]. This electric field is additive, meaning the local field potential from neighboring neurons will amplify it or, on the contrary, decrease it. In the case of synchronous activity of a large population of neurons (≈ 250000 [20]), the resulting local field potential will be strong enough to propagate beyond the tissues to reach the scalp. It is the activity that will be recorded by the scalp EEG. In addition, we can add that the primary currents generate local magnetic fields that will be recorded by the MEG. These two modalities can provide information on brain functions [21] and dysfunctions [22, 23] in the form of characteristic signals. These characteristic markers in the case of epilepsy are presented in the next section.

1.2 Epilepsy

Epilepsy is one of the most common neuronal diseases and concerns about 50 million people worldwide, according to the World Health Organization. It is a heterogeneous and multifactorial disease characterized by the sudden occurrence of a hypersynchronized electrical activity called paroxysmal discharge in one or more cortical regions. This discharge provokes a large variety of symptoms [24, 25], such as motor symptoms, loss of consciousness, and sensations (aura), depending on regions involved by this abnormal activity. However, it always leads to temporary dysfunctions of the electrical brain activity called seizures.

Each type of epileptic paroxysms forms a specific local field potential that will be transcribed into a characteristic grapho-element [23] on the EEG recording (see figure 1.6). Those grapho-elements are used to refine the diagnosis of the patient and used to locate the pathogenic zone. For example, ictal discharges (figure 1.6A) are grapho-elements that can be seen during a seizure. They last several seconds to a few minutes and are characterized by a rhythmic activity. Between seizures, brief paroxysms, called interictal spikes (figure 1.6B), occur in irregular intervals. They

are polymorphous and are represented by spikes or spikes-wave complex. This thesis will focus on interictal spikes as they are a marker of epilepsy between seizures and can be accurately modeled with the neural mass models (see section 1.4.2).

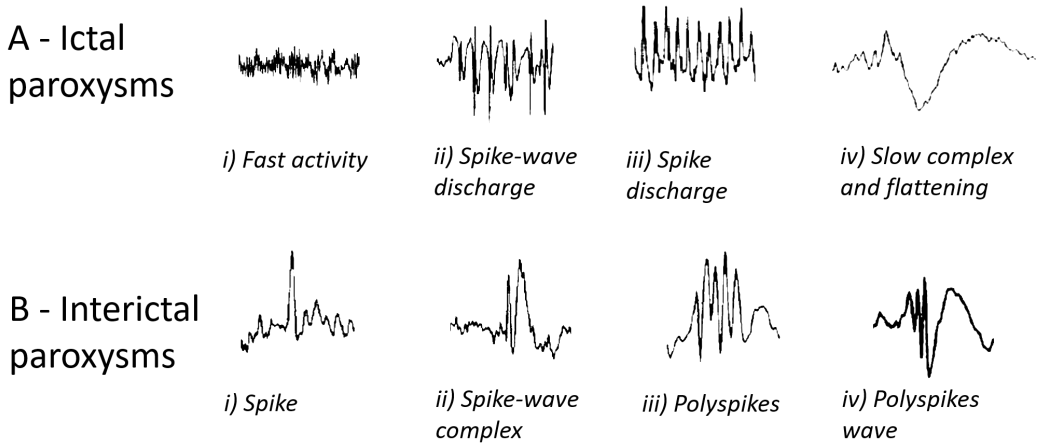


Figure 1.6: Example of characteristic grapho-elements in scalp EEG [3]. A: Examples of interictal paroxysms. B: Examples of ictal paroxysms.

In most cases, epileptic patients can be successfully treated with drugs, which prevents epileptic seizures or allows to reduce their frequency. However, of the 500.000 patients treated in France, 20-30% are drug-resistant, meaning they show no sign of improvement after medication [28]. These cases of drug-resistant epilepsy are among the most severe and often originate from one brain region or from a delimited network of connected regions called the Epileptogenic Zone (EZ). These brain areas can be removed without leading to critical functional deficiencies, and surgical intervention can be considered to stop the occurrence of seizures. In that case, it is clear that the precise delimitation of the EZ is crucial. To do so, patients must usually undergo extensive pre-surgical evaluation, including long-term EEG sessions with video monitoring as well as intracranial SEEG recordings.

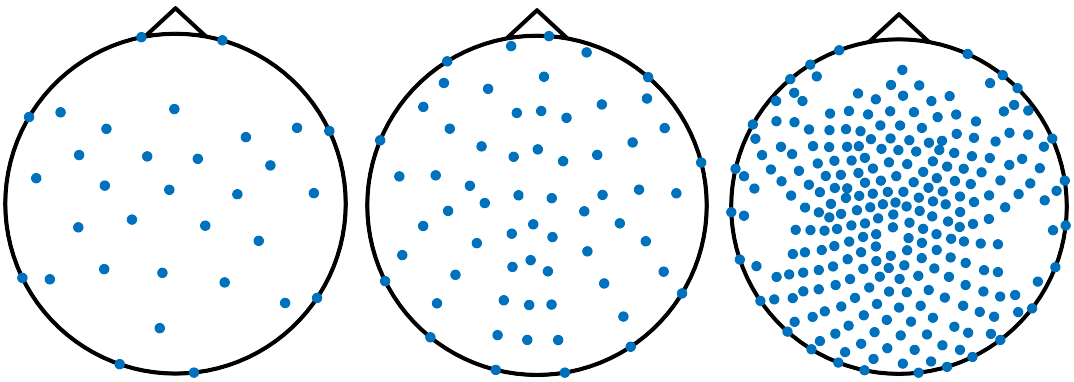
1.3 EEG applied to epilepsy

EEG is a non-invasive technique that measures the scalp's electric potential using a set of electrodes placed on the scalp. More precisely, it records the difference in the electric potential between each sensor and a reference electrode. It is the preferred modality for diagnosing epilepsy [28] as it is cheaper and less restrictive than other modalities like MEG or functional Magnetic Resonance Imaging (fMRI). Accordingly, it allows a more extended time recording section of the brain activity and in a different condition such as recording during patient sleep or a particular activity.

EEG headset used for standard clinical procedure comprises 32 electrodes, whereas

3. Pictures adapted from [26, 27]

dense EEG (also called high resolution-EEG (hr-EEG)) used for research can include up to 256 sensors. The electrodes are positioned on the scalp according to standardized placement systems since the configuration and the number of electrodes substantially impact various applications based on EEG (see [29] for an extensive review).



(a) 32 EEG headset (b) 64 EEG headset (c) 256 EEG headset

Figure 1.7: Example of EEG headsets.

EEG is directly related to the electrophysiological brain mechanisms (see section 1.1.2). Therefore, it offers an excellent temporal resolution of the order of milliseconds, which permits the analysis of the epileptic paroxysms (see the previous section). It is possible to model it physically and mathematically, which allows the development of methods to locate the epileptic zone.

1.4 Modeling epilepsy

1.4.1 Modeling the EEG

As described in section 1.1.3, the current flow generated by the active neurons can be modeled as a current dipole oriented from the dendrites to the cell body. From the scalp perspective, the activity of one neuron alone is unnoticeable as the current is not sufficient enough to go through all the brain tissues. In the particular case of pyramidal neurons (section 1.1.2), this dipole is oriented perpendicularly to the cortical surface. Moreover, since a population of pyramidal neurons forms a dense network of parallel-oriented cells, their activities will sum up to form an equivalent dipole which activity can become recordable at the level of EEG sensors.

A brain source is characterized by the number of current dipoles composing it, its amplitude, orientation, and position. Based on the assumptions we made on these parameters, we define a source model. Two source models (figure 1.8) are often found in the literature: i) the equivalent current dipole considers a point source with a free position inside the brain volume ii) the dipole layer considers a more extended source

with complex geometry confined on a surface layer. The former (figure 1.8a) is less prone to segmentation errors and can allow us to use more accurate diffusion priors but at the cost of a longer computation time. The latter (figure 1.8b) allows to consider the physiological prior formulated above and reduce the search space, which is error-prone when it is oversimplified (see section 1.4.3). Still, it offers a more physiologically plausible and time-saving solution than the former and, therefore, will be used in this thesis. Complementary assumptions will be made on the source model depending on the head model conception in section 1.4.3 and the method used to estimate source activities in section 1.5.2.

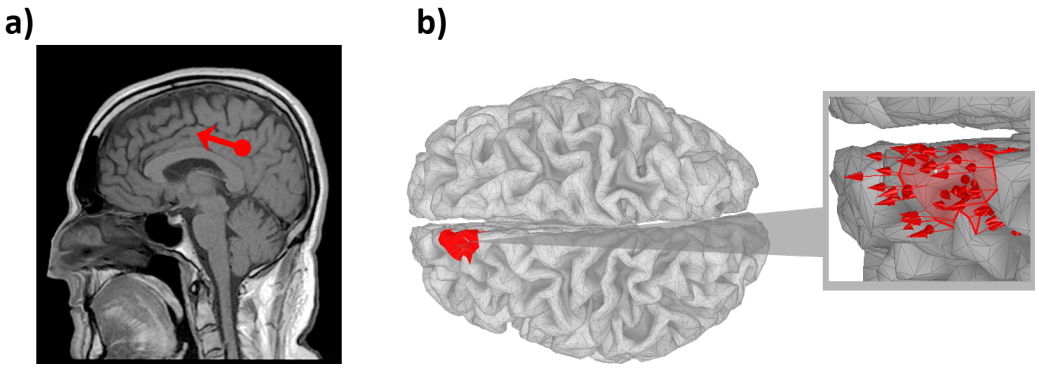


Figure 1.8: Representation of the two source models: a) Equivalent current dipole b) Dipole layer.

From a mathematical viewpoint, it is well established [30] that the EEG signals recorded at the head surface constitute a linear combination of the electric brain sources. Indeed, the EEG signals at time t are assumed to be a realization of an N-dimensional Second-Order wide-sense Stationary Ergodic (SOSE) random process $\{\boldsymbol{x}(t)\}$ given by:

$$\boldsymbol{x}(t) = \boldsymbol{G}\boldsymbol{s}(t) + \boldsymbol{\nu}(t) \tag{1.1}$$

where $\{\boldsymbol{s}(t)\}$ is an L-dimensional SOSE source random process and where $\{\boldsymbol{\nu}(t)\}$ is a SOSE noise random process. $\boldsymbol{G}$ is the so-called leadfield matrix, which models the electrical propagation of the sources to the scalp sensors (see section 1.4.3). The equation (1.1) will be used for simulations where all parameters are known and for real applications where source activity and noise are unknown. Variations of this formula are possible. They will be specified in the manuscript if used.

1.4.2 Modeling the epileptic electrical activity

As exposed in section 1.1.2, the neurons are organized in microscopic networks in the cortex. In a healthy brain, the activity of pyramidal neurons is finely regulated by GABAergic inhibitory interneurons [31]. In epilepsy, this system is disrupted, producing abnormal states that lead to the observed interictal spikes. The situation may gradually evolve into a general unstable system responsible for the seizures. These

networks can be studied by modeling approaches in order to understand how physiological or pathological (i.e. epileptic in our case) signals are generated [32]. Such modeling approaches have been developed in our team for several years. It is not the topic of this thesis to describe them in detail. In short, in macroscopic neural mass models, a population of neurons can be modeled as an electrical system involving transfer function between the three neurons type (pyramidal, slow Gamma-AminoButyric Acid (GABA) and fast GABA) [33, 32]. This system can then be mathematically expressed as a set of ordinary differential equations that need to be solved to create the physiologically plausible signal. This model involves fine-tuning parameters that generate either epileptic or normal brain activities. The latter are also called background activities, as they are not of interest in this thesis.

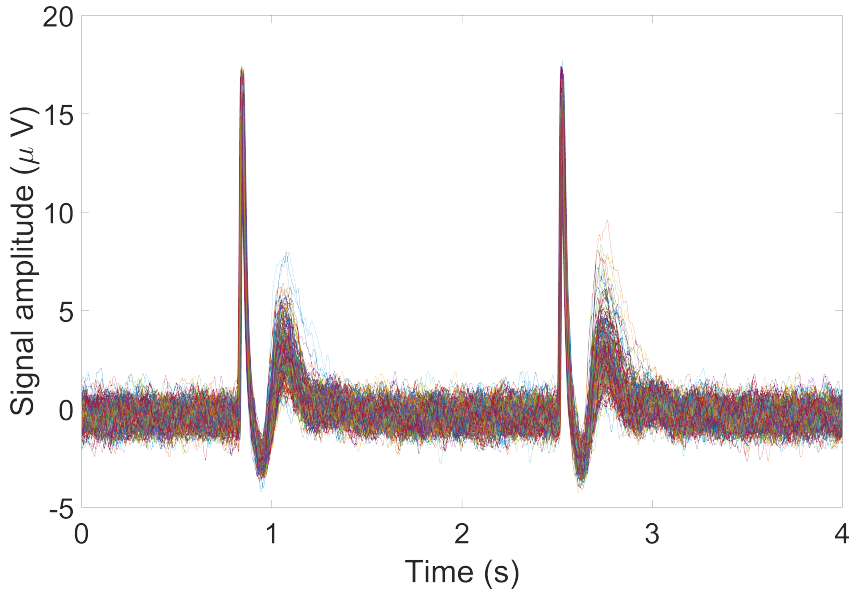


Figure 1.9: Example of 200 physiologically plausible interictal spike signals generated by our team and used in this thesis.

This model can be combined with a model of cortical geometry based on a biophysical model of dipolar sources [34] to simulate signals at the level of extended brain sources most appropriate source model for epilepsy (see section 1.5.2).

1.4.3 Source space, head model and leadfield matrix

As described in section 1.4.1, the source model chosen is a dipole layer. This source model constrains the dipole location to a source space defined by the cortical grey matter. Their orientation is generally constrained to the normal of the cortical surface, but it can remain free depending on the inverse problem methods (see section 2.1).

To accurately model the EEG, the geometry and the physical parameters (conductivity) that impact the diffusion of electric fields in the head (head model) must be considered. The leadfield matrix $\boldsymbol{G}$ carries this information (see equation 1.1) that

can be calculated by solving the forward problem. The simplest head model consists of several nested spheres with homogeneous conductivity representing the brain, the skull, and the scalp. Other surfaces like the cerebrospinal fluid or the spongious part of the skull can also be added to refine the model [35]. In MEG, this model is generally sufficient. However, it is considered an oversimplified model for EEG that is firmly dependent on conductivity variations. To obtain accurate source imaging results in the context of epilepsy (see the following section), a realistic head model should be derived by segmenting anatomical images (T1-Magnetic Resonance Imaging (MRI)). Moreover, the segmentation allows the categorization of the different brain tissues (the grey/white matter (section 1.1.2)) and the scalp. The remaining surfaces are estimated using other algorithms which we will not develop in detail here. In practice, we used in our studies the Freesurfer package [36, 37, 38].

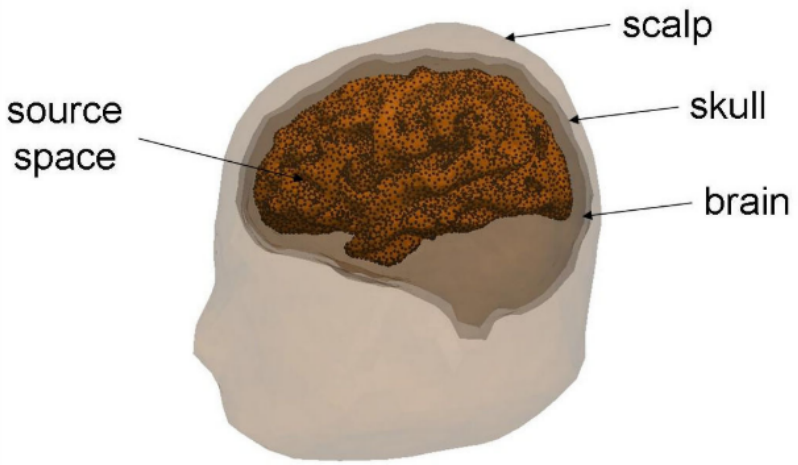


Figure 1.10: Illustration of a realistic head model with three compartments representing the brain, the skull, and the scalp and a source space that consists of a large number of dipoles (represented by black dots) [4].

The segmented brain tissues discretized the source space as triangles and vertices forming a more or less precise cortical surface mesh. Regarding the positioning, the dipole can be fixed at the mesh's vertices or at the centroid of the mesh triangles to complete the source model. The former contains fewer elements and leads to similar results with less computational effort than the latter. Therefore, dipoles are considered on the vertices in this thesis using software like Brainstorm [40]. From a strict anatomical point of view, this layer can represent the gray-white interface (called white) or an intermediate layer between this one and the surface of the cortex (called mid). One must be cautious of mesh errors attributable to the MRI or the segmentation algorithms like holes in the mesh or uneven representation of a brain region. Every modification on the discretized source space can lead us away from the proper

4. Picture from [39]

anatomy of the patient and, therefore, away from the true solution. There is, therefore, a trade-off between precision and condition that requires a consensus between the different manipulators.

Once the head model and the source space are defined, the lead field matrix ($\boldsymbol{G}$ in 1.1) can be computed numerically by solving quasi-static approximation of Maxwell equations [41]. This step is called the forward problem. Assuming no current flow outside the model, they mathematically link the source activities to potential recorded scalp electrodes. The method used to solve those equations depends on the conductivities considered, the head geometry and the source activity. In our case, the Boundary Element Methods (BEM) is used since we have surfaces with homogeneous and isotropic conductivities. Compared to the others, it is considered fast because it involves linear systems that can be easily reversed. However, it has two flaws that should be considered: i) in the presence of a significant difference in conductivity between two surfaces, it is subject to numerical errors ii) when two surfaces are too close, its accuracy drops drastically [41]. These problems can be circumvented depending on the implementation as in the OpenMEEG software [42] used in this thesis. However, it is essential to mention the importance of checking the surface before solving the inverse problem.

1.5 Solving the inverse problem

The source model, the leadfield matrix and the EEG recordings are the only information available for the localization of epileptic sources in a patient. In order to find the epileptic zone, the source activities must be estimated from (1.1) using a brain source imaging method.

1.5.1 Source activities reconstruction

BSI-EEG (Brain Source Imaging based on EEG) aims at reconstructing the electrical activity everywhere within the brain from EEG recordings. The location and the time courses of these electrical brain sources provide information which is essential i) to understand physiopathologies at different scales, ii) to provide a valuable aid for the diagnosis of some diseases and iii) to understand the brain (dys)functions [21, 22, 23] or interactions with our environment [43, 44].

In order to localize active sources and reconstruct their activities, a linear system of equations needs to be solved. This is achieved by solving the so-called EEG inverse problem. This problem is challenging because of its severe ill-posedness due to the attenuation of electrical activity when propagating from the brain to the head surface, and the small number of observations compared to the number of sources to be reconstructed. Thus the solution to such a problem is not unique, and constraints need

to be set in order to identify the appropriate solution among the multiple possible candidates.

1.5.2 State of the art of the inverse problem

Various BSI-EEG methods have been developed [45]. Based on the source model, they can be distinguished into two categories: parametric and non-parametric methods. The parametric approaches, also mentioned as equivalent source dipole methods, consider that few isolated dipoles represent the brain current sources with unknown positions and moments. For non-parametric approaches, also referred to as distributed sources methods, each brain current source is characterized by a large number of dipoles with possibly fixed orientations, which are distributed in the whole brain volume or cortical surface. It is well accepted that the latter is more physiologically plausible for the characterization of epileptic regions.

Regarding the parametric methods, they can broadly be classified into three types: dipole fitting approaches [46, 47, 48], subspace methods [49, 50], and beamforming [51]. The dipole fitting approaches aim at estimating a dipole moment (six non-linear parameters per dipole related to the position, the orientation and the amplitude see section 1.4.1) by minimizing a data-fit cost function of the residual energy. Among the subspace methods, Multiple-Signal Classification (MUSIC) [49] is the most popular. Based on the eigenvalue decomposition of the covariance matrix of measurements, the idea of MUSIC is to divide the vector space of the data into two orthogonal subspaces, the signal subspace and the noise subspace. One can also mention the $2q$-Recursively Applied and Projected MUSIC (RAP-MUSIC) [52] algorithm and its extensions to higher-order statistics [50], where MUSIC is successively applied after removing the contribution of the previously identified sources. The third class considers the beamforming techniques. Basically, the beamforming approaches exploit the data covariance matrix to perform a spatial filtering of the scalp data in order to estimate the temporal activity of the dipoles of interest, called Source(s) Of Interest (SOI(s)), while vanishing the contribution of the other dipoles.

The non-parametric methods can be categorized into three main classes, according to the exploited hypotheses and constraints. The first class is the well-known regularized least-squares method which aims at finding the solution that best describes the measurements in a least-squares sense. This class encompasses i) the Minimum Norm Estimate (MNE) and other more general minimum-norm algorithms such as weighted MNE (wMNE) [53], Low-Resolution Electromagnetic Tomography (LORETA) [54], standardized LORETA (sLORETA) [55], constrained LORETA (cLORETA) [56], Local Auto Regressive Average (LAURA) [57] ii) methods that consider the spatial sparsity of the SOIs [58] or sparsity in a transformed domain [59], where the representation of the sources is sufficiently sparse [45], and iii) techniques that

impose constraints in different domains [60, 61]. For instance $L_{1,2}$-Source Imaging based on Structured Sparsity ($L_{1,2}$-SISSY) [62] simultaneously imposes sparsity in a space transformed domain and smoothness in the time domain. The second class concerns the tensor-based approaches [63] which construct a q-dimensional data tensor ($q > 2$) by exploiting data variability over an additional tensor mode (i.e. frequency, wave-vector, subject or realization), in addition to conventional spatio-temporal methods. The last class covers the Bayesian methods [64, 65, 66], which are based on a probabilistic model of the data by considering the observations, the sources and the noise. More precisely, these techniques attempt to maximize the posterior distribution of the sources for given measurements. Depending on the way used to infer the hyper-parameter, different algorithmic frameworks can be designed such as empirical/variational Bayesian approaches [67]. Note that other distributed source algorithms have been proposed, such as $2q$-Extended Source-Multiple Signal Classification ($2q$-ExSo-MUSIC) [68], which can be considered as an extension of the $2q$-MUSIC method [69] to the case of distributed sources, where $2q$-th order statistics with $q > 1$ of the data are exploited, instead of second-order statistics.

1.6 Conclusion

As argued in the introduction, the choice of a suitable BSI-EEG method remains an open issue and strongly depends on the application context. For instance, this choice can be guided by using the adequate source model, the best trade-off between the complexity/accuracy of the used method and the assumptions made on the statistical relations between SOI(s). However, as stated, it is not uncommon to see parametric approaches being misused to locate distributed sources as they provide unique information. There would therefore be gain in adapting them. The next chapter will focus on the equivalent source dipole model and, more specifically, on beamforming methods [70, 71, 72, 73]. Indeed, under some assumptions, beamforming methods allow us to analytically and efficiently estimate the time course of the dipoles. More discussed in the MEG context than EEG, these parametric methods provide a good initialization to complex distributed source localization algorithms in order to improved confidence in the estimated sources. They can also be exploited to propose a hybrid combination of parametric and non-parametric approaches [74, 75, 76]. However, they present some severe drawbacks as they are known to cancel the electrical activity of correlated sources, which is, to a certain degree, incompatible with the distributed source model. In addition, the behavior of the beamforming techniques is strongly influenced by the preprocessing applied to the data and the metric used to localize SOIs (post-processing step) after filtering. Therefore, we must first carry out a state-of-the-art review of all the practices in beamforming.

Chapter 2

Beamforming for brain source imaging: a comparative performance study using epileptic EEG signals

Our first objective for this thesis is to conduct an extensive analysis of the performance of different beamforming pipelines in the context of EEG, both in terms of SOI localization and reconstruction accuracy. Currently, it can be noticed that all existing studies [77, 78, 79, 80] propose to evaluate either the beamforming filter calculation or the post-processing step, but never to assess the efficiency of the global beamforming pipeline, namely the preprocessing, the filtering (beamforming) and the post-processing. Several experiments are proposed on realistic simulated epileptic EEG data to determine the best beamforming pipeline. That one will be applied to 256-channel hr-EEG interictal spikes acquired from a drug-resistant partial epileptic patient who underwent a full presurgical evaluation.

This chapter is organized as follows: in section 2.1, we formulate the assumptions made throughout this chapter. Section 2.2 proposes a taxonomy of the beamforming pipelines (preprocessing methods, filtering and post-processing metrics). Section 2.3 is devoted to the proposition of adequate pipelines that will be evaluated in our application context. Then, we provide in section 2.4 the results of the conducted study on simulated and on real EEG data. Finally, in section 2.5, a new regularization is proposed to overcome a significant problem of source reconstruction on real data. This regularization, named Temporal-Smoothing-based beamforming, can be applied to all beamforming methods and makes a fine addition that concludes this chapter.

2.1 Assumptions

Lets us recall the mathematical model, the EEG recordings at time t are assumed to be a realization of an N-dimensional SOSE random process $\{\boldsymbol{x}(t)\}$ given by:

$$\boldsymbol{x}(t) = \boldsymbol{G}\boldsymbol{s}(t) + \boldsymbol{\nu}(t) \tag{2.1}$$

where $\{\boldsymbol{s}(t)\}$ is an L-dimensional SOSE source random process and where $\{\boldsymbol{\nu}(t)\}$ is a SOSE noise random process. $\boldsymbol{G}$ is the so-called leadfield matrix, which models the electrical propagation of the sources to the scalp sensors. It is computed from a realistic head model by solving Poisson equation with the BEM [81] (see section 1.4.3). The source orientation is fixed to the outward unit normal vector of the cortical surface. Thus, this thesis deals with the scalar beamforming methods as opposed to vector beamforming algorithms [77]. Moving from one to the other requires only a few changes in the notation; however, the latter adds an extra step of finding the best orientation like in the Synthetic Aperture Magnetometry (SAM) method [82]. This step may be the subject of further discussion [83, 84] which will be out of the scope of this thesis.

Beamforming spatial filters or beamformers can directly be applied to (1.1) to reconstruct each source activity independently:

$$\widehat{s}(\theta_\ell, t) = \boldsymbol{w}(\theta_\ell)^\mathsf{T}\boldsymbol{x}(t) \text{ with } \boldsymbol{w}(\theta_\ell)^\mathsf{T} = f(\widehat{\boldsymbol{C}_{\boldsymbol{x}}}, \boldsymbol{G}) \tag{2.2}$$

where $\boldsymbol{w}(\theta_\ell)$ stands for the band filtering weights for position θ_ℓ. Weight optimization usually leads to an analytical solution depending on the leadfield matrix $\boldsymbol{G}$ and the data Sample Covariance Matrix (SCM) $\widehat{\boldsymbol{C}_{\boldsymbol{x}}}$ defined as:

$$\widehat{\boldsymbol{C}_{\boldsymbol{x}}} = \frac{1}{T-1}\sum_{t=1}^{T}(\boldsymbol{x}(t) - \overline{\boldsymbol{x}})(\boldsymbol{x}(t) - \overline{\boldsymbol{x}})^\mathsf{T} \tag{2.3}$$

Due to the SCM, beamformers are also called linear data-driven estimators as opposed to data-independent estimators [85, 86].

2.2 Beamforming review

Each beamforming method can be decomposed into three steps: preprocessing, filter design and post-processing. These steps contain various techniques that will be carefully described in this section in a broad sense. However, we focus on methods with a solid theoretical foundation.

2.2.1 Preprocessing

All inverse methods start with a preprocessing step during which raw data are transformed to fit some required priors. Preprocessing is sometimes included, like in the array gain beamforming technique [77].

Leadfield normalization

Deep sources have a leadfield vector norm lower than the surfacic ones, which implies a spatial dependency. Normalizing each leadfield vector by its norm avoids localization errors due to low leadfield norms [77].

Spatial filtering

Scalp electrodes recordings present shared activities which can be filtered beforehand. The Surface Laplacian method [87, 88] diminishes shared activities coming from deep sources. A Laplacian matrix is applied to the system in order to focus on the evolution of the signal between the electrodes. Several variants exist based on the way to build the Laplacian. It improves the quality of the beamformer for superficial sources at the expense of deep sources.

Shared activities coming from noises can also be filtered using a whitening filter. Under the SOSE assumption, the noise SCM $\widehat{\boldsymbol{C}_{\boldsymbol{\nu}}}$ of $\{\boldsymbol{\nu}(t)\}$ can be computed from EEG recordings with no signal of interest. A whitening operator is then calculated from the inverse of the square root of $\widehat{\boldsymbol{C}_{\boldsymbol{\nu}}}$. Once applied to (1.1), a new system is obtained with a noise SCM equal to the identity matrix.

Data denoising

Raw recordings include physiological artifacts, which may impact source localization. Blind Source Separation (BSS) methods aim at reformulating data as a linear mixture, from which noise and artifacts can be identified and removed. Independent Component Analysis (ICA) [89, 90], Signal Space Projection (SSP) [91] and Signal Space Separation (SSS) [92] with its extensions [93, 94] have already been used for denoising in beamforming. Note that SSS approaches are restricted to spherical head models for MEG.

Principal Component Analysis (PCA) removes redundant information and improves the localization of deep sources [95, 96]. When applied to the data SCM and combined with meta-heuristic algorithms [97], it tackles the problem of correlated sources in beamforming. When other information like sensor space parcelization is available, Partial Least-Squares (PLS) [96] is used instead. In general, BSS methods improve the results [98] but have a major flaw: imposing a linear mixture model with fewer components than sensors makes the new data SCM be rank deficient (see 2.2.1).

Table 2.1: Preprocessing for beamforming

	Prior	Transformed input	Method and references
Leadfield normalization	Depth dependent direct problem	$\widetilde{\boldsymbol{g}}(\theta_\ell) = \frac{\boldsymbol{g}(\theta_\ell)}{\|\boldsymbol{g}(\theta_\ell)\|_2}$	Array gain beamforming [77]
Spatial filtering	Source at the surface of the cortex	$\widetilde{\boldsymbol{x}}(t) = \boldsymbol{L}\boldsymbol{x}(t)$	Surface Laplacian beamforming [88]
	Whitening	$\widetilde{\boldsymbol{x}}(t) = \widehat{\boldsymbol{C}_\nu}^{-\frac{1}{2}}\boldsymbol{x}(t)$	Noise whitening [102]
Data denoising	Linear mixture	$\widetilde{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{y}(t)$	ICA [90] / SSP [91] / SSS [92, 93, 94] PCA [97, 95] / PLS [96]
Spectral targeting	Narrowband signal	$\boldsymbol{z}(f_0)$	DICS beamforming [99]
	Broadband signal	$\boldsymbol{z}(f)$ for $f \in \Gamma$	FAB beamforming [100]
	Frequency band known	$\widetilde{\boldsymbol{x}}(t, f) = \boldsymbol{h}_f(t) * \boldsymbol{x}(t)$ for $f \in \Gamma$	Filtering [72]
	Time frequency analysis	$\widetilde{\boldsymbol{x}}(\tau, f) = \boldsymbol{h}_f(\tau) * \boldsymbol{x}(\tau)$ for $f \in \Gamma$ and $\tau \in t$	Time frequency beamforming [101]
Robust covariance estimation and reconditioning	Non full-rank covariance matrix	$\widetilde{\boldsymbol{C}_x} = \widehat{\boldsymbol{C}_x} + \mu I_N$	Diagonal loading [103, 77] Bayesian PCA [104]
	Bad covariance matrix estimation	$\widetilde{\boldsymbol{C}_x} = \alpha_1\widehat{\boldsymbol{C}_x} + (1-\alpha_1)\mu I_N$	Oracle Approximating Shrinkage [105] Ledoit shrinkage [100]

with $\boldsymbol{L}$, $\boldsymbol{A}$, $\boldsymbol{h}_f(t)$, $\boldsymbol{z}$ denote the surface Laplacian matrix, the mixing matrix, the impulse response for frequency f and the Fourier transform of $\boldsymbol{x}$, respectively.

Spectral targeting

The SOI may show specific attributes in the frequency domain. Using Fourier transform to enhance the EEG dataset, new spatial filters with a frequency specificity can be designed. In that case, the cross-spectrum matrix, the frequency counterpart of the SCM is used. Once the filters are computed, they are applied to the original EEG time series as in the Dynamic Imaging of Coherent Sources (DICS) beamforming method [99]. However, they are efficient only if the frequency of interest is unique or inside a small set.

In the case of broadband signals in the frequency domain, the filters are applied directly to an averaged cross-spectrum matrix [100]. Another way to deal with broadband signals is to apply multiple Finite Impulse Response (FIR) frequency filters. Data will then become specific to the frequency band given by the impulse response. Moreover, beamforming-based time-frequency analysis is also achievable by using those filters with a time window [101].

Robust covariance estimation and reconditioning

All beamforming methods use the SCM and mainly its inverse. When this matrix is rank deficient or poorly conditioned, computing its inverse is subject to errors. The regularization method of diagonal loading is used to ensure a correct inversion by adding a small value μ to the diagonal components of the SCM, which has to

Table 2.2: Filtering for beamforming

	(Piece-wise) stationnary and ergodic sources				Noises	Criteria
	Location	Known reference	Correlated	Required covariance	Spatially decorrelated	Beamformer type
Unit Gain [77] / SAM [82]	Unknown	no	no	no	no	MV
Unit Noise Gain [103, 77]	Unknown	no	no	no	yes	MV
Standardized MV [85, 79]	Unknown	no	no	no	no	MV
AOLS beamforming [107, 108]	Unknown	no	no	no	no	MV
LCMV [109]	Unknown	no	yes	no	no	MMSE
Common Spatial Pattern [110, 111]	Unknown	no	no	no	no	Max SNR
Spatial Notch [112, 113, 114]	Unknown	yes	no	no	no	MMSE
LMMSE [108]	Unknown	yes	no	no	no	MMSE
Nulling Beamformers [115, 116, 117, 118, 80]	Known	no	yes	no	no	MV
Dual-Source [119] / Dual-Core [120]	Known	no	yes	no	no	MV
MV-PURE [109]	Unknown	no	yes	yes	no	MMSE

be chosen by the user. The Bayesian PCA technique [104] can be used to select an appropriate value of μ.

Another problem arises when there are not enough time samples to compute the SCM, making it a bad estimator of the unknown theoretical covariance matrix. From Random Matrix Theory (RMT), shrinkage methods robustify the SCM estimator using priors. The Oracle Approximating Shrinkage (OAS) [105] and Ledoit's shrinkage [106] have already been used for beamforming applied to BSI-EEG.

Beamforming as a preprocessing

Since beamforming methods reconstruct source time-course activities, they have been used in other inverse problem methods as a preprocessing step. For instance, they are used in meta-heuristic methods as an update step [74, 76] or in minimum norm approaches with the Array-Gain Constraint Minimum Norm with Recursively Updating Gram matrix (AGCMN-RUG) beamformer [75, 79].

2.2.2 Beamforming

Choosing an appropriate beamformer constitutes the core of an inverse problem method based on beamforming. In BSI, three main classes of optimum beamformers [121] have been used:

— Minimum Variance (MV) / Minimum Power (MP)

— Minimum Mean Square Error (MMSE)

— Max Signal-Noise-Ratio (Max SNR)

Minimum Variance / Power Beamforming

The idea is to design a band-pass spatial filter that is as focused as possible on a specific location. Unfortunately, like frequency band-pass filters, it is nearly impossible to have a complete rejection of elements in the stop band. In order to attenuate the contribution from other directions, the output power of the filter must be minimized with respect to directional constraints. Hence, MV/MP beamformers are based on the same optimization:

$$\min_{\boldsymbol{w}(\theta_\ell)} \boldsymbol{w}(\theta_\ell)^\mathsf{T} \widehat{\boldsymbol{C}_\boldsymbol{y}} \boldsymbol{w}(\theta_\ell) \text{ s.t. } \boldsymbol{w}(\theta_\ell) \in \Omega_{\boldsymbol{w}}, \boldsymbol{G} \in \Omega_{\boldsymbol{G}} \tag{2.4}$$

where $\Omega_{\boldsymbol{w}}$ and $\Omega_{\boldsymbol{G}}$ denote the sets of constraints applied to the beamforming filters and the leadfield matrix, respectively. $\boldsymbol{w}(\theta_\ell)$ is the filter computed for position θ_ℓ. The choice of $\boldsymbol{y} \in \{\boldsymbol{\nu}, \boldsymbol{x}\}$ in equation (2.4) determines if a Minimum Variance ($\boldsymbol{\nu}$) or a Minimum Power ($\boldsymbol{x}$) problem is studied. This difference is sometimes only semantic as the optimization with suitable constraints leads to the same filter[80]. If the distinction is clearly mentioned in the signal processing field, the confusion is often made in BSI. In order to remain consistent with the literature, we decide to keep the term minimum variance even if $\boldsymbol{y} = \boldsymbol{x}$. Since the diversity of the MV methods comes from the choice of the constraints, which are discussed below.

Linear constraints Linear constraints have the following common formulation:

$$\Omega_{\boldsymbol{w}_\mathcal{L}} = \{\boldsymbol{w}(\theta_{\ell_1})^\mathsf{T} \boldsymbol{G}(\boldsymbol{\theta}) = \boldsymbol{r}^\mathsf{T}\}, \ \boldsymbol{\theta} = [\theta_{\ell_1}, \theta_{\ell_2}, \ldots, \theta_{\ell_I}] \tag{2.5}$$

where $\boldsymbol{\theta}$ contains the locations of interest, where $\boldsymbol{G}(\boldsymbol{\theta}) = [\boldsymbol{g}(\theta_{\ell_1}), \boldsymbol{g}(\theta_{\ell_2}), \ldots, \boldsymbol{g}(\theta_{\ell_I})]$ and where $\boldsymbol{r}^\mathsf{T} = [1, 0, \ldots, 0]$ is their desired filter response. A natural response for a spatial filter is to let the signal pass from the studied location θ_ℓ by giving it a unitary gain. This linear constrain named distortion-less constraint is the basis of beamforming and can be found alone in the Unit Gain (UG) method [77]. In order to refine the spatial filter and deal with interferences, null responses can be imposed on specific locations that require stronger priors (see appendix A.2).

Quadratic constraints Quadratic constraints are used to normalize or regularize the analytical solution. They are combined to linear constraints as follows:

$$\Omega_{\boldsymbol{w}_\mathcal{Q}} = \{\Omega_{\boldsymbol{w}_\mathcal{L}}\} \cup \{||\boldsymbol{w}(\theta_\ell)^\mathsf{T} \boldsymbol{B}||_2^2 = 1\}, \ \boldsymbol{B} \in \{I_N, \widehat{\boldsymbol{C}_\boldsymbol{x}}^{-\frac{1}{2}}\} \tag{2.6}$$

When $\boldsymbol{B}$ is the identity matrix, the beamformer is qualified by the Unit Noise Gain (UNG) term [77], initially called Borgiotti and Kaplan beamformer [122]. Otherwise, it is the Standardized minimum variance beamformer (Stdr) [85, 79]. Those normalizations are equivalent to dynamic Statistical Parametric Maps (dSPM) [123] and

sLORETA [55], which give a minimum norm inverse solution offsetting the depth bias. Their effect on the solution is still a matter of discussion regarding signal spreading [124].

Shrinkage constraint The shrinkage problem discussed in section 2.2.1 can also be applied to design beamforming filters. It consists of computing the optimal weights between the analytical solution of a UG beamforming filter ($\boldsymbol{w}_1$) and a matched filter ($\boldsymbol{w}_2$) based on a data-independent estimator [125]. The constraint is presented as follows:

$$\Omega_{\boldsymbol{w}_S} = \{\Omega_{\boldsymbol{w}_\mathcal{L}}\} \cup \{\boldsymbol{w}(\theta_\ell) = \alpha_1 \boldsymbol{w}_1(\theta_\ell, \widehat{\boldsymbol{C}_x}) + \alpha_2 \boldsymbol{w}_2(\theta_\ell)\} \tag{2.7}$$

The unknown theoretical covariance matrix will replace the data SCM in the cost function (2.4). The optimization will be done with respect to α_1 and α_2 assuming that this covariance matrix is known. In the end, an asymptotic convergence analysis will enable us to free ourselves from this knowledge. The resulting analytical filter, named Asymptomatically Optimal Linear Shrinkage (AOLS), is better suited to deal with short events as shown in [108]. Note that this method has not yet been introduced in BSI.

MMSE Beamformers

MMSE is defined as follows (see appendix A.3 for full mathematical development):

$$\begin{aligned} MMSE &= \mathbb{E}\left[||\boldsymbol{s}(t) - \hat{\boldsymbol{s}}(t)||_2^2\right] \\ &\overset{(2.2)}{=} \mathbb{E}\left[||\boldsymbol{s}(t) - \boldsymbol{W}^\intercal \boldsymbol{x}(t)||_2^2\right] \end{aligned} \tag{2.8}$$

When the L-dimensional reference signal $\{\boldsymbol{s}(t)\}$ is unknown, the Minimum Mean Square Error (MMSE) cannot be minimized right away. For that reason, a reformulation is required [126]:

$$\begin{aligned} MMSE &\overset{(1.1)+(2.2)}{=} \mathbb{E}\left[||\boldsymbol{s}(t) - \boldsymbol{W}^\intercal(\boldsymbol{G}\boldsymbol{s}(t) + \boldsymbol{\nu}(t))||_2^2\right] \\ &= \mathbb{E}\underbrace{\left[||(I_L - \boldsymbol{W}^\intercal \boldsymbol{G})\boldsymbol{s}(t)||_2^2\right]}_{\text{Bias}^2} + \underbrace{\text{Tr}\{\boldsymbol{W}^\intercal \boldsymbol{C}_\nu \boldsymbol{W}\}}_{\text{Variance}} \end{aligned} \tag{2.9}$$

with two conditions: i) the source random process $\{\boldsymbol{s}(t)\}$ and the noise random process $\{\boldsymbol{\nu}(t)\}$ are uncorrelated and ii) the noise $\{\boldsymbol{s}(t)\}$ is zero mean. From this equation, a new MV-like beamformer can be designed [127]:

$$\min_{\boldsymbol{W}} \quad \text{Tr}\{\boldsymbol{W}^\intercal \widehat{\boldsymbol{C}_\nu} \boldsymbol{W}\} \text{ s.t. } \boldsymbol{W}^\intercal \boldsymbol{G} = I_L \tag{2.10}$$

This optimization problem has multiple names depending on the field of study: Best Linear Unbiased Estimator (BLUE) for linear estimation [126], distortionless-constrained estimator for array processing [128], and Linear Constraints Minimum Variance (LCMV)

in BSI [80]. If the equivalence between the MV and MP problems is accepted, the noise SCM can be replaced with the data SCM. Note that it is a special case of MV beamformer (2.4) for which all the filters are computed simultaneously. A null constraint is applied to all positions except the one of interest making this matrix beamformer the perfect set of pass band filters. However, the underlying minimization problem is ill-posed, involving the inversion of a rank-deficient matrix. As an aside, we would like to point out that there are many uses of the term LCMV in BSI-EEG. In many cases, it refers to the UG with a free source orientation in relation to the introductory paper by Van Veen [71]. This is legitimate as LCMV is a vague term designing the characteristic of the constraint. In this thesis, it is used to define a beamformer that currently has no alternative name to the knowledge of the author.

Max SNR Beamformer

Previous beamforming methods do not take into account both noise and data covariance matrices. In the Brain-Computer Interface (BCI) field, it is common to design a filter that maximizes the Signal to Noise Ratio (SNR) at its output:

$$\underset{w}{\operatorname{argmax}}\,\mathrm{SNR} = \underset{w}{\operatorname{argmax}}\,\frac{\boldsymbol{w}^\top \widehat{\boldsymbol{C}_x}\boldsymbol{w}}{\boldsymbol{w}^\top \widehat{\boldsymbol{C}_\nu}\boldsymbol{w}} \tag{2.11}$$

This optimization problem can be solved using the Common Spatial Pattern (CSP) method [129], which does not involve the leadfield matrix. The absence of the leadfield matrix is unique to BCI [130, 105, 131]. On the one hand, it allows us to avoid physics modeling errors. But on the other hand, it does not ensure to reconstruct brain source activities.

Other beamformers

Beamformers presented above do not require strong priors. This section presents those which need extra priors in order to give better results.

Known reference Regarding (2.8), if a clean reference of the SOI is known, the optimization problem is easier to solve. Spatial notch beamformer [112, 113] is a stop-band filter using a P300 reference signal for cognitive task analysis. This method is also sensitive to correlated sources [113]. If only the source of interest variance is available, there is a more direct method named Linear Minimum Mean Squared Error (LMMSE) [107, 108].

Interferences In the case of interferences correlated with the SOI, the distortion-less constraint of the UG beamformer (see appendix A.2) will involve the cancellation of all of them (interferences and source of interest). This effect can be avoided if a null response filter [116, 132] is applied to every position of source correlated with the

source currently reconstructed. This method, called nulling beamforming, requires the knowledge of the correlated source position which appears as a too strong assumption in practice.

In the case of two correlated sources with known locations (θ_{ℓ_1} and θ_{ℓ_2}), their corresponding leadfield vectors can be merged [119, 120]:

$$\Omega_{\boldsymbol{G}} = \{\boldsymbol{g}(\theta_\ell) = \mathbf{B} \begin{bmatrix} \boldsymbol{g}(\theta_{\ell_1}) \\ \boldsymbol{g}(\theta_{\ell_2}) \end{bmatrix} \}, \mathbf{B} \in \{[\alpha,\ (1-\alpha)]\ , I_2\} \tag{2.12}$$

Those methods are used with the free-oriented source model, proposing a way of estimating the orientation. Otherwise, more generic strategies have been suggested to estimate the correlated source location: partial sensory coverage [133] or short-term energy estimation mapping [134].

Dealing with over-constrained system Adding nulling constraints to (2.5) will make the spatial constraint dimension grow. It will result in an over-constrained optimization problem that needs to be relaxed. Let's introduce the Reduced Rank (RR) beamformers with the following constraint:

$$\Omega_{\boldsymbol{w}_{\mathcal{RR}}} = \{||\boldsymbol{w}(\theta_\ell)^\mathsf{T}\boldsymbol{G}(\boldsymbol{\theta}) - \boldsymbol{r}^\mathsf{T}||_2^2 < \epsilon\} \tag{2.13}$$

This constraint named eigenvector constraint [121] requires to estimate the rank k of $\boldsymbol{G}(\boldsymbol{\theta})$. In order to find an analytical solution, the Truncated Singular Value Decomposition (TSVD) [117] is used:

$$\boldsymbol{G}(\boldsymbol{\theta}) = [\boldsymbol{U}_k^{\boldsymbol{G}} \boldsymbol{U}_{N-k}^{\boldsymbol{G}}] \begin{bmatrix} \Sigma_k^{\boldsymbol{G}} & 0 \\ 0 & \Sigma_{N-k}^{\boldsymbol{G}} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_k^{\boldsymbol{G}^\mathsf{T}} \\ \boldsymbol{V}_{N-k}^{\boldsymbol{G}\ \mathsf{T}} \end{bmatrix} \approx \boldsymbol{U}_k^{\boldsymbol{G}} \Sigma_k^{\boldsymbol{G}} \boldsymbol{V}_k^{\boldsymbol{G}^\mathsf{T}} \tag{2.14}$$

An improvement using subspace suppression has been proposed [118], but it is difficult to apply to real scenarios due to a high computational cost. Minimum Variance - Pseudo Unbiased low-Rank Estimator (MV-PURE) uses an approach similar to the LCMV technique. The constraint will not be limited to the use of the Euclidian norm but more generally it considers the union of unitarily invariant norms (Frobenius, spectral, etc.). The analytical solution of MV-PURE is relatively compact but it involves to estimate the source covariance matrix $\widehat{\boldsymbol{C}_s}$:

$$\begin{aligned} \boldsymbol{W}_{\text{MVPURE}} &= \boldsymbol{U}_k^{\boldsymbol{E}} \boldsymbol{U}_k^{\boldsymbol{E}^\mathsf{T}} \boldsymbol{W}_{\text{LCMV}} \\ \text{with } \boldsymbol{E} &= (\boldsymbol{G}^\mathsf{T} \widehat{\boldsymbol{C}_x}^{-1} \boldsymbol{G})^{-1} - 2\widehat{\boldsymbol{C}_s} \end{aligned} \tag{2.15}$$

The nulling MV-PURE [135] beamformer uses an extra nulling constraint to cancel connected interferences explicitly.

Alternative optimization Regarding the optimization, all the above beamformers are solved using the Lagrangian method in order to find an analytical solution. But beamforming optimization problems have also been rewritten using the Global Sidelobe Canceller (GSC) method [136, 137, 138] or the Frost algorithm [139] which is a way to create real adaptive filters [121]. Note that the iterative Second Order Cone Programming (SOCP) technique was used too [140].

Extended sources Nulling beamformers [117, 118], LCMV and MV-PURE are technically compatible with the case of spatially extended sources (see appendix A.2). But, specific solutions have been proposed either by changing the source model [141] or designing specific beamformers [142, 143]. Beamforming strategies can also be combined with minimum norm methods like in Sparsity And Beamforming for brain source Localisation and Estimation (SABLE) [4] that will be presented in Chapter 3.

2.2.3 Post-processing

Once beamformers have been computed, an estimate of the source activities is obtained using (2.2). Post-processing methods are used to ensure a correct reconstruction of sources and to highlight helpful information on the SOI.

SNR increaser

As discussed in section 2.2.1, the number of time samples selected for the computation of the data SCM is crucial in beamforming. It also influences the SNR of the reconstructed sources [144], which depends on the number of electrodes too [145]. In order to improve the SOI reconstruction, the beamformer can be projected onto the signal (plus interference) subspace. The projector is defined from the eigenvalue decomposition of the SCM leading to a MV-PURE-like solution (2.15). Note that it cannot be applied directly in the case of null constraints since the subspace spanned by leadfield vectors is not necessarily included in the signal subspace of the SCM. New strategies, including a Gram-Schmidt orthogonalization [77] or a patch constraint optimization [80], have been proposed for this aim. Those methods are compatible with the GSC beamformer [146] (see section 2.2.2).

Source analysis

Once the source time activities are reconstructed, a manual review can be done to find the SOI location. When the SOI presents characteristic grapho-elements like interictal spikes, High-Frequency Oscillations (HFO), Event-Related Potential (ERP), traditional signal processing tools can be used to locate them in time and space. Tools based on time-frequency analysis [147] or wavelet analysis [148] can be used. If an assumption-free exploration of the SOI is preferred, the source-space ICA method [90] decomposes the reconstructed sources into independent principal components

Table 2.3: Post-processing for beamforming

	Priors	Form	Name
SNR increaser	SNR degradation by noise	$\widetilde{W} = \boldsymbol{U}_k^{\widehat{C}_x} \boldsymbol{U}_k^{\widehat{C}_x\,\intercal} W$	Eigenspace projection [146, 77]
Source analysis	Linear mixture between SOI and noise	$\widetilde{S} = \boldsymbol{A}\boldsymbol{\Sigma}_k^{\widehat{S}} \boldsymbol{V}_k^{\widehat{S}}$	Source-space ICA [90]
Source mapping	High variance / power for sources	$\sigma(\theta_\ell)^2 = \boldsymbol{w}(\theta_\ell)^\intercal \widehat{C}_x \boldsymbol{w}(\theta_\ell)$ $p(\theta_\ell, f)^2 = \boldsymbol{w}(\theta_\ell)^\intercal \widehat{C}_z(f) \boldsymbol{w}(\theta_\ell)$	Variance [71] pSPM [99]
	High SNR filter output for sources	$SNR(\theta_\ell) = \frac{\boldsymbol{w}(\theta_\ell)^\intercal C \boldsymbol{w}(\theta_\ell)}{\boldsymbol{w}(\theta_\ell)^\intercal \widehat{C}_\nu \boldsymbol{w}(\theta_\ell)}$	output SNR [77]
	Non gaussian SOI	$\kappa(\theta_\ell) = \frac{1}{T}\sum_t^T \frac{(\hat{s}(\theta_\ell,t) - \bar{\hat{s}}(\theta_\ell))^4}{\sigma(\theta_\ell)^4}$	Excess kurtosis [149] SAM(g2) [150]
	Functional coupling between sources	$\gamma(\theta_\ell, \theta_{\text{ref}}) = \frac{\lvert \boldsymbol{w}(\theta_\ell)^\intercal \widehat{C}_z(f) \boldsymbol{w}(\theta_{\text{ref}}) \rvert^2}{p^2(\theta_\ell, f) p^2(\theta_{\text{ref}}, f)}$	Coherence [99]
	The UG analytic solution injected into the ouput SNR	$m(\theta_\ell) = \frac{\boldsymbol{g}(\theta_\ell)^\intercal C_1 \boldsymbol{g}(\theta_\ell)}{\boldsymbol{g}(\theta_\ell)^\intercal C_2 \boldsymbol{g}(\theta_\ell)}$	PZ [151] / NAI [71] ZER (SPA) [152, 153] rZER [78]
Multiple source localizers / Iterative beamformers	Nullify methods	$m_{\text{M}}(\theta_\ell) = \text{Tr}\{\boldsymbol{G}(\boldsymbol{\theta}_\ell^{\text{r}})^\intercal C_1 \boldsymbol{G}(\boldsymbol{\theta}_\ell^{\text{r}}) (\boldsymbol{G}(\boldsymbol{\theta}_\ell^{\text{r}})^\intercal C_2 \boldsymbol{G}(\boldsymbol{\theta}_\ell^{\text{r}}))^{-1}\}$ with $\boldsymbol{\theta}_\ell^{\text{r}} = [\boldsymbol{\theta}_{\text{ref}}, \theta_\ell]$	MPZ / MAI MER (SIA/MIA) rMER [78, 153]
	Deflation method	$m_{\text{RAP}}(\theta_\ell) = \frac{\boldsymbol{g}(\theta_\ell)^\intercal \widetilde{C}_\nu^{-1} \boldsymbol{g}(\theta_\ell)}{\boldsymbol{g}(\theta_\ell)^\intercal \widetilde{C}_x^{-1} \boldsymbol{g}(\theta_\ell)}$ with $\widetilde{C}_x = \boldsymbol{P}(\boldsymbol{\theta}_{\text{ref}}) \widehat{C}_x \boldsymbol{P}(\boldsymbol{\theta}_{\text{ref}})$ $\boldsymbol{P}(\boldsymbol{\theta}_{\text{ref}}) = I_N - \boldsymbol{G}(\boldsymbol{\theta}_{\text{ref}}) \boldsymbol{G}(\boldsymbol{\theta}_{\text{ref}})^\dagger$	Iterative/RAP beamformer [102]

with $\widehat{C}_z(f)$ the cross-spectrum matrix (see section 2.2.1), C the data SCM $\widehat{C}_x$ or the SCM $\widehat{C}_{\bar{x}}$ of the data averaged over several epochs and $\{C_1, C_2\}$ the couple $\{\widehat{C}_x^{-1}, \widehat{C}_x^{-1}\widehat{C}_\nu\widehat{C}_x^{-1}\}$, $\{\widehat{C}_\nu^{-1}, \widehat{C}_x^{-1}\}$ and $\{\widehat{C}_x^{-1}\widehat{C}_{\bar{x}}\widehat{C}_x^{-1}, \widehat{C}_x^{-1}\}$ for PZ, NAI and ZER, respectively. $\boldsymbol{\theta}_{\text{ref}}$ is the location of the previously identified SOI.

that can be reviewed. However, this BSS method cannot be applied in practice due to the high dimension of $\hat{\boldsymbol{s}}$ and alternative strategy must be considered.

Source mapping

As the number of sources is often too high to review each time course by hand, source mapping methods represent a simple topographic map of SOI properties in the source space. Weights generated from these mapping methods will also be used as metrics to localize SOIs mathematically. This section lists all the SOI properties currently used in beamforming.

Variance mapping The first source mapping introduced in EEG/MEG was the estimated source variance. The idea is that each SOI has a superior variance in comparison with that of background activities. The variance has dual metric in the frequency domain called source power Statistic Parametric Map (pSPM).

Output SNR mapping When the noise covariance matrix is known or well estimated, the variance mapping can be enhanced by considering an SNR mapping which decreases the variance of the sources outside the SOIs. Usually the SCM defined in equation (2.3) is used, whereas using averaged data between n epochs has been sug-

gested as a better alternative to process ERP signals.

$$\widehat{\boldsymbol{C}_{\overline{\boldsymbol{x}}}} = \frac{1}{T-1}\sum_{t=1}^{T}(\boldsymbol{x}_{\overline{n}}(t) - \overline{\boldsymbol{x}}_{\overline{n}})(\boldsymbol{x}_{\overline{n}}(t) - \overline{\boldsymbol{x}}_{\overline{n}})^{\mathsf{T}} \tag{2.16}$$

with $\boldsymbol{x}_{\overline{n}}(t) = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{x}_{(n)}(t)$ being the averaged data over N epochs and $\widehat{\boldsymbol{C}_{\overline{\boldsymbol{x}}}}$ the averaged data SCM.

Fourth-order mapping Since brain signals of interest such as interictal spikes are generally non-Gaussian, higher order statistics may be useful. Using the Excess kurtosis, source activity with extreme variations will have a higher weight.

Coherence mapping In the frequency domain, the coherence function is similar to the output SNR. However, it requires to determine a reference beforehand to reduce the computation time drastically. Finding this reference remains an issue and multiple solutions have been suggested [99] like taking for instance the maximum of a minimum norm [99, 80].

SNR-derived mapping Since the solution for $\boldsymbol{w}(\theta_\ell)$ is analytic, it can be inserted into the output SNR formula leading to a new metric. Using the UG solution (see section 2.3.3), the Pseudo-Z (PZ) metric is obtained directly using mathematical simplification. If the equivalence between the MV and MP problems is considered (section 2.2.2), then the PZ metric can be reformulated as the Neural Activity Index (NAI) metric [71]. The remaining metrics do not directly involve the beamformer in both cases and depend on the SCM and the leadfield vectors. Localizer to the Event-Related (ZER) and reduced localizer to the Event-Related (rZER) metrics are defined the same way using the averaged data SCM. The Single-step Peak Approach (SPA) is a search approach dedicated to the ZER metric [153].

Localizers

All the above source mappings can be used to localize one SOI by finding the maximum value of the corresponding metric. But in the case of multiple sources, an iterative (deflation) scheme has to be implemented. Localizers or iterative beamformers will localize multiple SOIs by sequentially deleting the contribution of the sources already localized (also called referenced sources). Although these methods were developed for SNR-derived methods, they may be adapted to all source mapping methods.

Nullify metrics The first method consists in inserting the referenced source leadfield vectors into the derived-SNR metric. Thus, from the PZ, NAI, ZER, and rZER metrics, we can derive the Multiple Pseudo Z (MPZ), Multiple Activity Index (MAI),

Multiple Event-Related (MER), and reduced Multiple Event-Related (rMER) metrics, respectively. The aim is to vanish the leadfield vector highly correlated to the referenced leadfield vector by means of the inversion. The reference location will grow progressively until the energy of new located sources is too weak to be considered as that of SOIs. The Single-step Iterative Approach (SIA) and Multiple-step Iterative Approach (MIA) are search approaches proposed for the MER metric [153]. taking into account the MV-PURE beamformer methodology [109], an improvement has been suggested for the MAI metric using a projector [154].

Deflation metric The second approach consist in removing the contribution of located dipoles to the data covariance matrix using a deflation scheme. An orthogonal projector $\boldsymbol{P}(\theta_{\text{ref}})$ derived from the RAP-MUSIC [52] method is defined using the leadfield vectors corresponding to the already founded location $\boldsymbol{\theta}_{\text{ref}}$.

2.3 Toward the BSI-EEG beamforming pipeline

This part summarizes the mathematical calculations and considerations for these methods from the three beamforming steps. The first part of this section explains which methods were not considered for the pipeline, while the second part gives algorithmic details to solve the inverse problem using beamforming.

2.3.1 Choice of the beamforming pipeline

Depending on the BSI-EEG application, we may have a certain amount of prior knowledge which can be exploited by beamforming methods like the knowledge of the cortical region involved in a specific task or the frequency support of the SOIs. We must discard some methods previously introduced in the review as we aim at offering a versatile study.

For the preprocessing (Table 2.1), the spatial filtering method named surface Laplacian is not considered as it assumes the position of SOI to be shallow. Data denoising methods are also not retained, as they demand to know the number of SOI in advance, which is particularly difficult to obtain in practice. Spectral targeting methods are put aside as they are linked to the signals or the studied cognitive task. Regarding the robust estimation of the data SCM, we decide to keep Ledoit's shrinkage method as it offers an analytical solution that does not imply the arbitrary selection of μ. All other methods have been kept and presented in the following section.

For the beamformers (Table 2.2), we remove methods that do not allow localization such as Max SNR beamformers. Like in preprocessing, methods that require knowing the location of the sources are removed. Likewise, we withdraw those requiring a reference or an accurate estimate of the source covariance matrix.

For the post-processing (Table 2.3), SNR increasers methods and source analysis are removed since they require to know the number of SOIs. However, the coherence metric has been adapted to be applicable in the signal domain called EEG similarity (see section 2.3.4). Finally, the mapping methods using the UG solution have all been removed since they are specific to one beamformer except the deflation scheme, which can technically work for any metric.

2.3.2 Preprocessing

The preprocessing methods kept for the experiment are the leadfield normalization method, the whitening method and Ledoit's shrinkage method. In addition, we have added two combinations that may be interesting to study: the first is called shrinkaged whitening and it considers the conditioning of the noise SCM while the second, named ALL, combined the Leadfield normalization, Ledoit's shrinkage and the shrinkaged whitening. From this point, we considered that the data SCM is computed using all the samples available in order to not advantage shrinkage methods. Mathematically, there are some elements to add to table 2.1 except for the leadfield normalization.

Whitening

For the whitening, the whitened operator is obtained using the Singular Value Decomposition (SVD) decomposition:

$$\widehat{\boldsymbol{C}_\nu}^{-\frac{1}{2}} = \boldsymbol{U}\Sigma^{-\frac{1}{2}}\boldsymbol{U}^\intercal \tag{2.17}$$

The baseline time samples used for the computation of the symmetric noise SCM are selected between spikes from the interictal EEG data.

Ledoit's shrinkage method

Ledoit's shrinkage method [106] aims at providing an estimate of the covariance matrix that is both well-conditioned and more accurate than the data SCM. To do so, it is needed to solve the following optimization problem:

$$\min_{\alpha_1} \quad ||\widetilde{\boldsymbol{C}_x} - \boldsymbol{C}_x||_{\mathsf{F}}^2 \ \text{ s.t. } \ \alpha_1\widehat{\boldsymbol{C}_x} + (1-\alpha_1)\mu I_L \tag{2.18}$$

where $\boldsymbol{C}_x$ is the unknown true covariance matrix. The solution is analytic and uses asymptotic convergence proofs to alleviate the presence of $\boldsymbol{C}_x$. Convergence results define μ as the mean of the diagonal of $\boldsymbol{C}_x$ and the scale parameter ρ_1 as:

$$\alpha_1 = \frac{1}{T^2}\sum_{t=1}^{T}\frac{||\boldsymbol{x}(t)\boldsymbol{x}(t)^\intercal - \widehat{\boldsymbol{C}_x}||_{\mathsf{F}}^2}{||\widehat{\boldsymbol{C}_x} - \mu I_L||_{\mathsf{F}}^2}, \alpha_1 \in [0,1] \tag{2.19}$$

Shrinkaged whitening

Usually, if the whitening operator is well defined, the whitened noise SCM should be equal to the identity matrix. However, like the data SCM, the noise SCM may be ill-conditioned or rank deficient, leading to an operator which will not properly clean the data. This error may have repercussions on beamforming localization and reconstruction results. Therefore, we use a shrinkaged whitening, based on Ledoit's shrinkage method on the noise SCM before computing the whitening operator. A comparison will be made between the traditional whitening and the shrinkaged whitening through the first experiment.

ALL

The ALL preprocessing scenario consist in merging all the previous preprocessing methods, conducted in the following order:

— The noise SCM is regularized by Ledoit's shrinkage method;

— The whitening operator is computed from the regularized noise SCM and applied to both the data and leadfield matrices;

— Ledoit's shrinkage method is applied on the whitened data SCM;

— The leadfield normalization is applied on the whitened leadfield matrix.

2.3.3 Beamformers

The beamformer used through the experiments appear in an opaque font in Table 2.2. This section presents the mathematical resolution of the optimization problem for each selected beamforming method and highlights a few commonalities.

Unit Gain family

Before dealing with the UG beamforming and its normalized version (UNG and Stdr), let us recall how to solve the following optimization problem:

$$\min_{\boldsymbol{w}(\theta_\ell)} \quad \boldsymbol{w}(\theta_\ell)^\intercal \widehat{\boldsymbol{C}_x} \boldsymbol{w}(\theta_\ell) \text{ s.t. } \boldsymbol{w}(\theta_\ell)^\intercal \boldsymbol{g}(\theta_\ell) = r \tag{2.20}$$

From an optimization point of view, this problem can be solved analytically by minimizing the Lagrangian function:

$$\mathcal{L}(\boldsymbol{w}(\theta_\ell), \alpha(\theta_\ell)) = \boldsymbol{w}(\theta_\ell)^\intercal \widehat{\boldsymbol{C}_x} \boldsymbol{w}(\theta_\ell) + \alpha(\theta_\ell)(\boldsymbol{w}(\theta_\ell)^\intercal \boldsymbol{g}(\theta_\ell) - r) \tag{2.21}$$

where $\alpha(\theta_\ell)$ is the Lagrangian multiplier and r is the filter response fixed to 1 in the case of the UG beamformer or relaxed to $\tau(\theta_\ell)$ in the case of the UNG and Stdr beamformers. Vanishing the derivative of the Lagrangian function leads to the following:

$$\boldsymbol{w}(\theta_\ell) = \widehat{\boldsymbol{C}_x}^{-1}\boldsymbol{g}(\theta_\ell)\alpha(\theta_\ell) \tag{2.22}$$

where only the Lagrangian multiplier remains to be found. It happens that we can inject this primary solution directly into the various constraint in order to compute it.

Unit Gain The UG constraint is:

$$\boldsymbol{w}(\theta_\ell)^\intercal \boldsymbol{g}(\theta_\ell) = 1 \tag{2.23}$$

Inserting (2.22) into (2.23) leads to the following value of the Lagrangian multiplier:

$$\alpha(\theta_\ell) = \frac{1}{\boldsymbol{g}(\theta_\ell)^\intercal \widehat{\boldsymbol{C}_x}^{-1}\boldsymbol{g}(\theta_\ell)} \tag{2.24}$$

Unit Noise Gain The UNG constraints are:

$$\boldsymbol{w}(\theta_\ell)^\intercal \boldsymbol{g}(\theta_\ell) = \tau(\theta_\ell) \text{ and } \boldsymbol{w}(\theta_\ell)^\intercal \boldsymbol{w}(\theta_\ell) = 1 \tag{2.25}$$

First, the solution (2.22) is inserted into the first constraint leading to the following solution:

$$\alpha(\theta_\ell) = \frac{\tau(\theta_\ell)}{\boldsymbol{g}(\theta_\ell)^\intercal \widehat{\boldsymbol{C}_x}^{-1}\boldsymbol{g}(\theta_\ell)} \tag{2.26}$$

Second, the value of $\tau(\theta_\ell)$ is computed using the second constraint. Then we obtain:

$$\alpha(\theta_\ell) = \frac{1}{\sqrt{\boldsymbol{g}(\theta_\ell)^\intercal \widehat{\boldsymbol{C}_x}^{-2}\boldsymbol{g}(\theta_\ell)}} \tag{2.27}$$

Standardized The Stdr constraints are:

$$\boldsymbol{w}(\theta_\ell)^\intercal \boldsymbol{g}(\theta_\ell) = \tau(\theta_\ell) \text{ and } \boldsymbol{w}(\theta_\ell)^\intercal \widehat{\boldsymbol{C}_x}\boldsymbol{w}(\theta_\ell) = 1 \tag{2.28}$$

The analytical solution is computed in the same way as for the UNG beamformer:

$$\alpha(\theta_\ell) = \frac{1}{\sqrt{\boldsymbol{g}(\theta_\ell)^\intercal \widehat{\boldsymbol{C}_x}^{-1}\boldsymbol{g}(\theta_\ell)}} \tag{2.29}$$

From (2.24) and (2.26), we can conclude that UG, UNG and Stdr approaches are equivalent within one scalar factor: $\tau(\theta_\ell)$. This factor depends on the location and is used to correct for depth bias, as discussed in paragraph 2.2.2.

AOLS

The AOLS method uses a shrinkage constraint which imposes a trade off between the matched and the UG filters. The corresponding optimization problem is written given by:

$$\begin{aligned} \min_{\alpha_1(\theta_\ell),\alpha_2(\theta_\ell)} \quad & \boldsymbol{w}(\theta_\ell)^\intercal \boldsymbol{C_x} \boldsymbol{w}(\theta_\ell) \text{ s.t. } \boldsymbol{w}(\theta_\ell)^\intercal \boldsymbol{g}(\theta_\ell) = 1 \\ \text{and} \quad & \boldsymbol{w}(\theta_\ell) = \widehat{\boldsymbol{C_x}}^{-1} \boldsymbol{g}(\theta_\ell)\alpha_1(\theta_\ell) + \boldsymbol{g}(\theta_\ell)\alpha_2(\theta_\ell) \end{aligned} \tag{2.30}$$

It can be reformulated as:

$$\min_{\boldsymbol{\alpha}(\theta_\ell)} \quad \boldsymbol{\alpha}(\theta_\ell)^\intercal \boldsymbol{\Gamma}(\theta_\ell)^\intercal \boldsymbol{C_x} \boldsymbol{\Gamma}(\theta_\ell) \boldsymbol{\alpha}(\theta_\ell) \text{ s.t. } \boldsymbol{\alpha}(\theta_\ell)^\intercal \boldsymbol{\Gamma}(\theta_\ell)^\intercal \boldsymbol{g}(\theta_\ell) = 1 \tag{2.31}$$

where $\boldsymbol{\alpha}(\theta_\ell) \triangleq [\alpha_1(\theta_\ell), \alpha_2(\theta_\ell)]^\intercal$ is the vector Lagrangian multiplier and $\boldsymbol{\Gamma}(\theta_\ell) \triangleq [\widehat{\boldsymbol{C_x}}^{-1}\boldsymbol{g}(\theta_\ell), \boldsymbol{g}(\theta_\ell)]$ is the filter basis. The optimization is solved exactly like (2.20) and leads to a familiar result:

$$\boldsymbol{\alpha}(\theta_\ell) = \frac{(\boldsymbol{\Gamma}(\theta_\ell)^\intercal \boldsymbol{C_x} \boldsymbol{\Gamma}(\theta_\ell))^{-1} \boldsymbol{\Gamma}(\theta_\ell)^\intercal \boldsymbol{g}(\theta_\ell)}{\boldsymbol{g}(\theta_\ell)^\intercal \boldsymbol{\Gamma}(\theta_\ell)(\boldsymbol{\Gamma}(\theta_\ell)^\intercal \boldsymbol{C_x} \boldsymbol{\Gamma}(\theta_\ell))^{-1} \boldsymbol{\Gamma}(\theta_\ell)^\intercal \boldsymbol{g}(\theta_\ell)} \tag{2.32}$$

Using the matrix inversion definition and developing the partition matrix $\boldsymbol{\Gamma}(\theta_\ell)$ leads to a first analytical result. However, it contains both the theoretical covariance matrix and the estimated one. By means of the random matrix theory and asymptotically convergence studies [108], equivalences are only expressed in terms of SCM and a ratio $c \in [0, 1]$ equals to the ratio of the number of electrodes over the number of time samples used to estimate the SCM:

$$\boldsymbol{\alpha}(\theta_\ell) = \begin{bmatrix} \alpha_1(\theta_\ell) \\ \alpha_2(\theta_\ell) \end{bmatrix} = \frac{\begin{bmatrix} (1-c)(a_\ell^+ a_\ell^- (1-c) - 1) \\ c a_\ell^- \end{bmatrix}}{a_\ell^- (a_\ell^+ a_\ell^- (1-c)^2 - 2(1-c) + 1)} \tag{2.33}$$

where $a_\ell^- = \boldsymbol{g}(\theta_\ell)^\intercal \widehat{\boldsymbol{C}}_{\boldsymbol{x}}^{-1} \boldsymbol{g}(\theta_\ell)$ and $a_\ell^+ = \boldsymbol{g}(\theta_\ell)^\intercal \widehat{\boldsymbol{C}}_{\boldsymbol{x}} \boldsymbol{g}(\theta_\ell)$.

LCMV

The LCMV presented here can handle correlated sources by means of nulling constraints on every other sources except the computed one:

$$\min_{\boldsymbol{W}} \quad \text{Tr}\{\boldsymbol{W}^\intercal \widehat{\boldsymbol{C}}_{\boldsymbol{x}} \boldsymbol{W}\} \text{ s.t. } \boldsymbol{W}^\intercal \boldsymbol{G} = I_L \tag{2.34}$$

Its solution is analytic, similar to (2.22) and allows us the computation of all filters at the same time:

$$\boldsymbol{W}_{\text{LCMV}} = \widehat{\boldsymbol{C}}_{\boldsymbol{x}}^{-1} \boldsymbol{G} (\boldsymbol{G}^\intercal \widehat{\boldsymbol{C}}_{\boldsymbol{x}}^{-1} \boldsymbol{G})^{-1} \tag{2.35}$$

However, this solution is not directly computable since we cannot guarantee that $\boldsymbol{G}^\intercal \widehat{\boldsymbol{C}_x}^{-1} \boldsymbol{G}$ is invertible. In order to have a good estimation, we use the pseudo inverse based on the singular value decomposition. Singular values close to the computer floating point accuracy ϵ are put to 0:

$$(\boldsymbol{G}^\intercal \widehat{\boldsymbol{C}_x}^{-1} \boldsymbol{G})^\dagger = \boldsymbol{U} \Sigma_{k>\epsilon}^{-1} \boldsymbol{U}^T \tag{2.36}$$

Towards correlated sources

As presented in [71], the cost function of beamforming can be developed to understand the underlying optimization mechanisms. It also highlights the reason why beamforming methods are sensitive to the case of correlated sources. Lets us take two correlated sources at positions θ_{ℓ_1} and θ_{ℓ_2} (section 2.2.2). It gives the following decomposition (see appendix A.2 for full development):

$$\begin{aligned} \boldsymbol{w}(\theta_{\ell_1})^\intercal \widehat{\boldsymbol{C}_x} \boldsymbol{w}(\theta_{\ell_1}) = & ||\boldsymbol{w}(\theta_{\ell_1})^\intercal \boldsymbol{g}(\theta_{\ell_1})||_2^2 \sigma^2(\theta_{\ell_1}) \\ & + \sum_{i=2}^{L} ||\boldsymbol{w}(\theta_{\ell_i})^\intercal \boldsymbol{g}(\theta_{\ell_i})||_2^2 \sigma^2(\theta_{\ell_i}) \\ & + 2\boldsymbol{w}(\theta_{\ell_1})^\intercal \boldsymbol{g}(\theta_{\ell_1}) \sigma(\theta_{\ell_1}) \sigma(\theta_{\ell_2}) \boldsymbol{g}(\theta_{\ell_2})^\intercal \boldsymbol{w}(\theta_{\ell_2}) \end{aligned} \tag{2.37}$$

where $\sigma(\theta_{\ell_1})\sigma(\theta_{\ell_2})$ is the cross-correlation between the two correlated sources. This cross-term is not guaranteed to be positive for $\boldsymbol{w}(\theta_{\ell_1})$. Recall that the beamforming criterion chooses a $\boldsymbol{w}(\theta_{\ell_1})$ to minimize (2.37) in accordance to spatial constraints. In the case of the UG family, the minimum will be obtained by i) convergence to a constant on the first term due to the distortionless constraint, ii) convergence to zero on the second term as it is guaranteed to be positive, iii) assignment of a negative value for the cross-term to compensate for the constant in the first term. If this occurs, then the source estimated variance is significantly less than its true value. This estimation error will lead to reconstruction and localization errors; thus, this method is incompatible with correlated sources. The same observations can be extended to the AOLS beamforming, which embedded the UG analytic results. Regarding the constraint of the LCMV as we defined it, they force the cross-terms between θ_{ℓ_1} and θ_{ℓ_2} to be null. Moreover, it forces every other cross-term to be null such that we do not have to know the position of the correlated sources in advance (see section 2.2.2). If the constraint is well respected, it will correctly estimate the variance of the source at θ_{ℓ_1}, as such, it is technically compatible with correlated sources.

2.3.4 Post-processing

The post-processing methods (table 2.3) kept for the experiments are the variance mapping, the output SNR mapping, the kurtosis mapping and an adapted version of the coherence mapping called EEG similarity mapping.

EEG similarity

It is well-accepted that interictal spikes appear in EEG signals as soon as they spark in the brain. These grapho-elements can be used as a reference to find the location of spike-like sources. EEG similarity is a two-step post-processing method that classifies estimated sources according to their similitude of shape (γ_1) and amplitude (γ_2) with the activity recorded on a reference electrode (n_{ref}):

$$\begin{aligned}\gamma_2(\theta_\ell, n_{\text{ref}}) &= \begin{cases} ||\boldsymbol{x}(n_{\text{ref}}) - g(n_{\text{ref}}, \theta_\ell)\hat{\boldsymbol{s}}(\theta_\ell)||_2^2 & \text{if } |\gamma_1(\theta_\ell, n_{\text{ref}})| > \delta_{\text{thr}} \\ 0 & \text{otherwise} \end{cases} \\ \text{with } \gamma_1(\theta_\ell, n_{\text{ref}}) &= \frac{1}{T-1} \frac{\hat{\boldsymbol{s}}(\theta_\ell) - \bar{\hat{\boldsymbol{s}}}(\theta_\ell)}{\sigma(\theta_\ell)} \left(\frac{\boldsymbol{x}(n_{\text{ref}}) - \bar{\boldsymbol{x}}(n_{\text{ref}})}{\sigma(n_{\text{ref}})} \right)^{\intercal} \end{aligned} \tag{2.38}$$

where δ_{thr} is a threshold set at a certain percentage of the maximum of $\gamma_1(\theta_\ell, n_{\text{ref}})$. In the rest of the chapter, it is fixed to 65 % of the maximum and the reference electrode n_{ref} is selected as the one with the highest activity at interictal spike maximal activity by direct observation.

Source localization

In order to localize one source, we will take the maximum of each metric as per established practice except for the EEG similarity. For this particular criterion, we need to take the maximum of its inverse as we are looking for the source that minimizes the MMSE for this metric. In the case of two sources, we will take both maxima of the variance, EEG similarity and kurtosis metrics. As for the output SNR, we will use the deflation method as discussed in section 2.2.3. However, the orthogonal projector will be applied on the primary system (2.4) instead of using it directly in the metric. Then, the filter is recomputed and will not contain a contribution of the first source.

$$\begin{aligned} m_{\text{RAP}}(\theta_\ell) &= \frac{\widetilde{\boldsymbol{w}}(\theta_\ell)^{\intercal} \widehat{\boldsymbol{C}_{\boldsymbol{x}}} \widetilde{\boldsymbol{w}}(\theta_\ell)}{\widetilde{\boldsymbol{w}}(\theta_\ell)^{\intercal} \widehat{\boldsymbol{C}_{\boldsymbol{\nu}}} \widetilde{\boldsymbol{w}}(\theta_\ell)} \\ \widetilde{\boldsymbol{w}} &= f(\boldsymbol{P}(\hat{\theta}_{SNR})\boldsymbol{x}, \boldsymbol{P}(\hat{\theta}_{SNR})\boldsymbol{G}) \end{aligned} \tag{2.39}$$

Normalization and post-processing

The normalization process modifies the reconstructed sources in order to avoid model artefacts in the case of leadfield normalization (section 2.2.1) and to raise the signal coming from deep sources for UNG and Stdr beamformers(section 2.2.2). In beamforming, its behavior can be directly studied since we have only an analytic solution.

Beamformer normalization In the case of normalized beamformers within the UG family, we concluded in section 2.3.3 that the filter is equivalent up to a factor $\tau(\theta_\ell)$ and, using the linear system in (2.2), the reconstructed sources as well. We aim at

verifying that it is the same for post-processing criteria. Using the UNG gain solution, we have:

$$\begin{aligned}
\sigma_{\text{UNG}}(\theta_\ell)^2 &= \tau(\theta_\ell)^2 \boldsymbol{w}_{\text{UG}}(\theta_\ell)^\intercal \widehat{\boldsymbol{C}_x} \boldsymbol{w}_{\text{UG}}(\theta_\ell) \\
&= \tau(\theta_\ell)^2 \sigma_{\text{UG}}(\theta_\ell)^2 \\
SNR_{\text{UNG}}(\theta_\ell) &= \frac{\tau(\theta_\ell)^2 \boldsymbol{w}_{\text{UG}}(\theta_\ell)^\intercal \widehat{\boldsymbol{C}_x} \boldsymbol{w}_{\text{UG}}(\theta_\ell)}{\tau(\theta_\ell)^2 \boldsymbol{w}_{\text{UG}}(\theta_\ell)^\intercal \widehat{\boldsymbol{C}_\nu} \boldsymbol{w}_{\text{UG}}(\theta_\ell)} \\
&= SNR_{\text{UG}}(\theta_\ell) \\
\kappa_{\text{UNG}}(\theta_\ell) &= \frac{1}{\text{T}} \sum_{t=1}^{\text{T}} \frac{\tau(\theta_\ell)^4 (\hat{s}_{\text{UG}}(\theta_\ell, t) - \bar{\hat{\mathbf{s}}}_{\text{UG}}(\theta_\ell))^4}{\tau(\theta_\ell)^4 \sigma_{\text{UG}}(\theta_\ell)^4} \\
&= \kappa_{\text{UG}}(\theta_\ell)
\end{aligned} \tag{2.40}$$

As we can see, only the Variance will show different results between the UG and the UNG beamforming methods. It can be extended to the Stdr beamformer as well. Regarding the EEG similarity mapping, the first step can be developed like for the Kurtosis metric and is identical for the UG, UNG and Stdr beamformers. As for as the MMSE criterion is concerned, we have:

$$\gamma_{2_{\text{UNG}}}(\theta_\ell, n_{\text{ref}}) = ||\boldsymbol{x}(n_{\text{ref}}) - g(n_{\text{ref}}, \theta_\ell) \tau(\theta_\ell) \hat{\mathbf{s}}_{\text{UG}}(\theta_\ell)||_2^2 \tag{2.41}$$

As we can see, the reconstructed sources and the EEG signals have not the same scale. Thus, they generate an amplitude bias. The EEG similarity metric is not directly applicable to the UNG and Stdr beamformers. The factor must be reabsorbed, which comes down to use the UG beamformer.

Leadfield normalization The leadfield normalization preprocessing method can be developed in a same manner as for the UNG and Stdr beamformers by using the normalized leadfield vector $\tilde{\boldsymbol{g}}(\theta_\ell)$ introduced in table 2.1. From equations (2.22)-(2.24), we derive the following normalized version of the UG filter:

$$\begin{aligned}
\widetilde{\boldsymbol{w}}_{\text{UG}}(\theta_\ell) &= \frac{\widehat{\boldsymbol{C}_x}^{-1} \tilde{\boldsymbol{g}}(\theta_\ell)}{\tilde{\boldsymbol{g}}(\theta_\ell)^\intercal \widehat{\boldsymbol{C}_x}^{-1} \tilde{\boldsymbol{g}}(\theta_\ell)} = \frac{||\boldsymbol{g}(\theta_\ell)||_2^2}{||\boldsymbol{g}(\theta_\ell)||_2} \frac{\widehat{\boldsymbol{C}_x}^{-1} \boldsymbol{g}(\theta_\ell)}{\boldsymbol{g}(\theta_\ell)^\intercal \widehat{\boldsymbol{C}_x}^{-1} \boldsymbol{g}(\theta_\ell)} \\
&= ||\boldsymbol{g}(\theta_\ell)||_2 \boldsymbol{w}_{\text{UG}}(\theta_\ell)
\end{aligned} \tag{2.42}$$

The leadfield normalization leads to the same behavior as the UNG and Stdr beamformers in terms of post-processing metrics: a change will be noticeable for the Variance criterion or a direct mapping of the source activity at a specific time. For this preprocessing, it means an absorption of model errors (see section 2.2.1). These assumptions are also valid for the AOLS and LCMV beamforming methods.

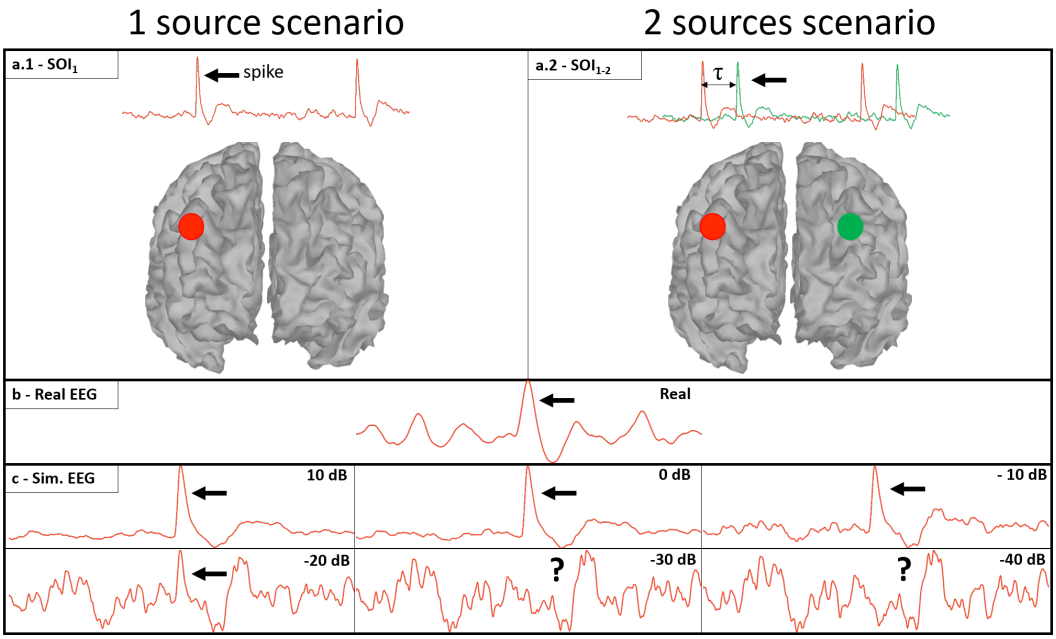


Figure 2.1: Data used in the simulations: a - Time courses and positions of the SOIs used in the one source scenario on the left side (a.1) and two sources scenarios on the right (a.2), b - Example of real interictal spike EEG data, c - Examples of simulated interictal spike EEG data for different input SNR values. When identifiable, interictal spikes are indicated using a black arrow.

As a side note, there is no interest in mixing the normalization preprocessing with the normalized beamforming methods since the latter embedded the former through the constraint relaxation $\tau(\theta_\ell)$.

2.4 Results

Three experiments were considered in this section to quantitatively and qualitatively evaluate the efficiency of beamforming pipeline. The first two experiments were performed on realistic simulated data in the context of focal epilepsy (figure 2.1). The objective of the first one was to assess the preprocessing methods and the post-processing metric performance involved in the beamforming pipeline. The second experiment aims to evaluate the second step of the pipeline, namely the used beamforming algorithm. Both experiments were conducted in two different scenarios: one source vs two sources. The use of the single source will permit us to evaluate the behaviors of the used pipelines as a function of SNR, and the two sources scenario will assess the influence of the correlation between the two sources on the used pipeline. Eventually, the third experiment was performed on human epileptic recordings to qualitatively measure the efficiency of the best beamforming pipeline on real hr-EEG signals.

2.4.1 Datasets and performance criteria

Generation of simulated data

To quantitatively evaluate the performance of the beamforming pipeline, we simulate a 257 channels EEG data (hr-EEG) using a realistic head model, representing the brain, the skull and the scalp, whose intermedium surfaces were extracted from the segmentation of a subject's anatomical 3D T1-weighted MRI. The source space consisted in a mesh made of 15002 vertices obtained from the segmentation of the grey-white matter interface from the same anatomical T1-weighted 3D-MRI using the software BrainVisa [155]. The source space mesh has been refined using the iso2mesh algorithm [156] which ensure that each vertex approximately cover the same amount of brain surface. Each vertex was associated to an elementary current dipole except for the vertices (≈ 1500) inside the Corpus Callosum since they are not of interest [157]. The time course of each source, was provided by the output of a computational neural mass model [33, 31], in which parameters can be adjusted to generate either background-like activity or interictal-spikes. Using this realistic head model, the forward problem was then numerically calculated for each vertex using BEM [42, 40] to obtain the lead field vectors contained in the matrix $\boldsymbol{G}$ (257×15002), which gives the contribution of each dipole of the mesh at the level of 257 scalp electrode positions. More particularly, for each experiment of simulated data, two source configurations were considered to generate epileptiform spike-like activity of 200 times samples with a sampling rate of 512 Hz (see figure 2.1): i) a single equivalent dipole source located in the right frontal lobe, and ii) two equivalent dipole sources located in the right and left frontal regions. For both experiment, the remaining vertices of the mesh were grouped into 1000 sets, called scouts. The dipoles within a scout were associated to neuronal population that generates background activity. The activity between scouts is uncorrelated. All simulations are repeated for 50 Monte Carlo (MC) trials with different spike-like signals and background activities.

Performance criteria

The performance of the used beamformer pipeline is evaluated in terms of source localization and source reconstruction. The localization performance is assessed by calculating: the Euclidean distance between the estimate source position ($\hat{\theta}_i$), and the true position (θ_i) over the 50 realizations:

$$d_{\theta_i}^{Euc} = ||\hat{\theta}_i - \theta_i||_2 \tag{2.43}$$

Regarding the evaluation of the times course reconstruction of the SOI, we compute the following Mean-Squared Error (MSE):

$$MSE = \frac{1}{T}\sum_t \left(\frac{s_e(\theta, t)}{\sigma_{\boldsymbol{s}_e}} - \frac{\widehat{s}_e(\widehat{\theta}, t)}{\sigma_{\hat{\boldsymbol{s}}_e}} \right)^2 \tag{2.44}$$

where $\widehat{s}_e(\widehat{\theta}, t)$ is the estimate time course at the estimate position and $s_e(\theta, t)$ is the time course at the true position. σ_{s_e} and $\sigma_{\hat{s}_e}$ are the standard deviation of the corresponding signals.

Simulations results will be presented in boxplot format in order to study both the accuracy and variability of each preprocessing and beamforming methods.

2.4.2 Simulated data

Experiment 1: Influence of the used preprocessing methods and post-processing metrics

The behaviors of the chosen preprocessing procedures and post-processing metrics are evaluated in terms of source localization and reconstruction as a function of SNR and correlation level. To this aims the beamforming filter is always calculated using the UG filter algorithm. More precisely six cases of preprocessing procedures, namely No-Preprocessing, Normalization, Whitening, Shrinkage, Shrinkage Whitening and ALL, and four post-processing metrics, namely Variance, SNR, Kurtosis and EEG similarity are compared (see sections 2.3.2 and 2.3.3 for details).

Single source scenario The simulated 257 channels EEG data are generated from a single epileptic source by varying the SNR values from -40 dB to 10 dB with a step of 10 dB. Figures 2.2a, 2.2b, 2.2c, 2.2d (left column) represent the Euclidian distances calculated between the true position of the spike-like source and the source position estimated using Variance, SNR, Kurtosis and EEG similarity metrics, respectively. For high SNR values (≤ 0 dB), the localization of the SOI is perfect ($d_{\theta_i}^{Euc} = 0$) whatever the used preprocessing algorithm and post-processing metrics. In the cases of realistic SNR values of -20 dB and -10 dB (see figure 2.1), all preprocessing methods permit an excellent localization when using SNR, Kurtosis and EEG similarity metrics. In contrast, the variance metric performs poorly, except when the ALL preprocessing procedure is applied. For very low SNR values (SNR $\in \{-30, -40\}$ dB), where the spike-like activity is entirely hidden in noisy simulated data as depicted in figure 2.1, we can see that all used pipelines fail to localize the SOI, except for an SNR value of -30 dB where i) the Kurtosis metric succeeds in the localization of the target source for all preprocessing procedures, and ii) the EEG similarity gives a quasi-perfect localization when the Whitening, Shrinkage Whitening and ALL are used.

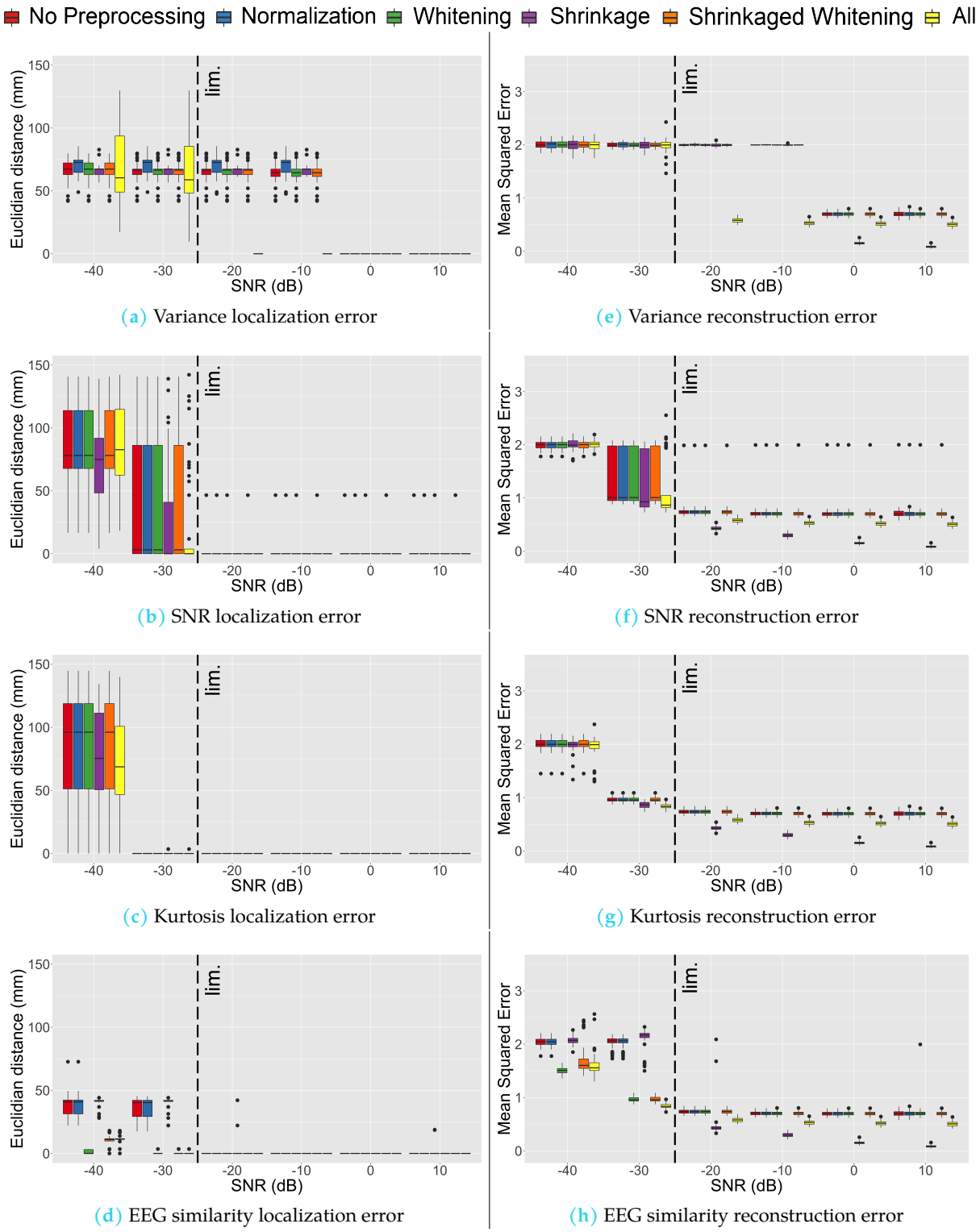
As far as the times course reconstruction quality is concerned, figures 2.2e, 2.2f, 2.2g, 2.2h (right column), the MSE values show clearly that the SNR, Kurtosis and EEG similarity outperform the Variance metrics, for all SNR and all preprocessing methods. We can also conclude that the preprocessing Shrinkage methods seem to be the most effective, followed by the ALL method preprocessing procedure. More finely,

even if the performances of the Kurtosis and the EEG similarity remain reasonable for the SNR of -30 dB, we globally observe that for very low SNR, of -40 dB and -30 dB, the signal reconstruction quality is poorer whatever the used pipeline.

Two sources scenario Here the simulated EEG data are generated using two epileptic sources (see figure 2.1). The aim is to access the influence of the preprocessing and the post-processing on the beamformer BSI-EEG pipeline in the presence of correlated sources. To do so, we fix the SNR value to -10 dB and vary the correlation coefficient between the two SOIs. The left column displays the Euclidian distances between the two true positions of the SOIs and the estimated ones. We observed that, for a correlation coefficient equal 1 (the two SOIs are perfectly correlated), all used pipelines fail in localizing the two sources, except the ones that using the Whitening, Shrinkage Whitening and ALL preprocessing methods and the EEG similarity post-processing metric (figure 2.3d). In fact, in those cases, even if the localization is not perfect, it still reasonable (Euclidian distances around 5 mm). Concerning the other correlation coefficient values (≥ 0.9), figure 2.3a, shows that the Variance metric gives lower performance than other post-processing metrics, except when the ALL preprocessing procedure is applied. This result is in agreement with the one obtained in the single source scenario. In figure 2.3b, we remark that the SNR metric gives perfect results (for correlation coefficient values ≤ 0.9) only when the Shrinkage and the ALL techniques are performed before applying UG algorithm. Indeed, for the four other cases, even if the boxplots present very small medians, the SNR metric fails to localize SOIs in several MC trials.

Regarding the MSE results, in agreement to the single source scenario, figures 2.3e, 2.3f, 2.3g, 2.3h, show that the Variance metric is less effective than other post-processing metrics. In addition, no pipeline is able to properly rebuild the SOIs for a perfect correlation of 1. We can also conclude that in the case of low correlations (≤ 0.5), the Shrinkage and the ALL preprocessing are more efficient that other preprocessing methods and gives interesting performances with MSEs around 0.5.

Discussion: part 1 All previous results conclude that the leadfield Normalization method does not improve performance. Indeed, it presents similar results when no preprocessing is applied, whatever the SNR and the coefficient correlation level. Noise whitening (Whitening), which aims to enhance the SNR (section III.A.2), shows some interest in the -30 dB and sometimes -40 dB scenarios, especially when a data-driven post-processing metric (EEG similarity) is applied. However, it is crucial to be aware that the Noise Whitening preprocessing needs to select a priori a segment of EEG with only background activity (without any SOIs) to estimate the noise covariance (section 2.3.2). In practice, it is very tricky to visually select such a part of EEG signal for bad SNR (see figure 2.1). The regularization method Ledoit is ap-



(a) Variance localization error

(e) Variance reconstruction error

(b) SNR localization error

(f) SNR reconstruction error

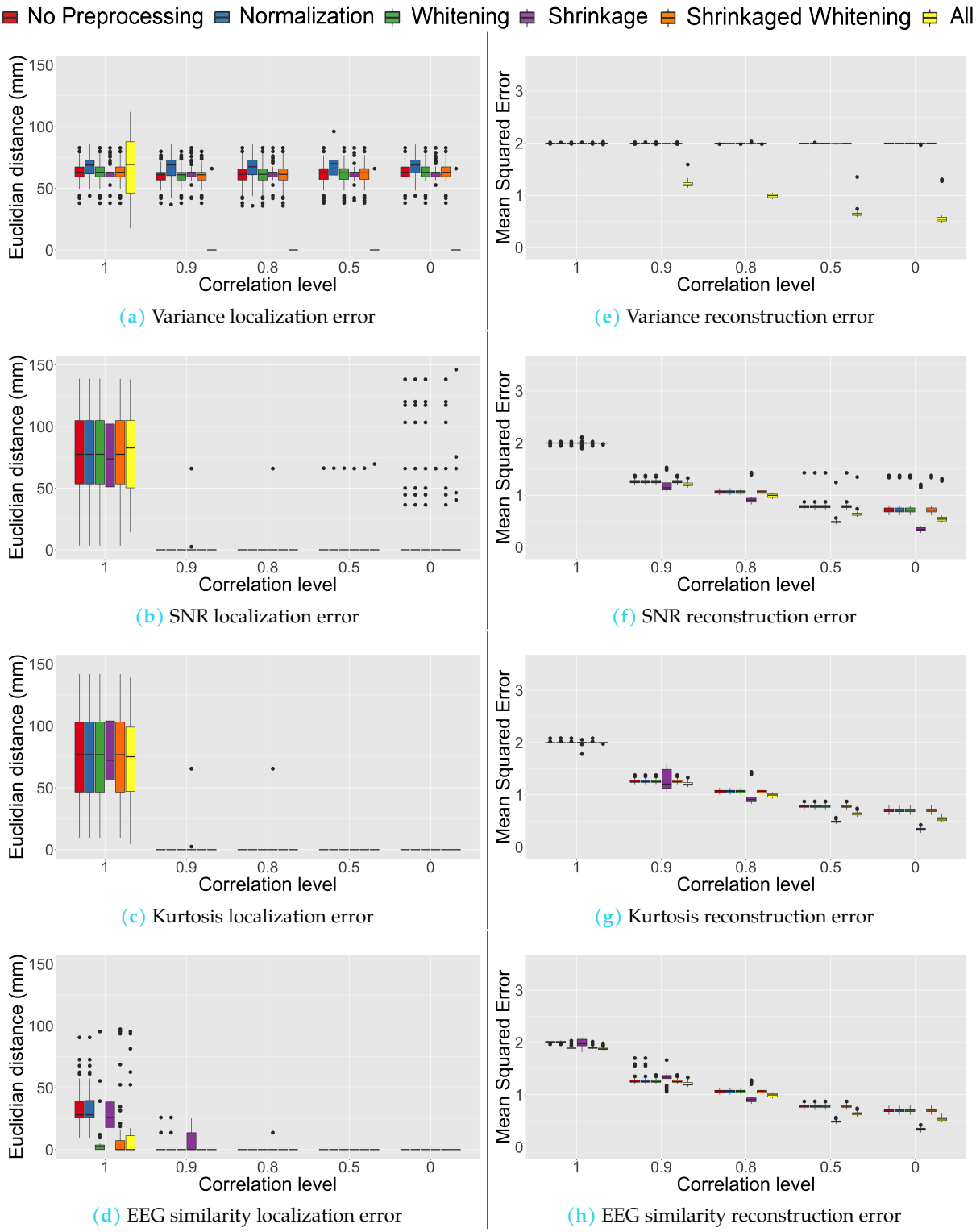
(c) Kurtosis localization error

(g) Kurtosis reconstruction error

(d) EEG similarity localization error

(h) EEG similarity reconstruction error

Figure 2.2: Effect of different preprocessing methods on the unit gain beamformer for one source: localization results are on the left and signal reconstruction on the right. The dashed lines represent the limit of baseline identification.



(a) Variance localization error

(e) Variance reconstruction error

(b) SNR localization error

(f) SNR reconstruction error

(c) Kurtosis localization error

(g) Kurtosis reconstruction error

(d) EEG similarity localization error

(h) EEG similarity reconstruction error

Figure 2.3: Effect of different preprocessing methods on the unit gain beamformer for two sources depending on the correlation level: localization results are on the left and signal reconstruction on the right.

plied to i) the SCM of the data in the case of Shrinkage algorithm, ii) the SCM of the noise before whitening in Shrinkaged Whitening method, and iii) on both SCMs of the data and the noise in the ALL procedure. In our case, the size of SCMs of the data is (257×257) and we numerically verified that using classical estimators leads to not well-conditioned SCMs matrices. Applying Ledoit's shrinkage method improves the conditioning of the SCMs of the data. It gives a more robust estimation of their inverse matrices, explaining the good performance for almost all SNR values and the tested correlations. However, when Ledoit's shrinkage algorithm is applied to the noise SCMs, we hardly notice any benefit. This can be explained by the fact that in the major part of the used pipelines, we do not need to inverse the noise SCMs, and therefore, we lose the interest of such preprocessing. Note that the use of the Shrinkaged Whitening and the ALL methods remains problematic in practice for the same reasons mentioned previously for the Whitening method, i.e. the difficulty in estimation of the noise SCMs.

Now, if we focus on the post-processing metrics, the Variance metric is less effective, especially for low SNR values ($\leq$ -10 dB). This is due to the fact that this metric exploits the SCMs (see table 2.3) of the data and seems to be less robust to their bad estimations. The same remark also remains valid for the SNR metrics, even if the latter appears to be more robust for realistic SNR values of -20 dB and -10 dB. Nevertheless, because of the need of the estimation of the noise variance in the calculation of the SNR metric, as showed in table 2.3, its application in practice is not straightforward. The good performances of the Kurtosis and the EEG similarity can be explained by the fact that: i) in the case of Kurtosis, interictal spikes, considered as SOIs, have higher fourth-order moments than the simulated noise, which is close to Gaussian noise, i.e. with very small fourth order moments, and ii) the EEG similarity differs from other metrics as it explicitly exploits the time course of the surface EEG. We do not need to estimate any data statistics, which makes it less dependent on the sizes of the data matrix (number of electrodes and number of time samples).

In the rest of the thesis, taking into account the results and conclusions obtained in Experiment 1, we decide to select the Shrinkage preprocessing method (Ledoit's shrinkage method on the SCM of the data) for all the following studied pipelines. Indeed, in addition to give very satisfactory results, it is suited to use it in practice. Regarding the choice of the post-processing method, we decide to keep two metrics, namely the Kurtosis and because the EEG similarity. The Variance metric is discarded because of these poor performances. For the SNR metric, even if it gives quite satisfactory results, we do not take it into account because its practical use is not obvious, as stated above. Since the UG, UNG and Stdr beamforming methods are equivalent for the Kurtosis and the EEG similarity metric is incompatible with the UNG and Stdr as demonstrated in 2.3.4, only the UG results will be shown in the next experiment with the denomination UG family.

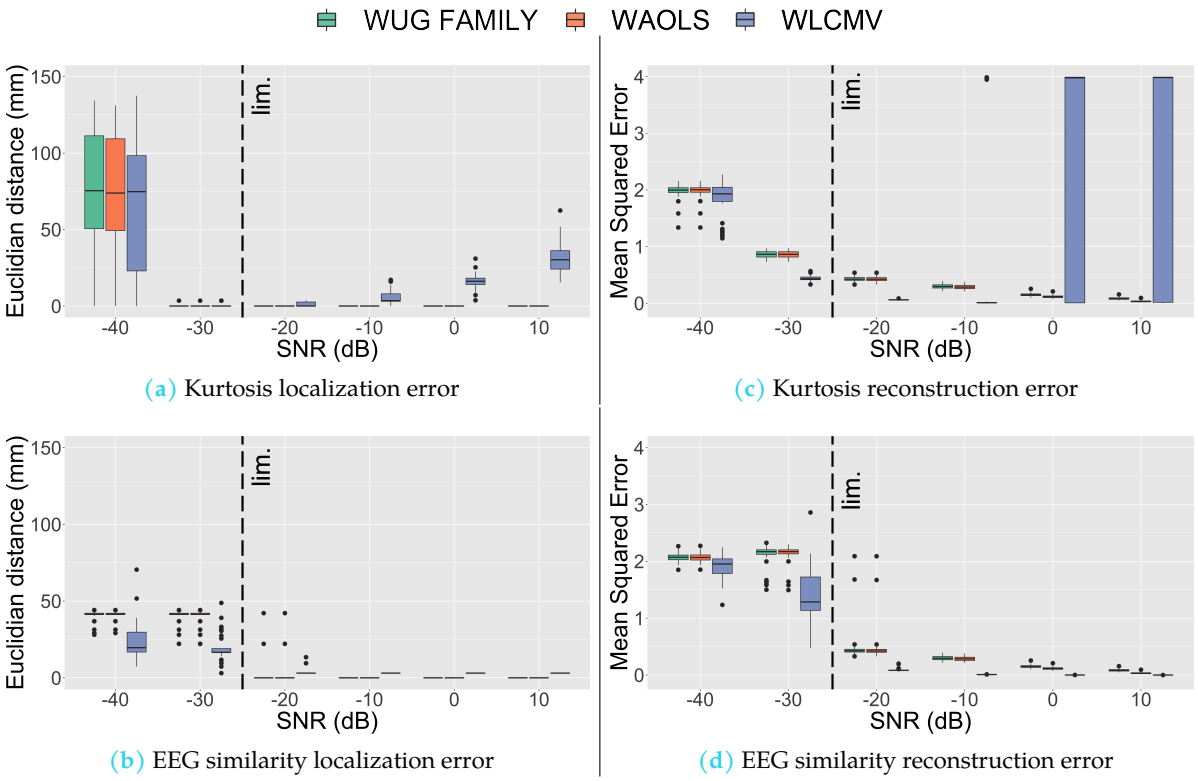
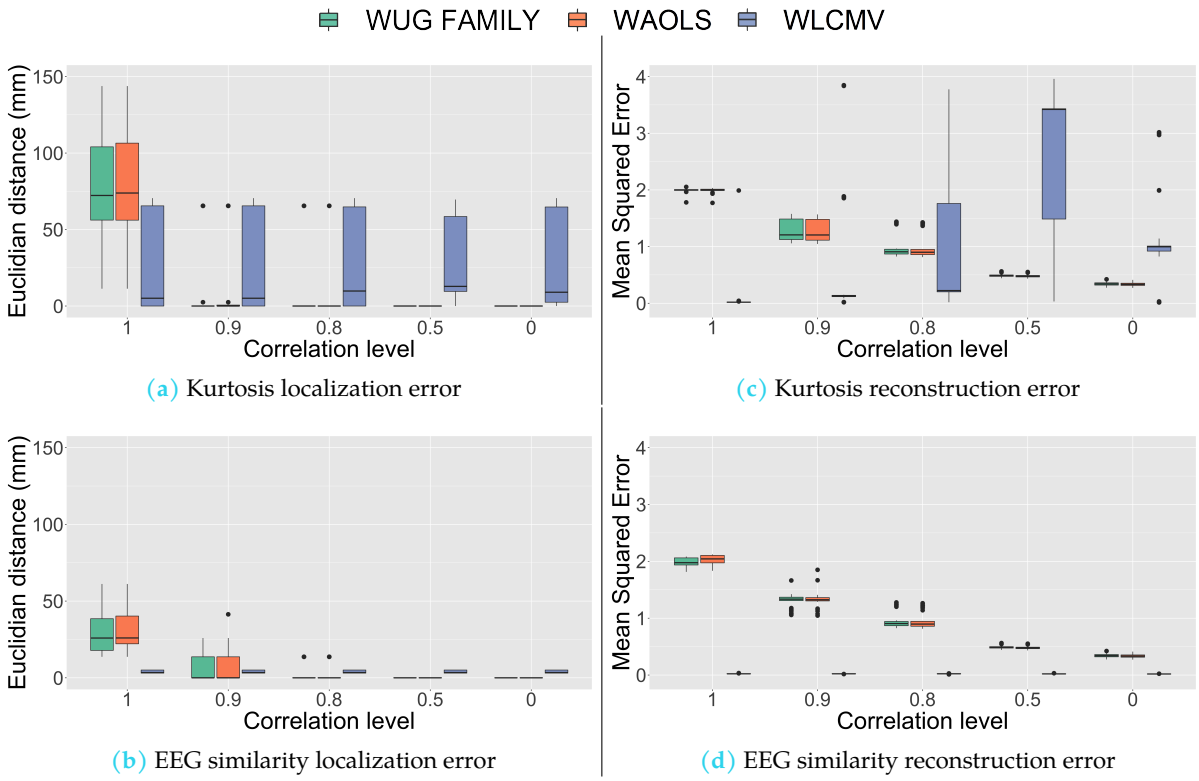


Figure 2.4: Effect of different beamforming methods with the Shrinkage preprocessing method in the single source scenario: localization results are on the left and signal reconstruction on the right. The dashed lines represent the limit of baseline identification.

Experiment 2: Beamforming algorithms efficiency

This second experiment aims to compare beamforming methods with the best preprocessing and post-processing determined in the previous section. As indicated in section (2.3.1), the aim of this chapter is to give a versatile study to evaluate different beamformers pipelines in the context of BSI-EEG. Thus, as justified in sections 2.3.4 and 2.4.2 only the UG family algorithms, AOLS and LCMV (table 2.2), are considered in the following. A similar scheme of experiment 1 is also adopted hereafter: i) single-source scenario to study the influence of the SNR, and ii) two sources scenario to address the influence of the correlation level between the two sources. In addition, as stated in section (2.4.2), only Shrinkage preprocessing and the Kurtosis and the EEG similarity metrics are used in the evaluated pipelines.

Single source scenario Figure 2.4 displays the Euclidian distance and the MSE criterion for the different studied pipelines as a function of SNR. We observe that UG and WAOLS present quasi-identical behaviors whatever the SNR level and the studied pipeline. WAOLS seems to be a little bit more efficient than the UG family to reconstruct the SOIs. For all SNR $\geq$ -20 dB, the performances of these two beamformers, in terms of SOIs localization, are perfect (Euclidean distance equal zero) and the SOIs reconstruction quality is also very promising for both post-processing



(a) Kurtosis localization error

(c) Kurtosis reconstruction error

(b) EEG similarity localization error

(d) EEG similarity reconstruction error

Figure 2.5: Effect of different beamforming methods with the Shrinkage preprocessing method in the two sources scenario: localization results are on the left and signal reconstruction on the right.

metrics. Regarding the LCMV, it seems that for SNR $\geq$ -20 dB, the algorithm is very effective when the EEG-Similarity metric is used. Indeed, the Euclidian distance and the MSE are very small. In the case where the Kurtosis metric is exploited in post-processing step, the LCMV gives very satisfactory results for realistic SNR values (SNR $\in \{-20, -10\}$ dB); but these performances deteriorate for the SNRs ≥ 0 dB. Regarding very low SNR value of -40 dB, clearly all beamformers fail to localize and reconstruct the SOIs. However, at -30 dB, it is interesting to show that all beamformers perfectly localize the SOIs when the Kurtosis metric is used. In addition, their reconstruction quality still also reasonable in this case, especially for LCMV.

Two sources scenario As in experiment 1, the SNR value is fixed to -10 dB and the coefficient correlation between the two SOIs is changed. As can be seen in figures 2.5a and 2.5b, the UG family and the AOLS algorithms, which are not designed to manage the correlated sources, give very interesting localization quality for all correlation coefficients ≤ 0.9. However, the reconstruction quality deteriorates for a perfect correction between the SOIs, especially for the pipelines where the Kurtosis metric is used. Concerning the reconstruction errors (figures 2.5c and 2.5d), UG and WAOLS exhibit acceptable results for correlation of 0.5 and totally decorrelated SOIs. Both methods tend to be less effective as the correlation coefficient increases. Contrary to UG and WAOLS filters, the LCMV beamformer was developed to process both uncor-

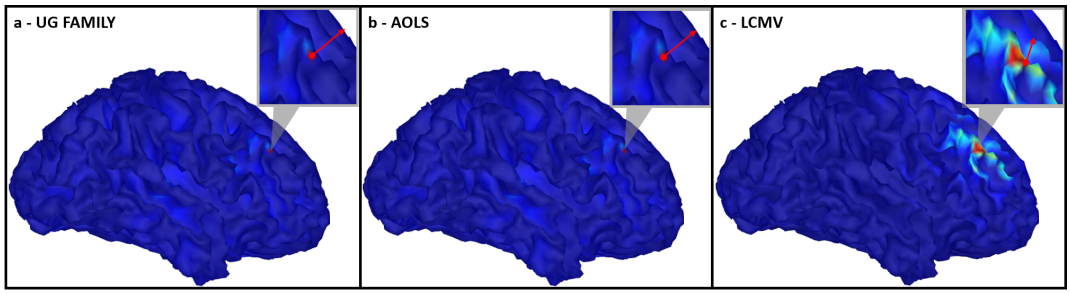


Figure 2.6: Absolute current density map of reconstructed sources at the maximum of the interictal spike before post-processing step for the three beamformers (a - UG Family, b - AOLS, c - LCMV). The dipole with the maximum activity is represented by a red arrow.

related and correlated SOIs. We can effectively see that, when the EEG similarity is used in the post-processing step, the localization (figure 2.5b) and the reconstruction (figure 2.5d) of both SOIs are very effective whatever the correlation level. However, the LCMV performances are questionable when the Kurtosis metric is exploited (figures 2.5a and 2.5c). Indeed, in the case of one source scenario, these performances tend to deteriorate for strong SNR values (figure 2.4c). Furthermore, as displayed in figure 2.5a, the localization quality is unsatisfactory for two source scenario.

Discussion: part 2 Experiment 2 shows that, as expected, the behaviors of the UG family methods are identical to those obtained in experiment 1, where the UG is used in the second step of the pipeline. Indeed, as discussed and demonstrated in section 2.3.3, the filters calculated by the UG, UNG and Stdr methods are equivalent up to a multiplicative scalar. However, the use of the Kurtosis and the EEG similarity metrics overcome this ambiguity (see section 2.3.4 for details). The performances of the AOLS beamformer are also quasi-equivalent with those of the UG family. This is explained by the fact that this algorithm is an all-in-one method where the Shrinkage preprocessing and the beamformer calculation steps are merged in one step (see section 2.3.3). Experiment 2 proves that a two-step strategy (Shrinkage + UG family) and the all-in-one strategy (AOLS) behave similarly.

The most intriguing results are those presented by the LCMV method followed by the Kurtosis metric. As pointed previously, in the case of one source scenario, these performances tend to deteriorate for strong SNR values. Likewise, the LCMV is not effective when the Kurtosis metric is used in post-processing, whatever the level of correlation between the two SOIs (two sources scenario). The behavior of the LCMV can be explained as follows: we have shown in section 2.3.3 that theoretically, the LCMV can handle correlated sources. In practice, the LCMV suffers from signal leakage. In other words, it generates an extended solution, i.e., a patch around the SOIs. This problem is illustrated in figure 2.6, where we clearly see that contrary to UG and AOLS family methods which give a very focal source solution, the LCMV generates an extended solution. The results of Experiment 2 show that the Kurtosis metric,

which is based only on the statistics of the estimated sources, is not robust to this signal leakage phenomenon. Indeed, the metric tends to select a source in the vicinity of the SOIs. On the other hand, the EEG similarity metric seems to be more robust to the signal leakage problem and can choose the good SOIs among different sources of the patch.

Considering the conclusions of experiments 1 and 2, we propose to process and compare the results of four different pipelines on the real data. These pipelines are built as follows i) in step 01, the Shrinkage algorithm is used to preprocess the data, ii) in step 02, AOLs and LCMV are retained, and iii) in step 03, the new proposed metric EEG similarity is compared to the Kurtosis one.

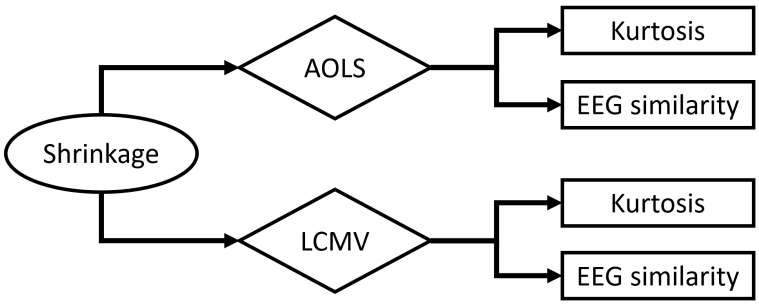


Figure 2.7: Representation of the four selected pipelines applied to real data.

2.4.3 Experiment 3: Application on real data

In this experiment, we propose to qualitatively measure the efficiency of the best beamforming pipelines on real hr-EEG data acquired from patients with drug-resistant partial epilepsy, for whom a surgical intervention can be considered to remove the epileptogenic brain regions. Indeed, in this case, a precise delineation of the epileptic regions is essential for the success of resective surgery. As indicated at the end of section 2.4.2, the beamformers AOLS and LCMV will be used on data preprocessed with the shrinkage method. The SOI localization will be done using the Kurtosis and the EEG Similarity metrics.

Data acquisition

A 256-channel hr-EEG was recorded for one hour, at 1000 Hz following the procedure approved by the National Ethics Committee for the Protection of Persons (CPP, agreement number 2012-A01227-36). The patient gave his written informed consent to participate in this study. This scalp recording revealed interictal spike activity at the left frontal basal electrode AFP9h, located above the subject left eyebrow (figure 2.7a). From this interictal epileptic activity, 8 spikes were visually selected away from the occurrence of any artefacts (muscle activity, blood pulsation, eye blinks). Each spike was centered in a 1 second window and all 8 windows were averaged to improve the

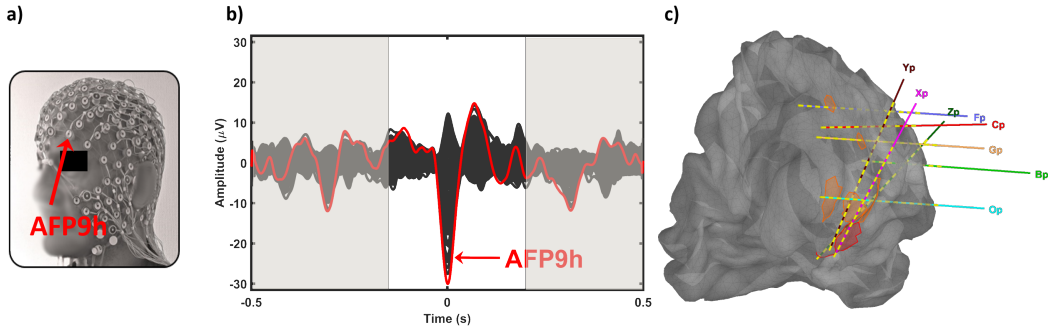


Figure 2.8: Patient-specific data. In a), the hr-EEG headset with 257 sensors. The sensor with the maximum of activity is point out in red. In b), the EEG recording in which the clear portion represents the time point selected for the SCM computation. In b), sensors on the depth electrodes are depicted in yellow. The area of spike maximal activity is represented in red. It covers approximately 5 cm^2 of cortex with parts of the left medial orbital sulcus, the left gyrus rectus and the left suborbital sulcus. The areas of spike propagation are depicted in orange and concerns multiple zone (left gyrus rectus, suborbital sulcus, anterior cingulate sulcus and medial anterior cingulate sulcus).

SNR ratio (figure 2.7b). As part of the presurgical evaluation, SEEG recordings were also performed using 8 intracerebral electrodes (10 ± 18 lead, length: 2 mm, diameter: 0.8 mm, 1.5 mm apart, sampling frequency: 512Hz) placed intracranially according to the Talairachs stereotactic method in the left frontal and temporal regions. The positioning of the electrodes was determined from an available non-invasive information and hypotheses about the localization of his EZ. The exact 3D coordinates of each electrode contact were determined after co-registering the CT scan showing the intracerebral leads onto the structural MRI image using a 6-parameters rigid-body transformation [158, 159]. Segmented anatomical data obtained with Freesurfer from the patient 3D-T1 MRI as well as the location of depth electrode contacts were imported within Brainstorm [40]. The areas of spike maximal activity and of spike propagation were delineated on the patient cortical mesh according to conclusions obtained after visual analysis of intracerebral interictal activity by the expert neurologist. They are represented in figure 2.7c in red and orange, respectively. Thanks to the presurgical evaluation, it has been determined that the EEG signal observed results from the activity of one source. This experiment is therefore similar to that of experiment 1 in section 2.4.2. The SCM matrix is computed using 350 time points around the spike (clear part in 2.7b).

Qualitative results

Figure 2.11 displays the localization and the SOIs reconstruction results obtained by the four pipelines. It can be seen that the AOLS + EEG Similarity pipeline offers the best localization performance. Indeed, the SOI is localized inside the targeted area delineated by the expert neurologist (red area on the Groundtruth). The three other pipelines are not able to well localize the SOIs. We can also remark that even if the reconstructed SOI of the AOLS + EEG Similarity pipeline seems to be contaminated by

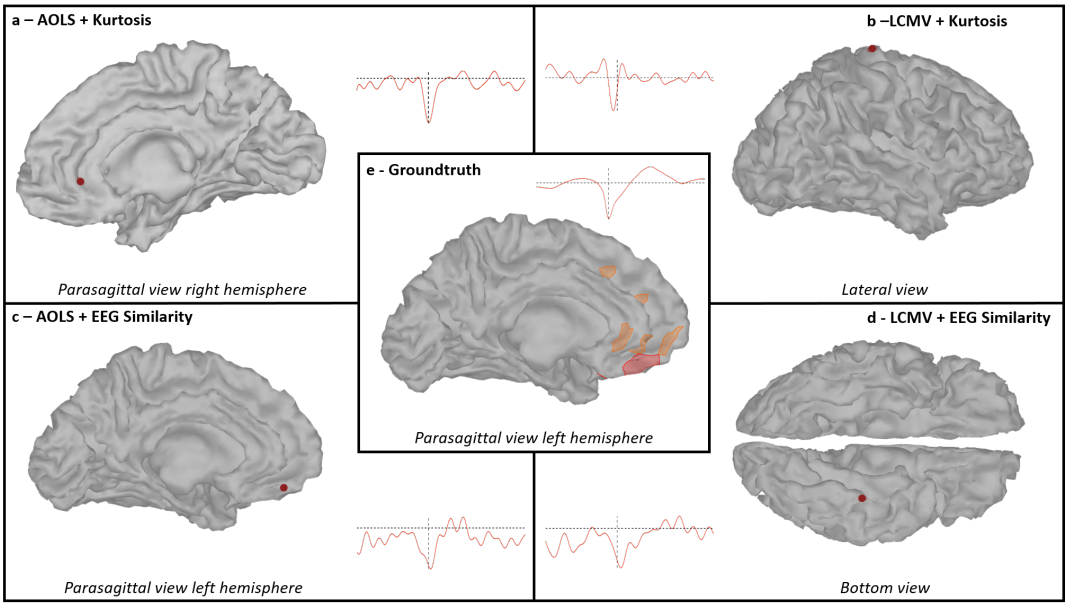


Figure 2.9: Effect of the AOLS and LCMV beamforming methods applied on real data. The upper part shows the localization results of the kurtosis criterion for the AOLS (a) and LCMV (b) methods, respectively. The lower part shows the localization results of the EEG similarity (c and d). The brain view has been carefully oriented to present the areas containing the SOI. On the side of the results is the estimated SOI activity. It is presented along a vertical and horizontal axis corresponding to the time index of the interictal spike maximum and the reference, respectively.

slow waves, it gives an interesting correlation level of 70% between the reconstructed SOI and the SEEG measured spike. The obtained results on real data are in agreement with those obtained in simulated data. Indeed, we can see that the good performance is given using an adequate combination of the beamformer and the post-processing metric, namely AOLS + EEG Similarity. Clearly, EEG Similarity which exploits the time course of the EEG data is very efficient to well localize the SOIs. Regarding the SOIs time course reconstruction, it can depend on either: i) the robustness of the response of the pipeline used with respect to noise, or ii) the considered source model. Indeed, for the latter, it is important to underline that even if the studied patient presents a very focal EZ, the targeted SOI is a distributed source. This can possibly explain that the shape of the reconstructed SOI is not perfect.

2.4.4 Conclusion

In this section, we propose to classify different beamforming pipelines to solve the BSI-EEG problem for an equivalent source model. Indeed, although the efficiency of such methods is extensively discussed in the MEG context, only a few studies are provided when EEG data are used. In addition, whatever the used data, most existing studies propose to evaluate either the beamforming filter calculation or the post-processing method, but never to assess the efficiency of the global beamforming pipeline, namely the preprocessing, the filtering (beamformer) and the post-processing. To found the more suitable pipeline, we surveyed all of those practices

to formulate different pipeline candidates. Then, we assess these candidates through various scenarios, in the context of simulated epileptic EEG data, to identify those that can be effective on real data. The results on simulated data show that Ledoit's shrinkage method, the AOLS beamformer and the new EEG similarity metric pipeline offers the best performances in term of EZ localization and sources reconstruction when only one source is to be found. In the case of correlated sources, the LCMV beamformer seems to offers the best result, but attention must be paid to leakage effect that can lead to poor localization. Obtained results on real data of one epileptic patient, where the observed epileptic EEG signature seems to be generated from the activity of one source, are in agreement with those obtained in simulated data. Indeed the Ledoit's shrinkage method, the AOLS beamformer and the new EEG similarity metric pipeline give the best approximation of the EZ. However, even if the quality of the reconstructed SOIs is quite satisfactory, we noted that their baselines are not smooth enough compared to the real sEEG. This improper reconstruction may be deleterious as it is involved in the localization criteria. To solve this problem, we propose to estimate the beamformer filter under an additional temporal-smoothing constraint.

2.5 Temporal-Smoothing-based beamforming method

Considering that the non-perfect reconstruction of the SOIs is due to a flaw in the beamforming methods, we propose to add a regularization term to promote the signal smoothness. This regularization term will be merged into the covariance matrix during the optimization step without too much methodological adjustment to the original methods already presented. In the sequel, beamforming methods containing this regularization term will be called Temporal-Smoothing-based beamforming methods.

2.5.1 Method

The main contribution of the Temporal-Smoothing-based beamforming methods is the sole addition of an extra constraint in the cost function. The Temporal-Smoothing constraint reduces the amplitude difference between each time sample, which can be smoothly done using a L_2 norm and a differential operator [160, 161, 162, 60]. Let us define the differential operator $\boldsymbol{D} \in \mathbb{R}^{T \times T}$, whose elements D_{t_1,t_2} are given by:

$$D_{t_1,t_2} = \begin{cases} 1 & \text{if } t_1 = t_2 \\ -1 & \text{if } t_1 = t_2 + 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.45}$$

Then a general definition of the Temporal-Smoothing-based beamforming can be stated as the following:

$$\begin{aligned} \min_{\boldsymbol{w}(\theta_\ell)} \quad & \boldsymbol{w}(\theta_\ell)^\intercal \boldsymbol{C_x} \boldsymbol{w}(\theta_\ell) + \lambda ||\boldsymbol{s}(\theta_\ell)\boldsymbol{D}||_\mathsf{F}^2 \\ \text{s. t.} \quad & \boldsymbol{w}(\theta_\ell)^\intercal \boldsymbol{g}(\theta_\ell) = r,\ \boldsymbol{s}(\theta_\ell) = \boldsymbol{w}(\theta_\ell)^\intercal \boldsymbol{X} \end{aligned} \tag{2.46}$$

where $\boldsymbol{X}$ is a matrix containing multiple realizations of the N-dimensional SOSE random process $\{\boldsymbol{x}(t)\}$. λ is a hyper-parameter that makes the trade-off between the beamforming cost-function and the Temporal-Smoothing constraint. The time support used on the observations $\boldsymbol{X}$ is chosen to contain the interictal spike corresponding to the time support used to compute the data SCM. As for all the methods in this chapter, we can solve the optimization problem using the Lagrangian method such as:

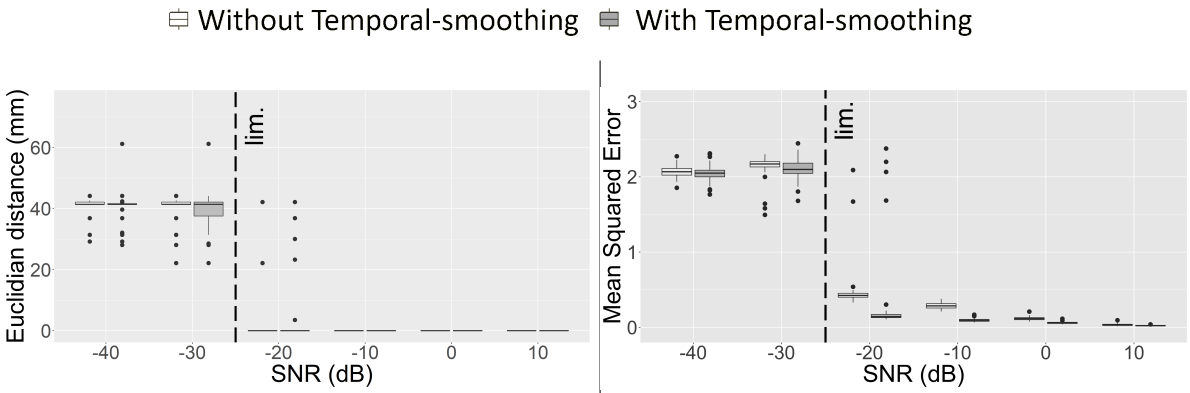
$$\mathcal{L}(\boldsymbol{w}(\theta_\ell), \alpha(\theta_\ell)) = \boldsymbol{w}(\theta_\ell)^\intercal \widehat{\boldsymbol{C_x}} \boldsymbol{w}(\theta_\ell) + \lambda ||\boldsymbol{w}(\theta_\ell)^\intercal \boldsymbol{X}\boldsymbol{D}||_\mathsf{F}^2 + \alpha(\theta_\ell)(\boldsymbol{w}(\theta_\ell)^\intercal \boldsymbol{g}(\theta_\ell) - \boldsymbol{r}) \tag{2.47}$$

where $\widehat{\boldsymbol{C_x}}$ is the data SCM and $\boldsymbol{r}$ is the desired beamformer response. Since all the elements in (2.47) are differentiable, the optimal point is reached where the Lagrangian derivative is null. Using the following property of the Trace [163] $\text{Tr}(ABC) = \text{Tr}(BCA)$, we get:

$$\begin{aligned} \frac{\delta\{\mathcal{L}(\boldsymbol{w}(\theta_\ell), \alpha(\theta_\ell))\}}{\delta\{\boldsymbol{w}(\theta_\ell)\}} &= 2\widehat{\boldsymbol{C_x}} \boldsymbol{w}(\theta_\ell) + 2\lambda \boldsymbol{X}\boldsymbol{D}\boldsymbol{D}^\intercal \boldsymbol{X}^\intercal \boldsymbol{w}(\theta_\ell) + \alpha(\theta_\ell)\boldsymbol{g}(\theta_\ell) = 0 \\ 2(\widehat{\boldsymbol{C_x}} + \lambda \boldsymbol{X}\boldsymbol{D}\boldsymbol{D}^\intercal \boldsymbol{X}^\intercal)\boldsymbol{w}(\theta_\ell) &= \alpha(\theta_\ell)\boldsymbol{g}(\theta_\ell) \\ 2\widetilde{\widehat{\boldsymbol{C_x}}} \boldsymbol{w}(\theta_\ell) &= \alpha(\theta_\ell)\boldsymbol{g}(\theta_\ell) \\ \boldsymbol{w}(\theta_\ell) &= \frac{1}{2}\widetilde{\widehat{\boldsymbol{C_x}}}^{-1} \boldsymbol{g}(\theta_\ell)\alpha(\theta_\ell) \end{aligned} \tag{2.48}$$

Excluding the new data SCM $\widetilde{\widehat{\boldsymbol{C_x}}} = (\widehat{\boldsymbol{C_x}} + \lambda \boldsymbol{X}\boldsymbol{D}\boldsymbol{D}^\intercal \boldsymbol{X}^\intercal)$, this final form is identical to the one formulated for the Unit Gain family of beamforming methods (2.22). Hence, the analytical solution proposed in the previous methodology for the UG, UNG and Stdr are directly applicable for the Temporal-Smoothing-based beamforming provided a new data SCM $\widetilde{\widehat{\boldsymbol{C_x}}}$. This development can be also applied to the LCMV and the AOLS as well. The latter being a trade-off between the Temporal-Smoothing-based UG beamformer and the matched filter.

In conclusion, the only condition to applied the Temporal-Smoothing-based beamforming is to provide the modified SCM $\widetilde{\widehat{\boldsymbol{C_x}}}$. It is interesting to note that the Temporal-Smoothing constraint can be viewed as a preprocessing step since the impact of the regularization modify $\widehat{\boldsymbol{C_x}}$(cf table 2.1). In the first instance, the hyper-parameter λ value has been fixed to 1.



(a) EEG similarity localization error

(b) EEG similarity reconstruction error

Figure 2.10: Comparison between the classic beamforming and Temporal-Smoothing-based beamforming on the preprocessing scenario for one source with the best pipeline: localization results are on the left and signal reconstruction on the right. The dashed line represents the limit of baseline identification.

2.5.2 Experimental results

In this section, a comparative study is proposed to assess the temporal smoothing effect. To do so, the best pipeline deduced in the previous experiments for the one-source scenario, namely the Ledoit's shrinkage method, the AOLS beamformer and the new EEG similarity metric, with and without temporal-smoothing, is studied on simulated and real data. Note that to satisfy readers' curiosity, all the experiments results are presented in figures 2.2 to 2.5 are also provided in appendix B, where the temporal-smoothing regularization term is added.

Simulated results

As for experiment 1, the simulated 257 channels EEG data are generated from a single epileptic source by varying the SNR from -40 dB to 10 dB with a step of 10 dB. To assess the sources localization behaviors, figure 2.10a (left column) represents the Euclidian distances calculated between the true position of the spike-like source and the estimated source position. The quality of the time courses reconstructions is depicted in figure 2.10b (right column). In both figures, black and gray colors stand to the best pipeline without and with temporal-smoothing, respectively. For all SNR values, the localization performances of the temporal smoothing-based beamformer are quasi-identical with those of the initial beamformer pipeline. This suggests that the temporal smoothing does not provide any advantage to an accurate delineation of the EZ. Regarding the sources reconstruction quality, we clearly observe that the temporal smoothing beamformer gives the best results for all SNR. This improvement is particularly visible for realistic SNR values of -20 dB and -10 dB.

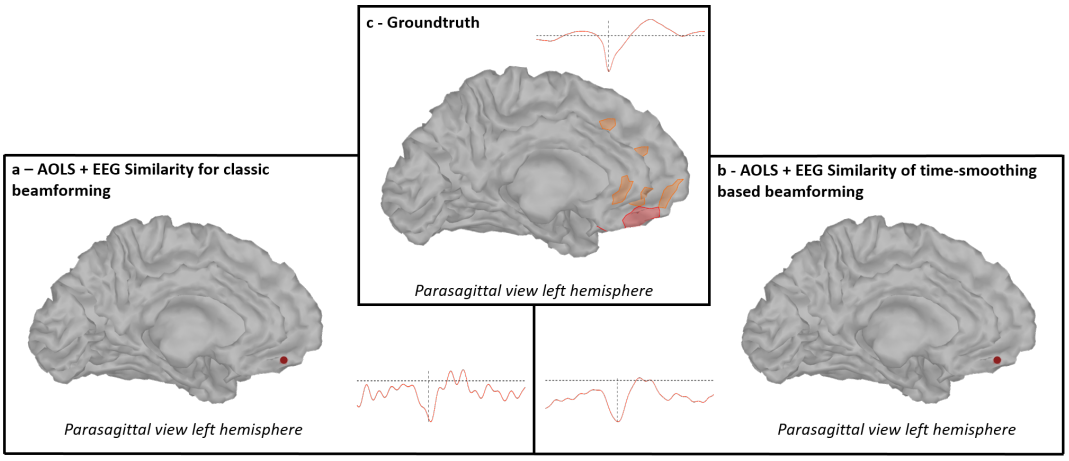


Figure 2.11: Effect of the Temporal-Smoothing-based beamforming methods applied on real data. The lower part shows the localization results of the best pipeline for the classic beamforming and the Temporal-Smoothing-based beamforming (a and b). The brain view has been carefully oriented to present the areas containing the SOI. On the side of the results is the estimated SOI activity. It is shown along a vertical and horizontal axis corresponding to the time index of the interictal spike maximum and the reference, respectively.

2.5.3 Real data analysis

The temporal-smoothing constraint is employed with the best pipeline on the real data. As we can expect from the simulations results, the localization does not change for the real source application. The SOI is still located in the EZ situated in the sub-orbital sulcus. As we can observe in figure 2.11, only the reconstruction results are different. Indeed, the reconstructed source using temporal smoothing (figure 2.11.b) looks more like the one recorded by SEEG (figure 2.11.c). No slow waves seem to be observed, contrary to the initial pipeline (figure 2.11.a). Moreover, the top of the spike perfectly matches the time sample of the maximum scalp activity (represented with the dashed line). We can conclude that the Temporal-Smoothing-based beamforming approach is more efficient than traditional beamforming to reconstruct real sources activity.

2.6 Conclusion

In this chapter, the main motivation is to give insights into how best to exploit beamforming to solve the BSI-EEG problem. Indeed, as stated before, beamforming is used mainly in MEG, but they are currently understudied in EEG. Another key question addressed here concerns the selection of the best beamforming pipeline by the analyzed data. Indeed, existing studies propose to evaluate the beamforming filter calculation or either the post-processing method, but never to assess the efficiency of the global beamforming pipeline, namely the preprocessing, the filtering (beamformer) and the post-processing. Hence, we first present a taxonomy of possible beamforming pipelines by giving crucial theoretical details. After that, a prior selection of the ad-

equate pipeline candidates is selected to be evaluated in the BSI-EEG context. More precisely, the choice of the beamformer pipelines is underpinned on practical considerations, such as i) the lack of knowledge of the position and the number of SOIs in EEG context, ii) the need or not of an additional reference and iii) the ability of the beamformer filter to both localize and reconstruct the SOIs. Some new improvements have also been proposed to complement existing ones. They consist of a new post-processing metric, named EEG similarity, which is particularly suitable in the EEG context, and a new beamformer filter that promotes the smoothness of the reconstructed SOIs.

In this thesis, the epileptic EEG context is chosen to study the behaviors of different pipelines. More precisely, a comparative analysis is first performed on realistic simulated epileptic EEG data to help us determine the most efficient pipeline. The result is that Ledoit's shrinkage method, coming from the random matrices theory, gives the best results as a preprocessing. The AOLS beamformer, which uses another form of shrinkage, also improves the quality of results for uncorrelated sources. Finally, the proposed EEG similarity metric offers the best localization results. This trio applied to real data allows an exact localization of the EZ. However, we noticed that real data's estimated source time courses are not as smooth as the ground truth, hindering correct localization. To solve this problem, the new Temporal-Smoothing-based beamforming pipeline, applied to real data, shows a significant improvement in reconstructing epileptic sources in all experiments.

Even if the best pipeline gives very promising results both in simulated and real data, all experiments assume an equivalent source dipole model. However, as stated in the introduction and chapter 1, the distributed source model is more physiologically suitable for SOIs localization and reconstruction in a different practical situation, such as epilepsy. The dipole equivalent model has been used in this chapter to understand and review the beamforming methods since it was designed for it. Despite the extensive review on the subject in section 2.2, it is troublesome to find a suitable beamforming method dedicated to distributed sources that do not require challenging prior knowledge, such as the SOIs position. Regarding the LCMV beamformer, which can theoretically manage both punctual and distributed sources, the results reveal a constant location bias even for advantageous scenarios. For this reason, this method is considered unreliable and is excluded as a valid candidate for distributed sources-based beamforming. However, we cannot rule out all beamforming methods. Indeed, in contrast to most existing BSI methods, they allow us to efficiently reconstruct the time course of the SOI with a mathematical solution usually straightforward and analytical. Therefore, we hypothesize that the extended source-based beamforming method will improve confidence in the estimated source. Thus, the objective of the next chapter will be to propose a new way to localize and reconstruct extended sources with beamforming.

Chapter 3

Beyond traditional beamforming methods: distributed source localization

As discussed in the previous chapter, beamforming methods efficiently estimate the time course of the sources in the equivalent dipole source model. The solution of the mathematical problem is usually straightforward and analytical. Its main flaw is its sensitivity to interference generated by correlated sources. Solutions have been proposed to limit their effect, but they are challenging to set up in the context of spatially distributed sources since they require strong priors such as the location of the dipoles within the SOIs (see section 2.2.2). Moreover, the beamformer specially designed to handle such a model of sources, or LCMV, requires to invert a rank-deficient matrix. As we underlined in the previous study, the algorithm proposed to solve such a problem presents a localization bias.

On the other hand, non-parametric methods are compatible with the distributed source model without the knowledge of any obliging priors. Some of them also present a straightforward and analytical solution to the inverse problem like wMNE and sLORETA. However, it has been shown that these methods are not the most suitable for the epileptic source location problem as they promote a smooth spatial distribution. On the contrary, regularized least-squares methods promoting sparsity such as Variation-Based Sparse Cortical Current Distribution (VB-SCCD) [59], Elastic Net [164], Source Imaging based on Structured Sparsity (SISSY) [62] give promising results. These require more complex algorithms to be solved, but the subject has been widely covered, and nowadays, they are pretty simple to implement. The advantage of regularization methods is that the more constraints they present, the better they work provided a correct trade-off between each constraint, i.e., the hyper-parameters are correctly tuned (see section 2.5.1). Hence these questions: could such approaches

be more efficient by regularizing them in a beamforming way? And how does such a method differ from traditional methods in terms of efficiency?

In order to answers those questions, this chapter is organized as follows: in section 3.1, we formulate the assumptions on the distributed source model made throughout this chapter. Section 3.2 presents a new distributed brain source method named SABLE (Sparsity And Beamforming for brain source Localization and Estimation) that shows how both beamforming steps (i.e., filtering and localization) can be merged to deal with spatially distributed sources. A preliminary set of numerical experiments show the promising good behavior of SABLE in the context of epilepsy. Section 3.3 is devoted to the proposition of an adequate pipeline to compare SABLE with other methods. This pipeline is inspired by those implemented in the software Brainstorm. Nevertheless, it proposes innovative solutions to problems occurring in the analysis of EEG data. Finally, in section 3.4, the localization results of Sparsity And Beamforming for brain source Localization and Estimation (SABLE) will be compared against three BSI methods, namely wMNE, sLORETA and coherent Maximum Entropy on the Mean (cMEM), on refined simulated data and real epileptic data.

3.1 Assumptions for the distributed source model

Compared to the previous chapter, we focus on an optimization problem rather than on statistical prior. Thus, we adopt a different notation system using vector and matrix form to ease the reading. Let $\boldsymbol{x}$ be an N-dimensional EEG vector recorded at a given point in time. This vector can be assumed to stand for one realization of an N-dimensional random vector with covariance matrix denoted by $\boldsymbol{C}_x$ such as:

$$\boldsymbol{x} = \boldsymbol{G}\boldsymbol{s} + \boldsymbol{\nu} \tag{3.1}$$

where $\boldsymbol{s}$ denotes the corresponding L-dimensional brain source activity vector and where $\boldsymbol{\nu}$ stands for an instrumental noise vector. $\boldsymbol{G}$ remains the so-called leadfield matrix and the source orientation is still fixed to the normal of the cortical surface. For the distributed source case, if we consider the hyper-synchronicity of the dipoles into small delimited cortical regions named patches, the EEG system can be further developed such as:

$$\begin{aligned}
\boldsymbol{x} &= \boldsymbol{G}\boldsymbol{s} + \boldsymbol{\nu} \\
&= \sum_{k=1}^{K} \sum_{\ell_k \in \Omega_k}^{L_k} \boldsymbol{g}(\theta_{\ell_k}) s(\theta_{\ell_k}) + \sum_{\ell \notin \cup_{k=1}^{K} \Omega_k} \boldsymbol{g}(\theta_\ell) s(\theta_\ell) + \boldsymbol{\nu} \\
&= \sum_{k=1}^{K} \boldsymbol{G}_k^{(e)} \boldsymbol{s}_k^{(e)} + \boldsymbol{G}^{(b)} \boldsymbol{s}^{(b)} + \boldsymbol{\nu}
\end{aligned} \tag{3.2}$$

$$= \boldsymbol{G} \sum_{k=1}^{K} \boldsymbol{\phi}_k \bar{s}_k^{(e)} + \boldsymbol{\nu}' \text{ with } \boldsymbol{\phi}_k \bar{s}_k^{(e)} = \boldsymbol{s}_k^{(e)}$$

where Ω_k is the set of indexes of the dipoles that belong to the k-th epileptic extended source. $s^{(e)}$ is the L_k-dimensional epileptic source activities vector of the k-th patch and $s^{(b)}$ is the complementary background activity vector. $\boldsymbol{G}_k^{(e)}$ is the matrix whose columns are the leadfield vectors of the epileptic dipoles inside the k-th distributed source. Equivalently $\boldsymbol{G}^{(b)}$ is a matrix whose columns are the leadfield vectors of the background dipoles. If the electrical hyper-synchronicity of the epileptic sources is assumed (section 1.2), $\boldsymbol{s}_k^{(e)}$ can be seen as a fixed time series allocated to different positions. Therefore, it can be considered as the product of two parameters, one for the time series and the other with the spatial information. This assumption is at the basis of the last proposition where $\bar{s}_k^{(e)}$ is the mean electrical activity of the k-th SOI and where the L-dimensional vector $\boldsymbol{\phi}_k$ is a sparse vector that informs us on the spatial support of that particular SOI.

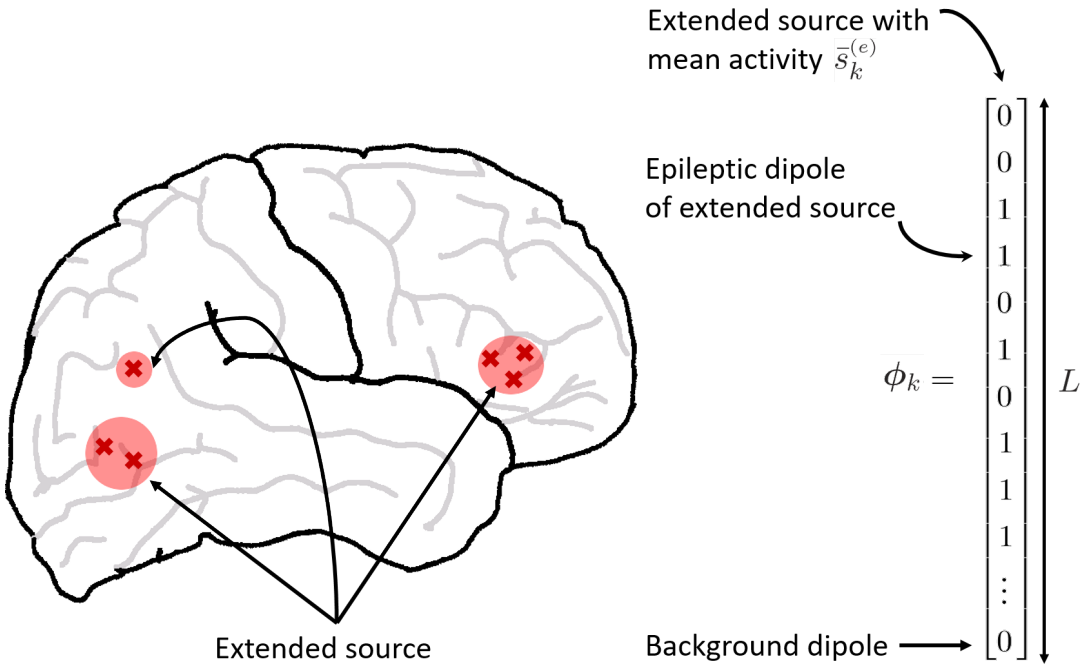


Figure 3.1: Schematic representation of the new proposed system with the role of $\boldsymbol{\phi}_k$.

The new method presented in this thesis aims at estimating both the spatial support $\boldsymbol{\phi}_k$ of all K epileptic sources and their mean activity $\bar{s}_k$. For now, we suggest to use a deflation scheme to iteratively estimate the SOIs as we already did for beamforming localization methods (see section 2.3). As depicted in figure 3.1, $\boldsymbol{\phi}_k$ will delimit Region(s) Of Interest (ROI(s)) that can be spatially distant; hence, K represents the number of SOIs within the electrical activity is the same instead of the number of ROIs.

3.2 SABLE: a novel distributed source beamforming method

In this section, the novel method called SABLE is presented and evaluated on a primary set of simulations to answer the first objective of this chapter. In particular, we will rely on the new mathematical model elaborated above and on the experimental pipeline presented in the previous chapter (see section 2.4).

3.2.1 Methods

At first, lets us recall the principle of the UG beamforming and present the SISSY approach on the basis of which SABLE is formed. As an aside, the UG beamforming has been chosen over the AOLS beamformer as its optimization problem is the most convenient handle for an initial proposition.

Unit Gain beamforming

Beamforming methods aim at creating a spatial filter $\boldsymbol{w}(\theta_\ell)$, which reconstructs the source time course at a given position θ_ℓ when it is applied to the EEG signals. The classical UG method solves the following minimization problem:

$$\min_{\boldsymbol{w}(\theta_\ell)} \ \boldsymbol{w}(\theta_\ell)^\top \boldsymbol{C_x} \boldsymbol{w}(\theta_\ell) \quad \text{s. t. } \boldsymbol{w}(\theta_\ell)^\top \boldsymbol{g}(\theta_\ell) = 1 \tag{3.3}$$

The constraint $\boldsymbol{w}(\theta_\ell)^\top \boldsymbol{g}(\theta_\ell) = 1$ ensures that the signal coming from the position θ_ℓ goes through the filter without distortion (see section 2.2). However, it also allows signals from other locations to pass through if and only if they are correlated with the one coming from θ_ℓ. This correlation is the cause of the source cancellation effect, which gradually deteriorates the reconstructed correlated sources.

In order to overcome this problem, it was proposed to combine the leadfield vectors of correlated sources [119, 120] and to calculate a single filter for these correlated sources provided that we know their number and their location. Such an approach cannot be used for spatially distributed sources since we do not know the position and the number of the dipoles constituting the SOI in advance.

Once the filter is applied, localization is performed by optimizing a metric computed from the estimated signals. In the equivalent dipole source model, we will look for an extremum while thresholding near the extremum will be used for spatially distributed sources. The latter solution is a little abusive, but it avoids formulating the aforementioned strong priors [142, 141, 143].

Source imaging based on structured sparsity

SISSY is a regularized least-squares method which aims at finding the solution that best describes the measurements in a least-squares sense. Initially named Sparse Variation-Based Sparse Cortical Current Distribution (SVB-SCCD) [165], the SISSY method solves the following optimization problem:

$$\min_{\boldsymbol{s}} ||\boldsymbol{x} - \boldsymbol{G}\boldsymbol{s}||_2^2 + \lambda(||\boldsymbol{T}\boldsymbol{s}||_1 + \alpha||\boldsymbol{s}||_1) \tag{3.4}$$

where the L_2 norm stands for the reconstruction error, where the L_1 norms promote sparsity in the (transformed) spatial domain and where $\boldsymbol{T}$ is the space gradient operator. This combination of L_1 norms, well-known as the fused Least Absolute Shrinkage and Selection Operator (LASSO) [166], promotes a piece-wise constant solution containing few active dipoles with few activity differences. Hyper-parameters λ and α play a balancing role between the three penalty terms and must be adjusted by hand. The Alternating Direction Method of Multipliers (ADMM) [167] is used to solve this minimization problem.

The ADMM algorithm is an efficient and straightforward algorithm for constrained convex optimization. It is based on the idea of alternatively updating the augmented Lagrangian variables and multipliers until convergence. Consequently, latent variables can be introduced and used to ease the update computation, which the original author of SISSY chose to deal with the L_1. Compared to others [62], SISSY was found to facilitate source separation and to better estimate the source spatial extent. Moreover, this method does not need priors on the source number/position.

3.2.2 The proposed SABLE method

The SABLE method aims at taking advantage of both techniques. Then it was designed by i) merging both steps of the beamforming procedure (i.e., filtering and localization) and ii) modeling the hyper-synchronous nature of the activity within a SOI. This idea led us to consider the following minimization problem:

$$\begin{aligned} \min_{\boldsymbol{w},\boldsymbol{\phi}} \boldsymbol{w}^\top \boldsymbol{C_x} \boldsymbol{w} + \lambda_1 ||\boldsymbol{x} - \boldsymbol{G}\boldsymbol{s}||_2^2 + \lambda_2(||\boldsymbol{T}\boldsymbol{\phi}||_1 + \lambda_3||\boldsymbol{\phi}||_1) \\ \text{s. t. } \boldsymbol{s} = \boldsymbol{\phi}\bar{s},\ \bar{s} = \boldsymbol{w}^\top \boldsymbol{x} \text{ and } \boldsymbol{w}^\top \boldsymbol{G}\boldsymbol{\phi} = 1 \end{aligned} \tag{3.5}$$

where λ_i is the i-th penalty parameter and $\boldsymbol{\phi}$ is a sparse vector that informs us on the spatial support of one of the SOI. In our first set of simulations, λ_1 will be set to 1 in order to give the same weight to both steps of the UG procedure (i.e., filtering and localization). We decided to solve the problem (3.5) using the ADMM optimization technique. Then let us consider the following augmented Lagrangian function:

$$\begin{aligned}\mathcal{L}_{\rho,\lambda}(\boldsymbol{w},\boldsymbol{\phi},\boldsymbol{\psi},\boldsymbol{\gamma},u,\boldsymbol{v},\boldsymbol{z}) = \boldsymbol{w}^\top \boldsymbol{C_x} \boldsymbol{w} + \lambda_1 \, ||\boldsymbol{x} - \bar{\boldsymbol{g}} \boldsymbol{w}^\top \boldsymbol{x}||_2^2 \\ + \lambda_2 ||\boldsymbol{\psi}||_1 + \lambda_2 \lambda_3 ||\boldsymbol{\gamma}||_1 + u(\boldsymbol{w}^\top \bar{\boldsymbol{g}} - 1) + \boldsymbol{v}^\top (\boldsymbol{T}\boldsymbol{\phi} - \boldsymbol{\psi}) \\ + \boldsymbol{z}^\top (\boldsymbol{\phi} - \boldsymbol{\gamma}) + \frac{\rho}{2}(||\boldsymbol{w}^\top \bar{\boldsymbol{g}} - 1||_2^2 + ||\boldsymbol{T}\boldsymbol{\phi} - \boldsymbol{\psi}||_2^2 + ||\boldsymbol{\phi} - \boldsymbol{\gamma}||_2^2)\end{aligned} \tag{3.6}$$

where ρ stands for a penalty parameter, $\bar{\boldsymbol{g}} = \mathbf{G}\boldsymbol{\phi}$ is the average leadfield vector and $u \in \mathbb{R}$, $\boldsymbol{v} \in \mathbb{R}^P$ and $\boldsymbol{z} \in \mathbb{R}^L$ are the scaled Lagrangian multipliers. Note that both constraints $\boldsymbol{s} = \boldsymbol{\phi}\bar{s}$ and $\bar{s} = \boldsymbol{w}^\top \boldsymbol{x}$ were directly inserted into the penalty term $||\boldsymbol{x} - \boldsymbol{G}\boldsymbol{s}||_2^2$. In addition, two latent variables $\boldsymbol{\psi} \in \mathbb{R}^P$ and $\boldsymbol{\gamma} \in \mathbb{R}^L$ were used to replace $\boldsymbol{T}\boldsymbol{\phi}$ and $\boldsymbol{\phi}$, respectively, in order to split the initial minimization problem. The augmented Lagrangian function (3.6) is then alternatively minimized using the following updating rules:

$$\begin{aligned}\boldsymbol{w}_{n+1} &= \left[2\lambda_1 \bar{\boldsymbol{g}}_n^\top \bar{\boldsymbol{g}}_n \otimes \boldsymbol{x}\boldsymbol{x}^\top + 2\boldsymbol{C_x} + 2\rho \bar{\boldsymbol{g}}_n \bar{\boldsymbol{g}}_n^\top\right]^{-1} \left[2\lambda_1 \boldsymbol{x}\boldsymbol{x}^\top \bar{\boldsymbol{g}}_n - \rho \bar{\boldsymbol{g}}_n u_n^\top + \rho \bar{\boldsymbol{g}}_n\right] \\ \boldsymbol{\phi}_{n+1} &= \left[2\lambda_1 \boldsymbol{G}^\top \boldsymbol{G} \otimes \boldsymbol{w}_{n+1}^\top \boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{w}_{n+1} + \rho(\boldsymbol{G}^\top \boldsymbol{w}_{n+1} \boldsymbol{w}_{n+1}^\top \boldsymbol{G}) + \rho(\boldsymbol{T}^\top \boldsymbol{T} + I_d)\right]^{-1} \\ &\quad \left[\lambda_1 \boldsymbol{G}^\top \boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{w}_{n+1} + \boldsymbol{G}^\top \boldsymbol{w}_{n+1}(\rho - u_n) + \boldsymbol{T}^\top(\rho\boldsymbol{\psi}_n - \boldsymbol{v}_n) + (\rho\boldsymbol{\gamma}_n - \boldsymbol{z}_n)\right] \\ \boldsymbol{\psi}_{n+1} &= \text{prox}_{\frac{\lambda_2}{\rho}}\left(\boldsymbol{T}\boldsymbol{\phi}_{n+1} + \frac{1}{\rho}\boldsymbol{v}_n\right) \\ \boldsymbol{\gamma}_{n+1} &= \text{prox}_{\frac{\lambda_2\lambda_3}{\rho}}\left(\boldsymbol{\phi}_{n+1} + \frac{1}{\rho}\boldsymbol{z}_n\right) \\ u_{n+1} &= u_n + \rho(\boldsymbol{w}_{n+1}^\top \bar{\boldsymbol{g}}_{n+1} - 1) \\ \boldsymbol{v}_{n+1} &= \boldsymbol{v}_n + \rho(\boldsymbol{T}\boldsymbol{\phi}_{n+1} - \boldsymbol{\psi}_{n+1}) \\ \boldsymbol{z}_{n+1} &= \boldsymbol{z}_n + \rho(\boldsymbol{\phi}_{n+1} - \boldsymbol{\gamma}_{n+1})\end{aligned} \tag{3.7}$$

Note that $\boldsymbol{\phi}$ is initialized as a unit vector with all its entries equal to 1, which implies that all positions are eligible candidates at the beginning. The update of the latent variables $\boldsymbol{\psi}$ and $\boldsymbol{\gamma}$ is formulated using the proximity operator defined as:

$$\text{prox}_\beta(\boldsymbol{a}) = \underset{\boldsymbol{b}}{\text{argmin}}\{\frac{1}{2}||\boldsymbol{a} - \boldsymbol{b}||_2^2 + \beta||\boldsymbol{b}||_1\} \tag{3.8}$$

The proximity operator of the L_1 norm is well-known as the soft thresholding operator, which is a shrinkage operator (i.e., it moves a point toward zero) [168]. The SABLE algorithm finishes either after convergence of the variables $\boldsymbol{w}$ or $\boldsymbol{\phi}$, or if a maximum number of iterations is reached.

3.2.3 Primary results on simulations

In this part, the performance of the SABLE algorithm is studied on realistic simulated epileptic interictal EEG signals compared to the UG and SISSY methods.

Experiment and performance criterion

The simulated data are generated following the model presented in equation (1.1). First, a surface head model is computed from a T1 MRI of a reference patient with a segmentation method. The head tissues have been separated into 5 boundaries: grey/white matters, inner/outer skulls and scalp. The source space corresponding to the middle of the grey matter is composed of $L = 15002$ vertices on each of which a current dipole oriented perpendicularly has been assigned. The source matrix is created using signals generated using the neural mass model [31]. The epileptic source region is modeled by epileptic patches including 33 grid dipoles on the right gyrus parietalis inferior with a part inside the superior temporal sulcus. The remaining dipoles are distributed among 1000 scouts with the same background activity within a scout and uncorrelated activity between scouts. The scalp EEG of $N = 257$ electrodes, $T = 5120$ samples and the sampling rate of 512 Hz is generated. In addition, to assess the efficiency of the proposed method, different intensities of background EEG are added to the simulated epileptic data leading to different values of SNR (SNR $\in \{-10, 0, 10\}$ dB), and all presented results are averaged over 25 Monte Carlo runs. Regarding the hyper-parameters involved in the methods, λ_1 is fixed to 1, λ_2 is adjusted by hand to a value proportional to the noise level and λ_3 is set to 0.07 as recommended in [62].

Concerning the performance criteria, we first determine the active patch by thresholding the intensity of estimated sources at its sample with maximal power, which corresponds to the maximum of the epileptic spike. To do so, thresholding is applied for all methods in which all sources with an intensity superior or equal to 75% of the absolute maximal intensity are retained. Then, the Dipole Localisation Error (DLE) is used to quantify the source localization accuracy of the compared methods. The DLE presents a measure of similarity between the original and the estimated source configuration. The DLE is defined as:

$$\text{DLE} = \frac{1}{2L_{\mathcal{M}}} \sum_{m \in \Omega_{\mathcal{M}}} \min_{k \in \Omega_{\widehat{\mathcal{M}}}} ||\boldsymbol{\theta}_m - \boldsymbol{\theta}_k||_2 + \frac{1}{2L_{\widehat{\mathcal{M}}}} \sum_{m \in \Omega_{\widehat{\mathcal{M}}}} \min_{k \in \Omega_{\mathcal{M}}} ||\boldsymbol{\theta}_m - \boldsymbol{\theta}_k||_2$$

where $\Omega_{\mathcal{M}}$ and $\Omega_{\widehat{\mathcal{M}}}$ denote the original and the estimated sets of indexes of all dipoles of an active patch, respectively. $L_{\mathcal{M}}$ and $L_{\widehat{\mathcal{M}}}$ are the numbers of original and estimated active dipoles. $\boldsymbol{\theta}_m$ is the position of the m-th source.

Computer results

The DLE results of the three approaches, namely SABLE, UG and SISSY, are presented in Table 3.1 as a function of SNR. As it can be seen, SABLE clearly outperforms the other methods for SNRs equal to 0 and 10 dB. For an SNR value of -10 dB, the results of SABLE and UG are not satisfactory (even if the DLE of SABLE is

better than the one of UG). The bad performance of SABLE for weak SNR values is mainly due to the beamforming term included in its cost function. In other words, the hyper-parameters λ_1 and λ_2 should be tuned more finely in order to best balance the localization step. Figure 1 illustrates an example of BSI results obtained with the different methods for an SNR value of 10 dB. Obviously, the SABLE algorithm provides a better localization than the SISSY approach which tends to under-estimate the patch. Besides, figure 3.1 shows the difficulties encountered with the UG algorithm to handle the distributed source, even if we consider a set of estimated sources associated with the maximum selected by the thresholding step.

Table 3.1: Performance of SABLE in terms of DLE

	DLE		
Scenario	-10dB	0dB	10dB
UG	25.48	16.91	6.98
SABLE	16.54	2.97	1.88
SISSY	4.98	5	4.96

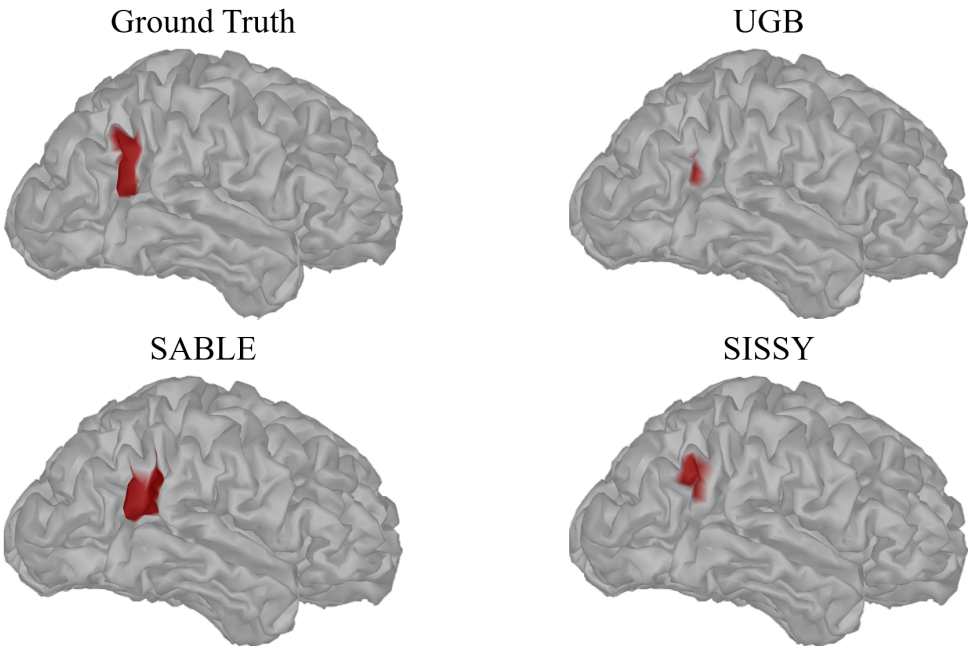


Figure 3.2: Source localization results at the output of UG, SABLE and SISSY for an SNR value of 10dB.

Conclusions

In this section, we propose a new method, named SABLE, for EEG-based brain source imaging. The SABLE method establishes a bridge between beamforming approaches and regularized least-squares methods promoting sparsity. The comparative study conducted using realistic EEG data, in the context of drug-resistant epilepsy tends

to show that SABLE is a promising method for the localization of brain spatially distributed sources. In order to mature this new method, we need to propose a refined pipeline to extend these preliminary results. To this aim, we will rely on methodologies already embedded in a computer application that is majorly adopted by the neuroscience community. That way, we approach the real conditions of applications and the possible diffusion of the SABLE algorithm. Our attention went to Brainstorm thanks to its ease of use and its data management system among the many existing softwares [169]. But before even considering the diffusion of this new method, we must first evaluate its performances compared to reference methods. The following section is dedicated to present the pipeline that we decided to apply on SABLE and the inverse method chosen from Brainstorm for a preliminary comparative study.

3.3 Toward the SABLE pipeline and the comparative study

The Brainstorm software [40] is a user-friendly open-source tool that brings together a large community of active users. From a technical point of view, it gathers essential functions for analyzing EEG and MEG data [170], such as i) the management of the database, ii) the resolution of the forward problem (i.e., the calculation of the leadfield matrix) and iii) the display of the results on a 3D mesh. The latter includes thresholding and analysis tools that facilitate interpreting the results. In fact, we used them to produce most of the imaging results in this thesis. The application also encapsulates external libraries like Brain Entropy in space and time (BEst) [171], MNE [172, 173] or OpenMEEG [81, 174]. In this section, we aim at realizing a comparative study between the mature BSI methods implemented in Brainstorm and both experimental methods SISSY and SABLE to assess their localization efficiency.

3.3.1 BSI methods of the Brainstorm software

Three BSI methods of Brainstorm have been retained: wMNE [53], sLORETA [55] and cMEM [175, 176]. The first two methods are classical and very used in EEG. They propose a quick and direct solution to the inverse problem, but the latter is not always physiologically plausible. As far as cMEM is concerned, it presents an exciting approach using parcellization and probability distribution to estimate the activation map of the brain. The smoothness is local and restricted inside the parcel, making it more physiologically plausible for locating epileptic sources.

Regarding the remaining BSI methods of Brainstorm, those specially designed for the equivalent source model have been discraded since they are not appropriate to the current BSI problem (see results of UG in the previous section). For the beamforming method LCMV, the Brainstorm implementation has not been retained as it strongly differs from the implementation given in the original paper. Indeed, the proposed beamformer seems to have been coded as a UG beamformer. Both remaining entropy

methods wMEM [177] and rMEM [178], from the BEst library, are not considered since they use priors based on wavelets that are out of the scope of this thesis.

After mastering the software structure, the Brainstorm methods can be called from an automated script regrouping the other methods and the other steps of the pipeline. Before presenting more deeper the pipeline, let us recall the assumptions and resolution process used by the Brainstorm methods.

wMNE

The wMNE method originates from the MNE [53] method, the first inverse problem method introduced for BSI. It has been proposed as an improvement to correct the depth bias of MNE, which favors the superficial source over the deep sources. To circumvent this bias, wMNE imposes a weight on each leadfield vector to equalize their energy. It is a methodology similar to the array gain beamforming presented in the previous chapter (see section 2.2). From an optimization point of view, it is a regularized least-squares based on the L_2 norm only contrarily to SISSY:

$$\min_{\boldsymbol{s}} ||\boldsymbol{x} - \boldsymbol{GKs}||_2^2 + \lambda||\boldsymbol{s}||_2 \tag{3.9}$$

where $\boldsymbol{K} \in \mathbb{R}^{L\times L}$ is the diagonal weight matrix of leadfield whose element K_{ℓ_1,ℓ_2} are defined as:

$$K_{\ell_1,\ell_2} = \begin{cases} \frac{1}{||\boldsymbol{g}(\theta_{\ell_1})||_2} & \text{if } \ell_1 = \ell_2 \\ 0 & \text{otherwise} \end{cases} \tag{3.10}$$

Unlike the L_1 norm, which promulgated sparsity, the L_2 norm in the constraint promotes a smooth spatial distribution. The solution is analytical and, in the case of whitened data, has the following form:

$$\hat{\boldsymbol{s}}_{\text{wMNE}} = \boldsymbol{K}^\intercal\boldsymbol{G}^\intercal(\boldsymbol{GKK}^\intercal\boldsymbol{G}^\intercal + \lambda I_E)^{-1}\boldsymbol{x} \tag{3.11}$$

sLORETA

The sLORETA method is a topographic method, meaning that it does not aim at reconstructing the true source activity but rather a property map of of the latter. It relies on the wMNE solution to highlight the source standardized current density. The proposition (3.11) can be reformulated as:

$$\hat{\boldsymbol{s}}_{\text{wMNE}} = \boldsymbol{W}_{\text{wMNE}}^\intercal\,\boldsymbol{x} \;\text{ with }\; \boldsymbol{W}_{\text{wMNE}}^\intercal = \boldsymbol{K}^\intercal\boldsymbol{G}^\intercal(\boldsymbol{GKK}^\intercal\boldsymbol{G}^\intercal + \lambda I_E)^{-1} \tag{3.12}$$

where $\boldsymbol{W}_{\text{wMNE}}$ can be called the filter or the kernel of wMNE. The main idea behind sLORETA is that this filter brings some variability to the original source [73]. Indeed,

the resolution matrix generated by the product $\boldsymbol{W}_{\text{wMNE}}^{\top}\boldsymbol{G}$ leads to spread the sources since it corresponds to the estimated source variance. Thus, by dividing the filter by each variance induced by the resolution matrix, sLORETA reconstructs a standardized version of the sources:

$$m_{\text{sLORETA}} = \boldsymbol{W}_{\text{wMNE}}^{\top}\boldsymbol{K}'\boldsymbol{x} \tag{3.13}$$

where m symbolizes a SOI metric like NAI in beamforming (see section 2.2.3) and $\boldsymbol{K}' \in \mathbb{R}^{T\times T}$ is the diagonal weight matrix whose elements K'_{ℓ_1,ℓ_2} are defined as follows:

$$K'_{\ell_1,\ell_2} = \begin{cases} \frac{1}{|\boldsymbol{w}_{\text{wMNE}}(\theta_{\ell_1})^{\top}\boldsymbol{g}(\theta_{\ell_1})|} & \text{if } \ell_1 = \ell_2 \\ 0 & \text{otherwise} \end{cases} \tag{3.14}$$

Similar to the UNG and Stdr beamforming filters, sLORETA is a method which can be also used as post-processing.

cMEM

cMEM is a stochastic method [179] that treats the measurements, the sources, and the noises as realizations of stochastic random processes. In this context, the reconstruction of the sources consist in estimating the source distribution. The EEG system is rewritten as follows:

$$\boldsymbol{x} = \int \boldsymbol{G}\boldsymbol{s}dp(\boldsymbol{s}) \tag{3.15}$$

where $dp(\boldsymbol{s})$ is the source probability distribution. For the generic case of the Maximum Entropy on the Mean (MEM) framework, the goal is to select a distribution $dp(\boldsymbol{s})$ that provides the maximum missing information carried by the data or μ-entropy [180, 175, 176, 181]. The estimation of the source distribution $d\hat{p}(\boldsymbol{s})$ is achieved by optimizing a convex function. The primary interest of this method is that we can consider the distribution probabilities of the sources as the product of a probability density $f(\boldsymbol{s})$ and a reference distribution $dv(\boldsymbol{s})$. The latter promotes source properties assumed by the user. The contribution of the cMEM method is twofold. First, the parcellization of the brain into P patches [182, 181] allows us to consider $dv(\boldsymbol{s})$ as a joint probability:

$$dv(\boldsymbol{s}) = \prod_{p=1}^{P} dv_p(\boldsymbol{s}_p) \tag{3.16}$$

where $dv_p(\boldsymbol{s}_p)$ is the distribution for the parcel p. Second, the distribution of each parcel is defined as a Bernouilli-Gaussian distribution [183, 184]:

$$dv_p(\boldsymbol{s}_p) = (1-\rho_p)\delta_p + \rho_p\mathcal{N}(\mu_p, \boldsymbol{C}_p)(\boldsymbol{s}_p)ds_p \tag{3.17}$$

where ρ_p is the probability of activation of the parcel p. δ_p and $\mathcal{N}(\mu_p, \boldsymbol{C}_p)$ provide the parcel distribution: a Dirac distribution that shuts down the parcel or a Gaussian

distribution of mean μ_p and variance $\boldsymbol{C}_p$, respectively.

3.3.2 Preprocessing

Our goal is to compare the localization results of the mature methods wMNE, sLORETA and cMEM with the experimental SISSY and SABLE techniques. Then, we position ourselves as a standard user for the former and as an expert user for the latter. Hence, the pipelines proposed in Brainstorm remain untouched for the wMNE, sLORETA and cMEM methods, but the used variables are identical. In this part, we describe the pipeline used upstream of SISSY and SABLE. This pipeline is inspired by those implemented natively in Brainstorm for cMEM. The two exceptions are i) the exclusive use of whitening for wMNE and sLORETA which is essential for the estimation of the hyper-parameter [1] and ii) the normalization preprocessing for cMEM, SISSY and SABLE that relies on a preliminary calculation of the wMNE solution.

Defining an extended portion of signal On a long EEG recording session, the whole signal is not necessarily of interest. Indeed, some parts contain artifacts or do not present SOIs. In this case, the manipulator can extract some portions of the raw signal and concatenate them to create a mixed signal. For our application, the signals of interest are the interictal spikes. However, the direct junction of both portions can lead to jumps in the data, which is not aesthetic and even misleading. We propose an alternative solution called cross-fading. It is a blending method used in music to smooth the transition between two audio signals. Assuming that both signals have the same mean and variance [185, 186], the cross-fading operation progressively reduces the end of the outgoing signal. It increases the beginning of the ongoing signal without loss of the signals characteristics. This operation corresponds to the application of complementary weights on both signals, as shown in figure 3.3. The weights applied on both signals are determined using specific functions such as tangent or linear that respect some desire constraints (see [185] for details). We recommend taking at least a portion of signal of $0.5s$ to combine both signals for acceptable results. Preferably, this fusion should be done on a portion that has not any grapho-element of interest (i.e., avoiding the spike fusion). In practice, we observe slight changes in the values of the noise SCM as well as in the distribution of its eigenvalues.

Removing Corpus Callosum In chapter 1, we stated that the corpus callosum was only composed of nerve fibers which make the junction between both cerebral hemispheres. Those nerves are wrapped into a myelin sheath, an insulating material that increases the transmission of neural activity. Therefore, there is a very low probability that the sources which generate the EEG activity are at this location. By default, the segmentation method represents the corpus callosum because it allows us a closed

1. The whitening function is commented in the cMEM solver.

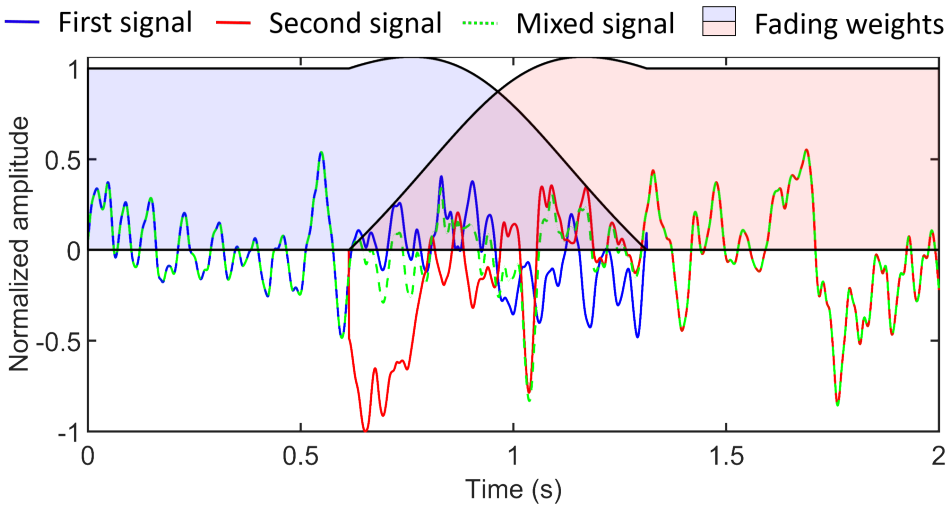


Figure 3.3: Example of signal mixing using cross-fading technique between two baseline signals on one electrode.

3D structure that can be easier to manipulate for specific algorithms.

However, for the studied inverse problem, it does not seems to be essential. On the contrary, since the dipoles constituting this anatomical structure are deep, they have an insignificant leadfield vector. The removal of such dipole allows us to limit the appearance of artifacts due to low leadfield norms. Moreover, we decrease the number of dipoles to estimate and thus increase the computational speed and the conditioning of inverse problem solutions.

Applying average reference Regarding the conditions of acquisition of real data, the EEG headset records the difference of potentials according to a reference electrode. As for the leadfield matrix, it is computed according to an infinite reference. Thus, we have to provide the same reference the EEG data and the leadfield matrix in order to have a valid mathematical system. The question of a good re-referencing is subject to discussion [187]; we opt for an averaged re-referencing like in Brainstorm. The average referencing corresponds to an artificial reference obtained by subtracting the time signal averaged over all sensors from the data of each channel. The following development shows how we achieve an identical referencing between the elements:

$$\boldsymbol{X}_{257} = \boldsymbol{G}_{\infty}\boldsymbol{S}$$

$$\begin{pmatrix} \boldsymbol{x}_1 - \boldsymbol{x}_{257} \\ \boldsymbol{x}_2 - \boldsymbol{x}_{257} \\ \vdots \\ \boldsymbol{x}_{257} - \boldsymbol{x}_{257} = 0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{g}_1 - \boldsymbol{g}_{\infty} \\ \boldsymbol{g}_2 - \boldsymbol{g}_{\infty} \\ \vdots \\ \boldsymbol{g}_{257} - \boldsymbol{g}_{\infty} \end{pmatrix} \boldsymbol{S}$$

$$\begin{pmatrix} \boldsymbol{x}_1 - \boldsymbol{x}_{257} \\ \boldsymbol{x}_2 - \boldsymbol{x}_{257} \\ \vdots \\ 0 \end{pmatrix} - \frac{1}{N}\sum_{n=1}^{N}(\boldsymbol{x}_n - \boldsymbol{x}_{257}) = \begin{pmatrix} \boldsymbol{g}_1 - \boldsymbol{g}_\infty \\ \boldsymbol{g}_2 - \boldsymbol{g}_\infty \\ \vdots \\ \boldsymbol{g}_{257} - \boldsymbol{g}_\infty \end{pmatrix} - \frac{1}{N}\sum_{n=1}^{N}(\boldsymbol{g}_n - \boldsymbol{g}_\infty)\boldsymbol{S}$$

$$\begin{pmatrix} \boldsymbol{x}_1 - \boldsymbol{x}_{257} - \sum_{n=1}^{N}\frac{1}{N}\boldsymbol{x}_n + \boldsymbol{x}_{257} \\ \boldsymbol{x}_2 - \boldsymbol{x}_{257} - \sum_{n=1}^{N}\frac{1}{N}\boldsymbol{x}_n + \boldsymbol{x}_{257} \\ \vdots \\ -\sum_{n=1}^{N}\frac{1}{N}\boldsymbol{x}_n + \boldsymbol{x}_{257} \end{pmatrix} = \begin{pmatrix} \boldsymbol{g}_1 - \boldsymbol{g}_\infty - \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{g}_n + \boldsymbol{g}_\infty \\ \boldsymbol{g}_2 - \boldsymbol{g}_\infty - \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{g}_n + \boldsymbol{g}_\infty \\ \vdots \\ \boldsymbol{g}_{257} - \boldsymbol{g}_\infty - \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{g}_n + \boldsymbol{g}_\infty \end{pmatrix}\boldsymbol{S}$$

$$\begin{pmatrix} \boldsymbol{x}_1 - \sum_{n=1}^{N}\frac{1}{N}\boldsymbol{x}_n \\ \boldsymbol{x}_2 - \sum_{n=1}^{N}\frac{1}{N}\boldsymbol{x}_n \\ \vdots \\ \boldsymbol{x}_{257} - \sum_{n=1}^{N}\frac{1}{N}\boldsymbol{x}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{g}_1 - \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{g}_n \\ \boldsymbol{g}_2 - \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{g}_n \\ \vdots \\ \boldsymbol{g}_{257} - \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{g}_n \end{pmatrix}\boldsymbol{S}$$

$$\boldsymbol{X}_{\text{avg}} = \boldsymbol{G}_{\text{avg}}\boldsymbol{S}$$

where $\boldsymbol{X}_{257}$, $\boldsymbol{G}_\infty$, $\boldsymbol{X}_{\text{avg}}$ and $\boldsymbol{G}_{\text{avg}}$ represent the data matrix referenced on the 257-th electrode, the leadfield matrix referenced to infinity and their equivalent re-referenced on the average, respectively. $\boldsymbol{x}_i$ and $\boldsymbol{g}_i$ are the data vector and the leadfield vector for the electrode i, respectively. It should be noted that the average reference is sensible to the dipole source orientation and the electrodes coverage [96]. There are potentially other solutions to improve this preprocessing method, but this is not the subject of this thesis.

Covariance matrix computation All methods require either the noise or signal covariance matrix. Therefore, we provide the best estimates of these matrices, i.e., the SMC of noise and signal refined with Ledoit's shrinkage method, as argued in Chapter 2.

Normalization by value In optimization, it is well-accepted that problems with normalized inputs are easier to solve. An example that often comes up by means of machine/deep learning is the gradient descent, in which the convergence rate of gradient descent is linear and depends on the condition number of the Hessian [188]. Hence, the problem is better bounded when the ratio of the largest and smallest eigenvalues of the Hessian is small. ADMM uses a gradient ascent to update the Lagrangian multipliers, which will influence the updates of the other variables. Therefore, their convergence rate is essential. Otherwise, one constraint that converges faster than the other may take the lead and result in an over-constraint solution. The normalization suggested in the cMEM pipeline proposes to scale the system as fol-

lows:

$$\frac{\boldsymbol{x}}{||\boldsymbol{G}||_\infty ||\boldsymbol{s}||_\infty} = \frac{\boldsymbol{G}}{||\boldsymbol{G}||_\infty}\frac{\boldsymbol{s}}{||\boldsymbol{s}||_\infty} + \frac{\boldsymbol{\nu}}{||\boldsymbol{G}||_\infty ||\boldsymbol{s}||_\infty} \tag{3.18}$$

$$\widetilde{\boldsymbol{x}} = \widetilde{\boldsymbol{G}}\widetilde{\boldsymbol{s}} + \widetilde{\boldsymbol{\nu}} \tag{3.19}$$

The maximum amplitude of the sources $||\boldsymbol{s}||_\infty$ can be estimated using the solution of the wMNE method. Indeed, this analytical solution method is not subject to convergence speed problems. Naturally, the inverse problem method will estimate a normalized source activity which we will have to denormalize through the post-processing.

New rank selection method As defined in section 3.1, SABLE is a deflation method. It means that each source is successively estimated and removed from the data until no more SOI is left. However, knowing the number of SOI in advance is challenging to evaluate and a few strategies have been proposed. In the cMEM pipeline, the number of SOI is considered to be equal to the number of eigenvalues whose sum constitutes 95% of the total sum of the eigenvalues. This criterion, called inertia, is efficient, but we observed that it overestimates the rank when the eigenvalues dispersion is narrow. Thus, we propose in this thesis a new measure inspired by the field of RMT (Random Matrix Theory), which fits our mathematical model and give a more precise thresholding value. Let us make the following assumptions :

A.3.1 The empirical spectral distribution of the eigenvalues of the Gaussian random SCM matrix is close to a deterministic distribution when $\frac{N}{T} \to c > 0$ for $N, T \to \infty$;

A.3.2 The deterministic distribution of the eigenvalues of the Gaussian random SCM matrix corresponds to the Marchenko-Pastur (M-P) distribution;

A.3.3 Eigenvalues of non-Gaussian activities lie outside the M-P distribution support;

A.3.4 The EEG data matrix in our application is a Gaussian random matrix perturbed by small rank matrices containing non-Gaussian activities.

Assumption **A.3.1** is one of the bases of RMT that starts in nuclear physics. The empirical spectral measure of the eigenvalues of a Gaussian random matrix is close to the semi-circle distribution [189]. For a covariance matrix, which is considered a particular case of the Wishart matrix [190, 191], the empirical distribution has been shown to follow the M-P distribution [192], hence the assumption **A.3.2**. As the mathematical demonstration is long, we accept those assumptions without evidence other than the vast bibliography. We recommend the book [193] with almost complete mathematical development to introduce such a matter.

Assumption **A.3.3** extends the first ones to the analysis of real signals and defines what we called the "spiked models" (unrelated to interictal spikes). In this model,

a small number of eigenvalues are outside the M-P support with a multiplicity 1 [194]. Those eigenvalues represent any activities that are not Gaussian and possess a significant higher-order statistic value (skewness, kurtosis).

Assumption **A.3.4** is compatible with our mathematical model (3.2) if we consider that $\boldsymbol{\nu}'$ is a random Gaussian vector [195, 196, 197]. It should be noted that some high eigenvalues coming from structured noises can appear. In the case of uncleaned data with muscle activity, we will stray further from the spiked model and assumption **A.3.3**. In that case, the problem is trickier [2] and will be reviewed in another work. For the moment, these assumptions are verified and the number of SOI can be obtained by counting the number of eigenvalues outside the M-P distribution support of the data SCM.

The M-P distribution p is defined as follows:

$$p(\Sigma_i^{\mathbf{C}_x}) = \begin{cases} \frac{1}{2\pi\Sigma_i^{\mathbf{C}_x} c\sigma^2}\sqrt{(b-\Sigma_i^{\mathbf{C}_x})(\Sigma_i^{\mathbf{C}_x}-a)} & \Sigma_i^{\mathbf{C}_x} \in [a,b] \\ 0 & \text{otherwise} \end{cases} \tag{3.20}$$

where $\Sigma_i^{\mathbf{C}_x}$ is the i-th eigenvalue of the data SCM and $a = \sigma^2(1-\sqrt{c})^2$ and $b = \sigma^2(1+\sqrt{c})^2$ are the limit of the support of the M-P distribution. c is the number of electrodes divided by the number of time points already defined for Ledoit's shrinkage method. Finally, all the eigenvalues superior to the value of b can be considered as SOI. The choice/estimation of the variance value (σ^2) will influence the threshold, we prefer to take the maximum variance of the data SCM (σ_X). In that case, we can assure to take only the most essential SOI. If we take the maximum variance of the noise SCM (σ_N), some structured noises will be considered. The application of this new selection method is illustrated in figure 3.4 in the case of one SOI and a few structured noises.

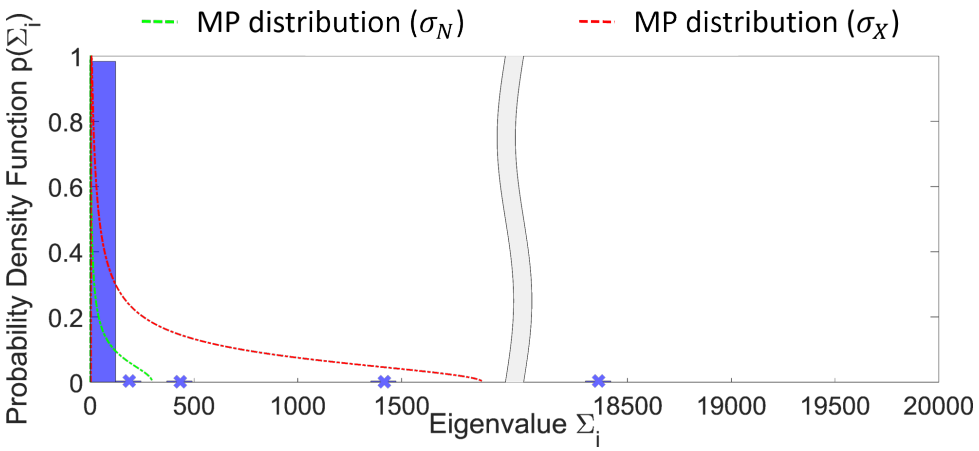


Figure 3.4: Eigenvalues repartition and Marchenko-Pastur distributions depending on the considered variance.

3.3.3 Post-processing

Once methods have been used, an estimate of the source activities or a statistical map is obtained. Post-processing methods are used to ensure a correct reconstruction of sources and highlight helpful information on the SOI.

Source reconstruction At the end of each method, we obtain:

— The kernel for wMNE and the standardized kernel for sLORETA;

— The reconstructed sources for SISSY;

— The probability distribution of the sources for cMEM;

— The filter and the spatial support for SABLE.

In order to reconstruct the sources for wMNE and sLORETA, we have to apply the kernel on the data using (3.12) and (3.13). For cMEM, we use the output of the method without modification. For SABLE, we reconstruct the source time series at each position using the following equation:

$$\hat{\tilde{\boldsymbol{s}}} = \sum_{k=1}^{K} \boldsymbol{\phi}_k \boldsymbol{w}_k^{\intercal} \tilde{\boldsymbol{x}} \tag{3.21}$$

Denormalization During the preprocessing step, input variables have been normalized for cMEM, SISSY and SABLE. For the former, the post-processing is already handled inside the BEst software. For SISSY and SABLE, the true values of the sources are restored by using the amplitude previously normalized (see previous section called normalization by value):

$$\hat{\boldsymbol{s}} = \hat{\tilde{\boldsymbol{s}}} ||\boldsymbol{s}||_{\infty} \tag{3.22}$$

3.4 Results

In this part, the performance of the SABLE algorithm is studied on realistic simulated epileptic interictal EEG signals compared to the wMNE, sLORETA, cMEM and SISSY methods. The simulation pipeline has been refined in order to simulate a more complex scenario with multiple ROIs with hyper synchronized activity.

3.4.1 Datasets and performance criteria

Generation of simulated data

In the same way as the previous simulations, we simulate a 257 channels EEG data (dense-EEG) using a realistic head model, representing the brain, the skull and the scalp, whose intermedium surfaces were extracted from the segmentation of a subject's anatomical 3D T1-weighted MRI. The source space consisted in a mesh made

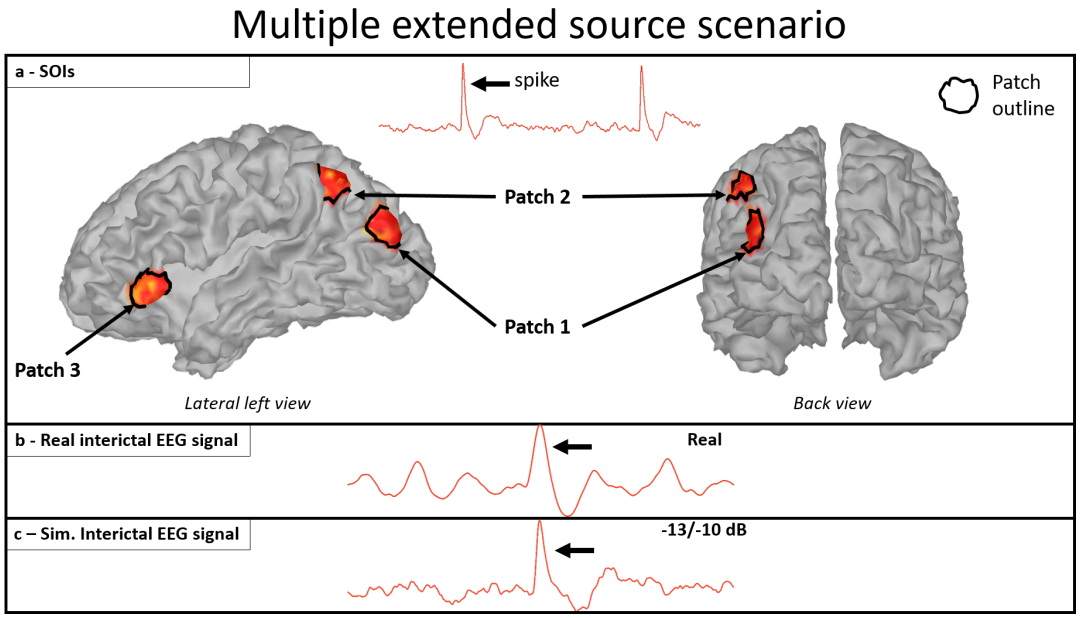


Figure 3.5: Data used in the simulations: a - Time courses and positions of the used extended sources, b - Example of real interictal EEG data, c - Examples of simulated interictal EEG data once the spikes have been extracted and averaged. When identifiable, interictal spikes are indicated using a black arrow.

of 15002 vertices before the removal of the corpus callosum obtained from the segmentation of the grey-white matter interface from the same anatomical T1-weighted 3D-MRI using the software BrainVisa [155]. The source space mesh has been refined using the iso2mesh algorithm [156] which ensures that each vertex covers the same brain surface approximately. The time course of each source, was provided by the output of a computational neural mass model [33, 31], in which parameters can be adjusted to generate either background-like activity or interictal-spikes. Using this realistic head model, the forward problem was then numerically calculated for each vertex using BEM [42, 40] to obtain the lead field vectors contained in the matrix $\boldsymbol{G}$ (257×15002), which gives the contribution of each dipole of the mesh at the level of 257 scalp electrode positions.

In this set of simulation, the epileptic source region is modeled by three epileptic patches including 50 grid dipoles each on the middle occipital gyrus, the inferior parietal gyrus and the triangular part of the inferior frontal gyrus. The remaining dipoles are distributed among 1000 scouts with the same background activity within a scout and uncorrelated activity between scouts. In addition, special scaling has been applied to obtain a scenario similar to the case of real data. Indeed, the samples were scaled so that i) the maximum value of the spike is approximately $100\mu V$ and ii) the maximum value of the background is at half the height of the spike. The corresponding SNR for such a scenario ranges from -13db to -10dB, validating the previous chapter assumption. From this simulated EEG data, each spike was selected and centered in a 1 second window and all widows were averaged to improve the SNR

like for real data. All presented results are averaged over 50 Monte Carlo runs. The hyper-parameters involved in SISSY and SABLE have been tuned by hand, including λ_1.

Performance criteria

The localization results are analyzed from a different angle using various qualitative and quantitative criteria for this last experiment. First, the overall efficiency will be assessed using the Receiver Operating Characteristic (ROC) curves and the Area Under Curve (AUC). After a reasonable choice of thresholding values, the analysis will be refined using criteria that will evaluate the similarity (Dice) [198, 199], the localization error (DLE) [200, 62] and the Spatial Dispersion (SD) [201, 202] of the estimated extended source. The remaining energy outside the estimate will also be considered using the Ratio of distant Spurious Activity (RSA) [203]. Finally, we propose a new criterion that informs on the energy distribution between the patches.

ROC and AUC The ROC is a graphical representation of the classifier efficiency as a function of the variation of a specified parameter. In this experiment, the parameter will be the coefficient of amplitude thresholding. For each coefficient, we compute the number of correctly identified source dipoles (called True Positive (TP)) and the number of source dipoles erroneously associated with the distributed sources (called False Positive (FP)). Once divided by the dimension of their respective domain, they form the True Positive Fraction (TPF) and False Positive Fraction (FPF) used to form the ROC:

$$\begin{aligned} \text{TPF} &= \frac{\text{TP}}{\#\Omega_{\mathcal{M}}} = \frac{L_{\mathcal{M}\cap\widehat{\mathcal{M}}}}{L_{\mathcal{M}}} \\ \text{FPF} &= \frac{\text{FP}}{\#\overline{\Omega_{\mathcal{M}}}} = \frac{L_{\widehat{\mathcal{M}}} - L_{\mathcal{M}\cap\widehat{\mathcal{M}}}}{L - L_{\mathcal{M}}} \end{aligned}$$

where $\#\overline{\Omega_{\mathcal{M}}}$ is the cardinal of the complement set of $\Omega_{\mathcal{M}}$ and where $L_{\mathcal{M}\cap\widehat{\mathcal{M}}}$ is the number of dipoles that are both in the original and estimated sets. The interpretation for such a criterion is qualitative as we are looking for the curve that is the closest to the top left corner of the axis. This corner corresponds to a perfect solution of 100% TPF for a 0% FPF. The ROC presented in the result are plotted for an FPF ranging from 0 to 20% for visibility.

The AUC [176] is a quantitative criterion calculated from a ROC. It can be view as a general accuracy ratio to evaluate the quality of an extended source localization method. Indeed, a high value of AUC underlines an earlier and faster increase of the ROC. It is estimated by using numerical integration on the whole ROC, although it is possible to concentrate the calculations on a smaller portion. It should be noted that a value of 80% is considered as a good detection accuracy [176, 181].

Thresholding choices For the other analyzing criteria, the threshold value must be fixed beforehand. Contrarily to the previous experiment (see section), it will be inappropriate to choose a fixed percentage of amplitude as the methods do not have the same priors. Indeed, the source estimates can vary considerably in size and form depending on those priors. For the sake of fairness, we propose two ways to delineate the estimated solution:

1. By taking the thresholding value that corresponds to the best ratio TPF/FPF in the ROC. It comes down to find the point with the minimum distance to the top-left corner;
2. By taking a solution composed with the 150 dipoles with the highest energy. This threshold will show which method is the fastest to locate sources.

With these two ways of thresholding, we can consider the equity of treatment of each localization method. An equivalent TPF/FPF will be provided for both corresponding threshold values alongside the other criterion.

Dice similarity coefficient (Dice) The Dice coefficient, also called Sorensen-Dice similarity coefficient [198, 199], is a complementary criterion of the TPF and FPF criteria. It is a statistic which evaluates the degree of equivalence between two binary images, in our case, between original and estimated extended sources. Unlike both other statistics, Dice considers the number of sources which are not localized by the source localization method, i.e. False Negative (FN) value.

$$\text{Dice} = \frac{2L_{\mathcal{M}\cap\widehat{\mathcal{M}}}}{L_{\mathcal{M}} + L_{\widehat{\mathcal{M}}}} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}$$

A high value of Dice underlines a good localization (TP) without overestimation (FP) or underestimation (FN).

Dipole Localization Error (DLE) Identical to the one used in the previous experiment, the DLE is used to quantify the source localization accuracy . Let us recall its definition:

$$\text{DLE} = \frac{1}{2L_{\mathcal{M}}} \sum_{m\in\Omega_{\mathcal{M}}} \min_{k\in\Omega_{\widehat{\mathcal{M}}}} ||\boldsymbol{\theta}_m - \boldsymbol{\theta}_k||_2 + \frac{1}{2L_{\widehat{\mathcal{M}}}} \sum_{m\in\Omega_{\widehat{\mathcal{M}}}} \min_{k\in\Omega_{\mathcal{M}}} ||\boldsymbol{\theta}_m - \boldsymbol{\theta}_k||_2$$

A low value of DLE signifies that the estimated indexes are inside or near the original extended source.

Spatial Dispersion (SD) The SD criterion [201, 202] measures the spatial spread of the estimated source distribution around the original source location. Its definition

is the following:

$$\text{SD} = \sqrt{\frac{\sum_{m \in \Omega_{\widehat{\mathcal{M}}}} d_m^2 \widehat{s}(\theta_m)^2}{\sum_{m \in \Omega_{\widehat{\mathcal{M}}}} \widehat{s}(\theta_m)^2}} \quad \text{with} \quad d_m = \min_{k \in \Omega_{\mathcal{M}}} ||\boldsymbol{\theta}_m - \boldsymbol{\theta}_k||_2$$

where d_m is the minimal distance between the estimated source and the set of all original sources. A low value of SD means that the dipoles with high energy are inside the original extended source.

Ratio of distant Spurious Activity (RSA) The RSA criterion [203] provides information on any high activity located outside the estimated sources:

$$\text{RSA} = \frac{\sum_{m \notin \Omega_{\widehat{\mathcal{M}}}} \widehat{s}(\theta_m)^2}{\sum_{\ell=1}^{L} \widehat{s}(\theta_\ell)^2}$$

A low value of RSA means that there is no more energy outside the estimated patches. It should be noted that all the above criteria depend on the size of the estimated sets of indexes, hence on the threshold method.

Patch amplitude ratio (Par) The patch amplitude ratio evaluates the energy distribution between the different distant ROIs. It provides information on the equality of treatment of each ROI by the method and the existence of a localization bias. The former ensures that a manipulator may not overlook potential sources due to its low amplitude caused by a method defect. For the latter, there is a localization bias if the 100% value is missing. The metric is defined as the follows:

$$\text{Par}_k = \frac{|\widehat{s}(\theta_m)|}{||\widehat{\boldsymbol{s}}||_\infty} \text{ with } m \in \Omega_{\mathcal{M}_k} \tag{3.23}$$

where $\Omega_{\mathcal{M}_k}$ is the set of indexes inside the k-th patch. A high value in each patch means that there is a common treatment. We consider that a value under 50% underlines a difference of treatment.

3.4.2 Simulated data

The behavior of the five chosen BSI-EEG methods is evaluated in terms of source localization with a realistic SNR value. To this aim, multiple quantitative criteria have been proposed illustrating different aspects of the results. Two thresholding methods have been selected for the measures requiring a thresholded solution (see the previous section for details). Criterion results have been synthesized into figure 3.6 and tables 3.2 to 3.5. Examples of thresholded results have been presented in figures 3.7 and 3.8, and in appendix B.3 with a different angle.

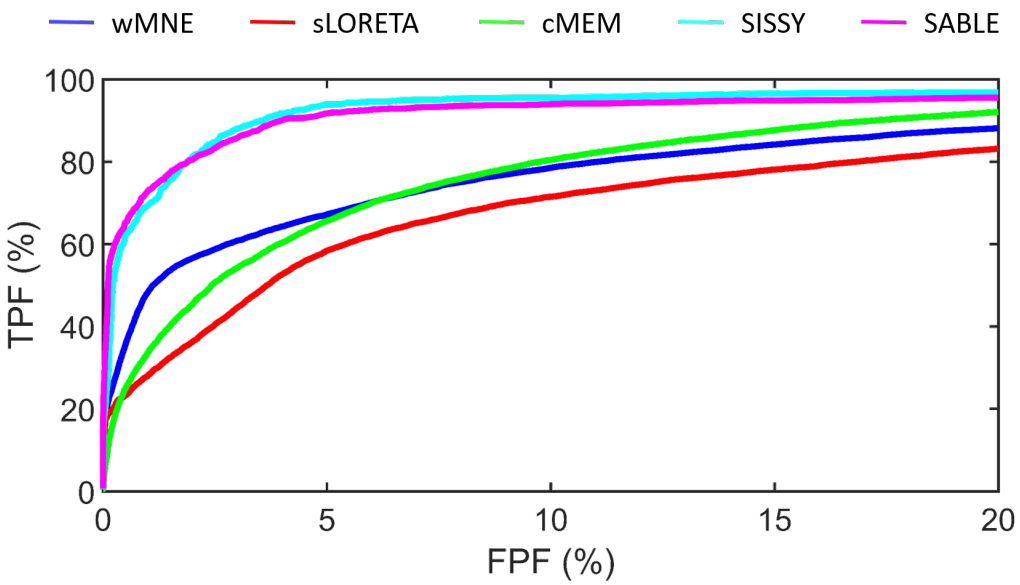


Figure 3.6: The ROC criterion at the output of 5 BSI methods.

Broad analysis using ROC and AUC Figure 3.6 presents the ROC criterion of all the five BSI-EEG methods for FPF values ranging from 0 to 20%. The ROC depicts an apparent distinctive behavior between the traditional methods and the experimental ones. Indeed, for FPF values under 20%, the ROC criterion at the output of wMNE, sLORETA and cMEM is systematically lower than that of SISSY and SABLE. The ROC criterion at the output of SABLE has the highest start of all methods and is the first to reach the TPF value of 80%. It is caught up soon after by the one given by SISSY method, which will remains above for this point on. The sLORETA method does not outperform wMNE in this kind of scenario, probably because there are limited deep sources. The cMEM performance improves after an FPF value of 5%.

Table 3.2: The AUC criterion at the output of 5 BSI methods

	wMNE	sLORETA	cMEM	SISSY	SABLE
AUC	91.48%	88.75%	93.79%	97.20%	96.70%

Table 3.2 quantitatively summarizes the observations derived from figure 3.6. sLORETA shows the worst localization performance but still has an AUC of 80%, which means that it remains a good localization method according to the hypotheses formulated in the previous section. sLORETA is followed by wMNE and cMEM with an AUC of approximately 91.5% and 94%. SISSY and SABLE present astonishing results close to 97%. The slight difference of 0.5% between these two methods can be partially interpreted by means of figure 3.6, where a slight difference in efficiency appears at an FPF value of 5% and remains until 100%.

As we can understand from those results, SISSY is the best method for overall lo-

Table 3.3: Comparison of 5 BSI methods through 6 performance criteria using the best threshold

	wMNE	sLORETA	cMEM	SISSY	SABLE
TPF	84.8%	82.1%	88.6%	**95.3%**	**93.6%**
FPF	14.3%	17.5%	13.9%	**4.8%**	**4.7%**
Dice	11.8%	9.5%	**12.6%**	**31.7%**	**31.7%**
DLE	9.36	10.79	8.90	**5.40**	**6.28**
SD	12.79	17.5	9.69	**5.75**	**5.26**
RSA	0.09	0.12	**0.01**	**0.05**	**0.05**

calization performances for this scenario. However, it is followed very closely by SABLE, of which it is partially at the origin. Let us recall that SABLE presented much worse outcomes than SISSY for an SNR value of -10dB in the preliminary results (see section 3.2.3) due, as we claimed to the beamforming term. We can observe that this is no longer the case by means of a fine-tuning of the hyper-parameter λ_1. Moreover, SABLE overcomes SISSY for low FPF values, which may explain its better performance for specific fixed thresholding values. For such a scenario, methods implemented in Brainstorm, namely wMNE, sLORETA and cMEM, present decent overall localization results. cMEM is the best performing of the three methods and sLORETA the worst. Now, let us fix the thresholding value to that providing the best ratio TPF/FPF, using the ROC criterion to refine our analysis.

Quantitative results with the best thresholding Table 3.3 summarizes the performance achieved with the different source imaging algorithms in terms of TPF, FPF, similarity, DLE, SD and RSA for the optimal thresholding value. All values show that SABLE and SISSY still outperform the other distributed source localization methods for the considered scenario.

Regarding the coefficient of index proportion, table 3.3 reveals that the three traditional methods are grouped with an FPF value approximately equal to 15%. However, cMEM can be considered as the best traditional method with a TPF value of 88.5%. On the other hand, SISSY and SABLE have a sensibility greater than 90% for only a specificity of 5%. Those observations corroborate with the ones presented in the previous session. If we consider the similarity coefficient or the Dice criterion, which is more reliable since, in addition to encompassing the FP and TP values, it also considers the FN value. The performance values for traditional methods are below 15%, while SISSY and SABLE exhibit a similarity value of 31.7%.

Regarding the other quantitative criteria, cMEM has the best DLE, SD and RSA compared to the traditional methods. The main thing to notice is its very low value of RSA, which means that this method presents very focused results, probably due to

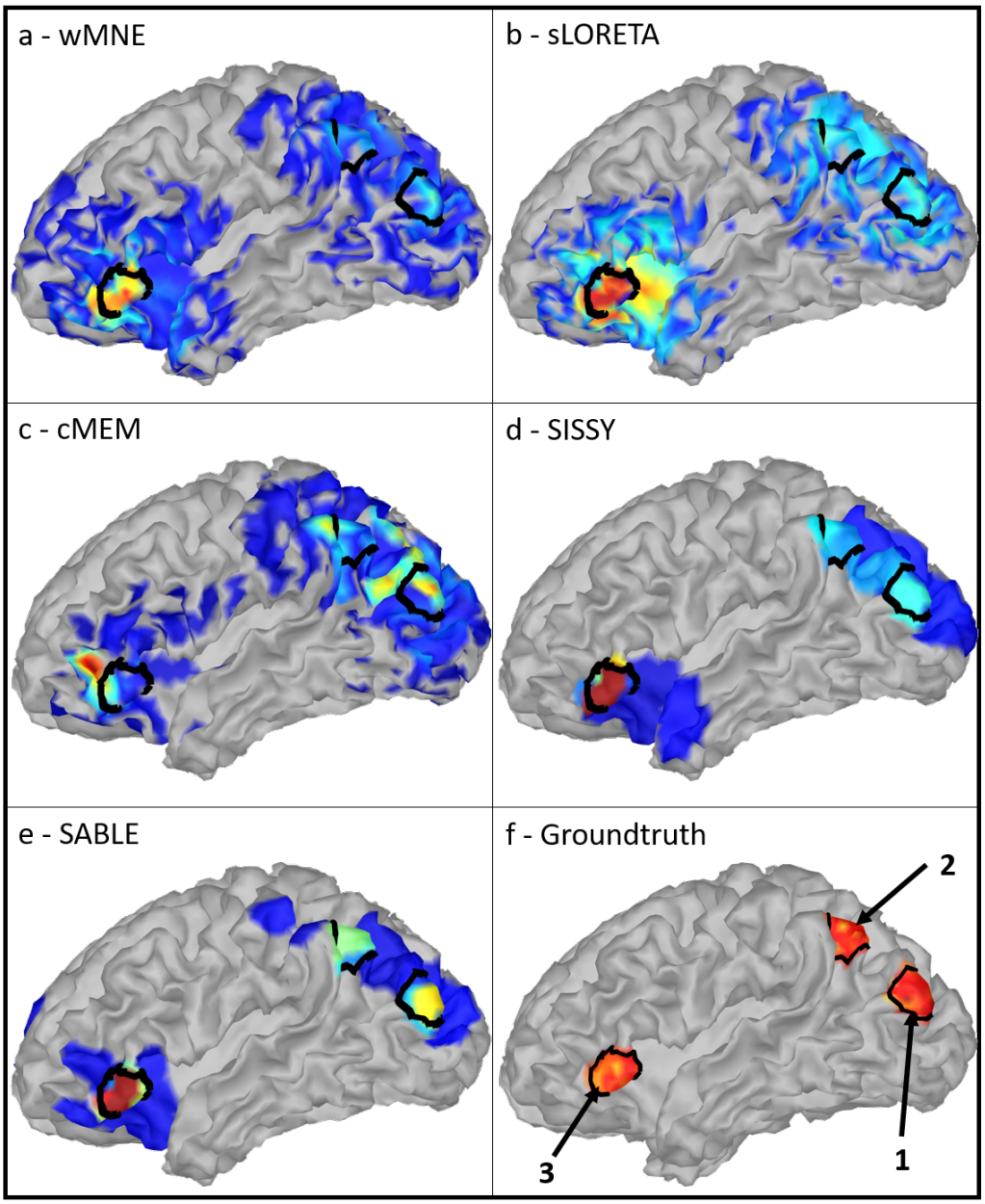


Figure 3.7: Source localization results using the best threshold at the output of wMNE, sLORETA, cMEM, SISSY and SABLE for an SNR value of approximately -10dB (Sagittal left). The patches outline is represented by a plain black line.

how it assigns values to parcels. However, we must recall that RSA is a criterion influenced by the number of indexes taken into account in the thresholded solution. Thus, it cannot be directly computed at the output of SABLE and SISSY since the latter have about half the number of indexes that cMEM (see section 3.4.1). On the other hand, sLORETA has a high RSA value due to a general increase in the energy of all estimated dipoles induced by the normalization. For the experimental methods, SABLE shows an SD value inferior to that given by SISSY for the same RSA value, which suggests a more focused solution for the former. SISSY has a better DLE value than

Table 3.4: Comparison of 5 BSI methods through 6 performance criteria using the 150 first indexes threshold

	wMNE	sLORETA	cMEM	SISSY	SABLE
TPF	40.9%	26.8%	31%	**60.1%**	**64%**
FPF	0.67%	0.82%	0.78%	**0.45%**	**0.41%**
Dice	40.9%	26.9%	31%	**60.1%**	**63.9%**
DLE	3.96	7.28	5.66	**2.08**	**2.22**
SD	5.65	8.71	6.45	**2.14**	**2.02**
RSA	0.51	0.72	0.31	**0.23**	**0.14**

SABLE which can be explained by the fact that SABLE will preferably locate sources further away than extending the borders of already found sources. These observations can be supported by a direct observation of the threshold solutions provided in figure 3.7.

Regarding the source localization solutions presented in figure 3.7, all methods except cMEM localized the third patches correctly. We can observe a high amplitude dipole located outside the patch underlying a localization bias for the latter. For the first and second patches, wMNE and sLORETA cannot differentiate both patches as they consider them like one large patch. This behavior is a problem known from the L_2 norm constraint method. For cMEM, dipoles with relatively high amplitude (about half of the max) are located inside both patches. However, we can also observe other high activities outside the patch outlines which would be considered as spurious distant sources for another threshold. For SISSY and SABLE, we notice that both of them assign relatively strong energy inside both patches. However, it is more visible for SABLE than SISSY. We can also confirm the assertion made in the previous paragraph that SISSY will preferably expand the already localized sources and SABLE will seek new sources.

Quantitative results on 150 first indexes Table 3.4 summarizes the performance achieved by each source imaging method for a thresholded solution containing only the 150 dipoles with the highest amplitudes. Since the number of estimated dipoles is now fixed, criteria depending on this value are now identical. Globally, we can see that the SABLE method gives the best results in terms of coefficient (PSF, FPF and Dice). Indeed, as we claimed by looking at the ROC results, SABLE is the quickest to locate sources because it is more focused than SISSY. Moreover, it is also more concentrated as it presents the lowest SD and RSA values. Contrary to the results of the previous methods, wMNE presents the best results among the traditional methods like the ROC results let it suggested. It should be noted that the RSA and the Dice

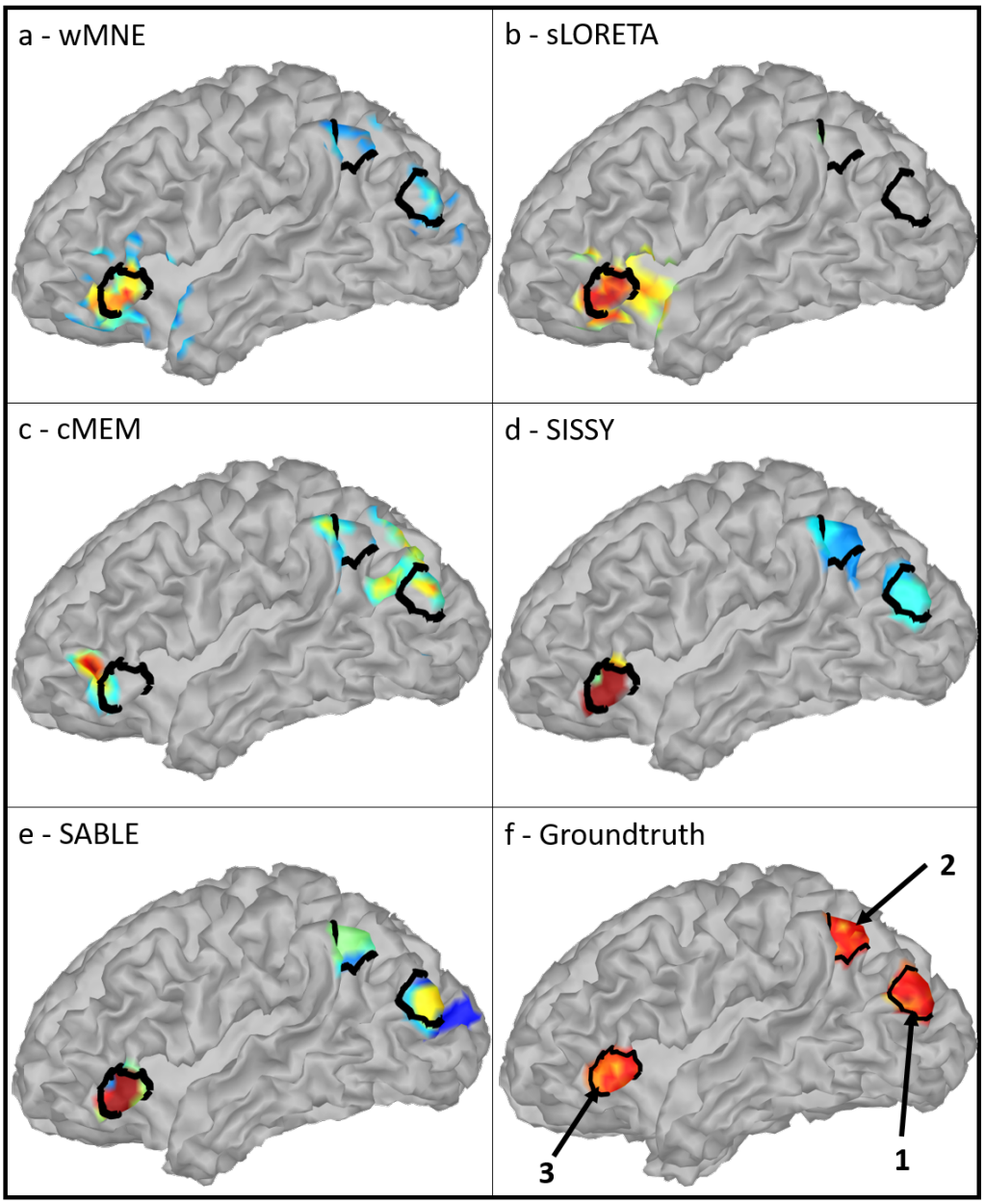


Figure 3.8: Source localization results thresholded with the first 150 indexes at the output of wMNE, sLORETA, cMEM, SISSY and SABLE for an SNR value of approximately -10dB (Sagittal left). The patches outline is represented by a plain black line.

values are superior to those mentioned in the previous table due to the size of the estimated index.

Regarding the source localization solutions presented in figure B.7, all methods except sLORETA present activities in the three patches. The sLORETA solution exhibits a wide activity concentrated around the patch alternative thresholding. cMEM presents a localization bias for such a scenario. The wMNE solution presents diffuse activity outside the patches. SABLE shows a more focus solution than SISSY has the coloration indicates.

Patch amplitude ratio analysis The table 3.5 quantitatively summarizes the amplitude ratio for each method. We can observe that cMEM never achieves 100%, which means that the maximum amplitude is located outside of the three patches. Thus it confirms a localization bias for a certain number of MC runs. Otherwise, its energy distribution between each patch is pretty good as it does not fall under 30%. Besides cMEM, all other methods attribute the maximum of energy to the third patch. Moreover, SABLE is the one which distributes its energy more uniformly. Indeed, we observe a difference of 45% between patch 3 and the least energetic patch 1 while it is 65% for the wMNE and SISSY methods and 55% for sLORETA. From these observations, we can conclude that SABLE will be more conducive to highlight distant correlated sources. For sLORETA, the standardized version of wMNE, we can see a net amelioration in terms of energetic distribution with respect to the original one.

Table 3.5: Energy ratio of each patch

	wMNE	sLORETA	cMEM	SISSY	SABLE
Par_1	44.9%	45.1%	81%	35.5%	71%
Par_2	33.9%	55%	77.1%	45.2%	54.8%
Par_3	100%	100%	52.29%	100%	100%

Discussion on simulated data In this section, we assess the behavior of three traditional methods and two experimental methods in terms of localization results for a scenario with three strongly correlated extended sources: one isolated and two very close to each other. A multitude of quantitative criteria was used to study the effects from all perspectives, and a new one was introduced. From the results, we can say that, for this kind of scenario, the considered traditional methods are not efficient enough. Indeed, wMNE and sLORETA consider both nearby sources as one large extended source and cMEM presents a localization bias which is not insignificant. SISSY provides a good solution but may have some flaws when the energy is dispersed between several sources. Indeed, this method will give more weight to a single extended source which can be deleterious during a thresholding operation. The experiment shows that the SABLE method overcomes this problem by presenting more focus activity inside each patch due to the beamforming constraint. However, we have to pay attention to the hyper-parameter setting that makes the weight between both cost functions which compose it as shown in the previous experiments. Considering those results, we propose to process and compare the results of the five methods on real data using the same pipeline. The subject studied is the same as the one we presented in the previous chapter since now we explore the hypothesis that a new source model could improve the localization and reconstruction results.

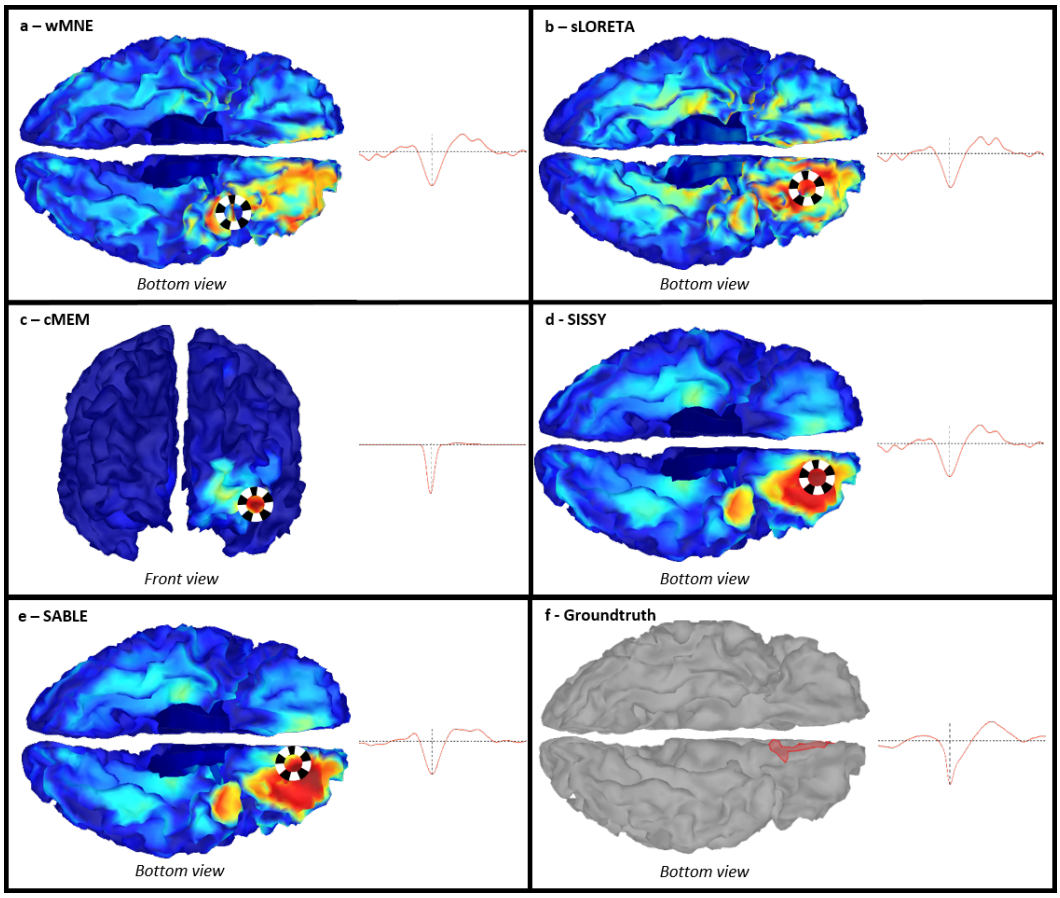


Figure 3.9: Source localization results at the output of wMNE (a), sLORETA (b), cMEM (c), SISSY (d) and SABLE (e) on real data. The estimated ground-truth is presented alongside in (f). The brain view has been carefully oriented to display the areas of maximum activity located in the center of the black and white ring. On the side of the results is the estimated SOI activity. It is shown along a vertical and horizontal axis corresponding to the time index of the interictal spike maximum and the reference, respectively.

3.4.3 Application on real data

In this experiment, we propose to measure qualitatively the efficiency of each method efficiency on real hr-EEG data acquired from the patient studied in chapter 2. Lets us start by recalling its characteristics before presenting the results.

Data acquisition

A 256-channel hr-EEG was recorded for one hour, at 1000 Hz following the procedure approved by the National Ethics Committee for the Protection of Persons (CPP, agreement number 2012-A01227-36). The patient gave his written informed consent to participate in this study. This scalp recording revealed interictal spike activity at the left frontal basal electrode AFP9h, located above the subject left eyebrow. From this interictal epileptic activity, 8 spikes were visually selected away from the occurrence of any artefacts (muscle activity, blood pulsation, eye blinks). Each spike was centered in a 1 second window and all 8 windows were averaged to improve the SNR

ratio. As part of the presurgical evaluation, Stereotactic EEG (SEEG) recordings were also performed using 8 intracerebral electrodes (10 ± 18 lead, length: 2 mm, diameter: 0.8 mm, 1.5 mm apart, sampling frequency: 512Hz) placed intracranially according to the Talairachs stereotactic method in the left frontal and temporal regions. The positioning of the electrodes was determined from an available non-invasive information and hypotheses about the localization of his EZ. The exact 3D coordinates of each electrode contact were determined after co-registering the CT scan showing the intracerebral leads onto the structural MRI image using a 6-parameters rigid-body transformation [158, 159]. Segmented anatomical data obtained with Freesurfer from the patient 3D-T1 MRI as well as the location of depth electrode contacts were imported within Brainstorm [40]. The areas of spike maximal activity and of spike propagation were delineated on the patient cortical mesh according to conclusions obtained after visual analysis of intracerebral interictal activity by the expert neurologist. Thanks to the presurgical evaluation, it has been determined that the observed EEG signal results from the activity of one source. The new rank selection method presented in section 3.3.2 confirmed this assumption. The hyper-parameters has been tuned by hand.

Qualitative results

Figure 3.9 displays the localization and the SOI reconstruction results obtained by the five methods. It can be seen that the SABLE method is the one which presents its maximum activity and which is the closest to the EZ. Indeed, although a little higher than the suborbital sulcus, the white striped circle can be seen on the other side of the gyrus rectus. Regarding the other methods, sLORETA and SISSY present a solution a little further from the estimated EZ in the orbital gyrus/sulcus. For the former, we can see a discontinuous band of high intensity which goes towards the interior of the brain. In contrast, we can see its activity propagating in the orbital sulcus like SABLE for the latter. For wMNE and cMEM, the maximum activity is located on the temporal lobe and the fronto-marginal gyrus, respectively. Their solution is therefore found too far for the EZ. For wMNE, we observe the same wideband as for sLORETA. Concerning the general dispersion of the activities, we find a similar attitude as for patches 2 and 3 of the simulations: wMNE and sLORETA have a very spread solution while SISSY and SABLE are very focused. Furthermore, cMEM seems to perform the localization with a localization bias but a focus solution. Regarding the SOI time course reconstruction, all methods faithfully reproduce the activity observed in sEEG. Contrarily to the one observed with the traditional beamforming, none of them shows strong oscillations in their reconstruction.

Figure 3.10 displays the localization results obtained by the five methods with a different view inside the brain. This view is the most suited for observing the estimated EZ as stated in the previous chapter. As we can see, all methods show some activ-

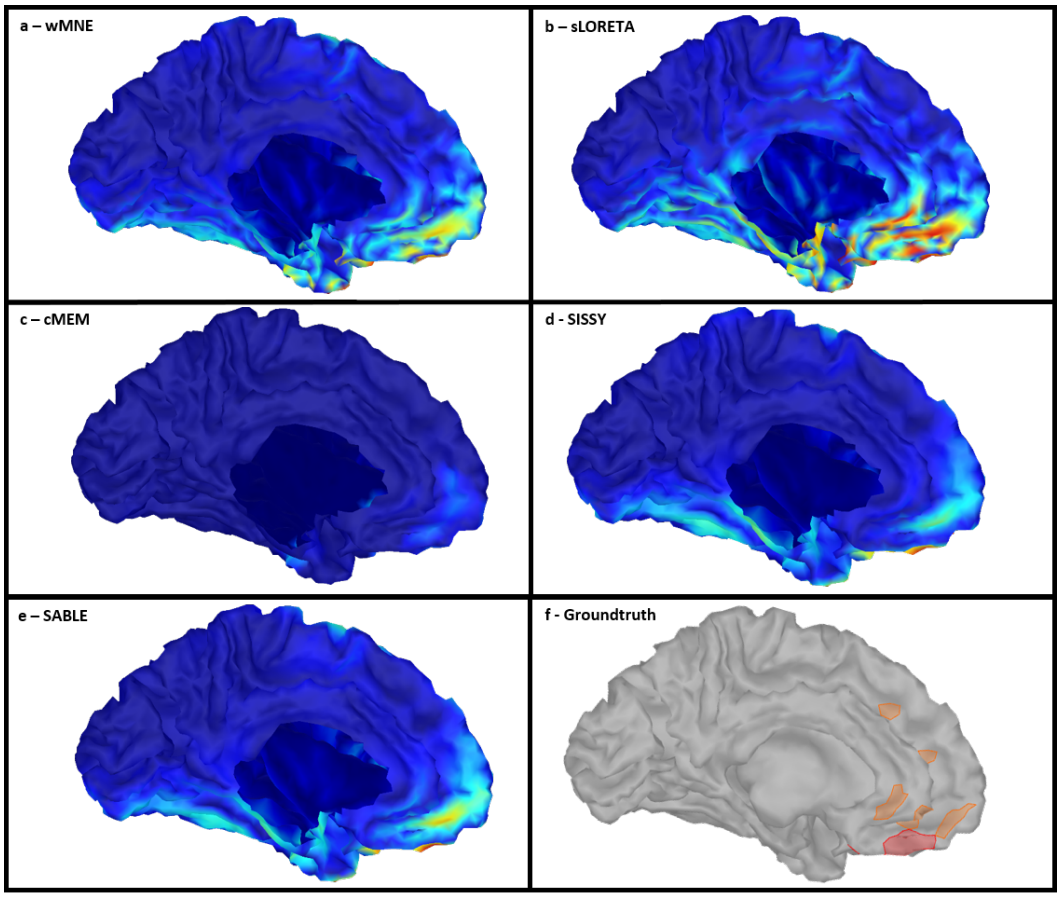


Figure 3.10: Source localization results at the output of wMNE (a), sLORETA (b), cMEM (c), SISSY (d) and SABLE (e) on real data. The estimated ground-truth is presented alongside in (f) with, in red, the estimated Epileptic Zone (EZ) and in orange the spike propagation. The brain view has been oriented to the parasagittal view of the left hemisphere.

ity inside the cingular gyrus. For wMNE and sLORETA, it is again a spread activity that diffuses into the EZ and the spike propagation area. On the contrary, SABLE displays a very focused activity near a considered spike propagation zone. For SISSY and cMEM, the activity is shallow but still present. From these two sets of figures, we would be tempted to assume a 1 SOI for a 3 ROI scenario: one in the temporal lobe, one around the orbital gyrus and one in the anterior cingular gyrus. In that case, we can use the Par values in table 3.5 previously determined in the simulation to threshold the solution.

Figure 3.11 displays the thresholded localization results obtained by the five methods with both different views. Although we are not in the same configuration as the simulations, we can make similar observations. Indeed, SABLE shows three zones of high energies contrarily to SISSY for which it is more diffuse. Moreover, the third patch in SISSY can be mistaken for a background activity as the thresholded solution contained multiple zones with similar energies. cMEM displays only one zone even after thresholding which is always far from one of the identified ROIs in sEEG.

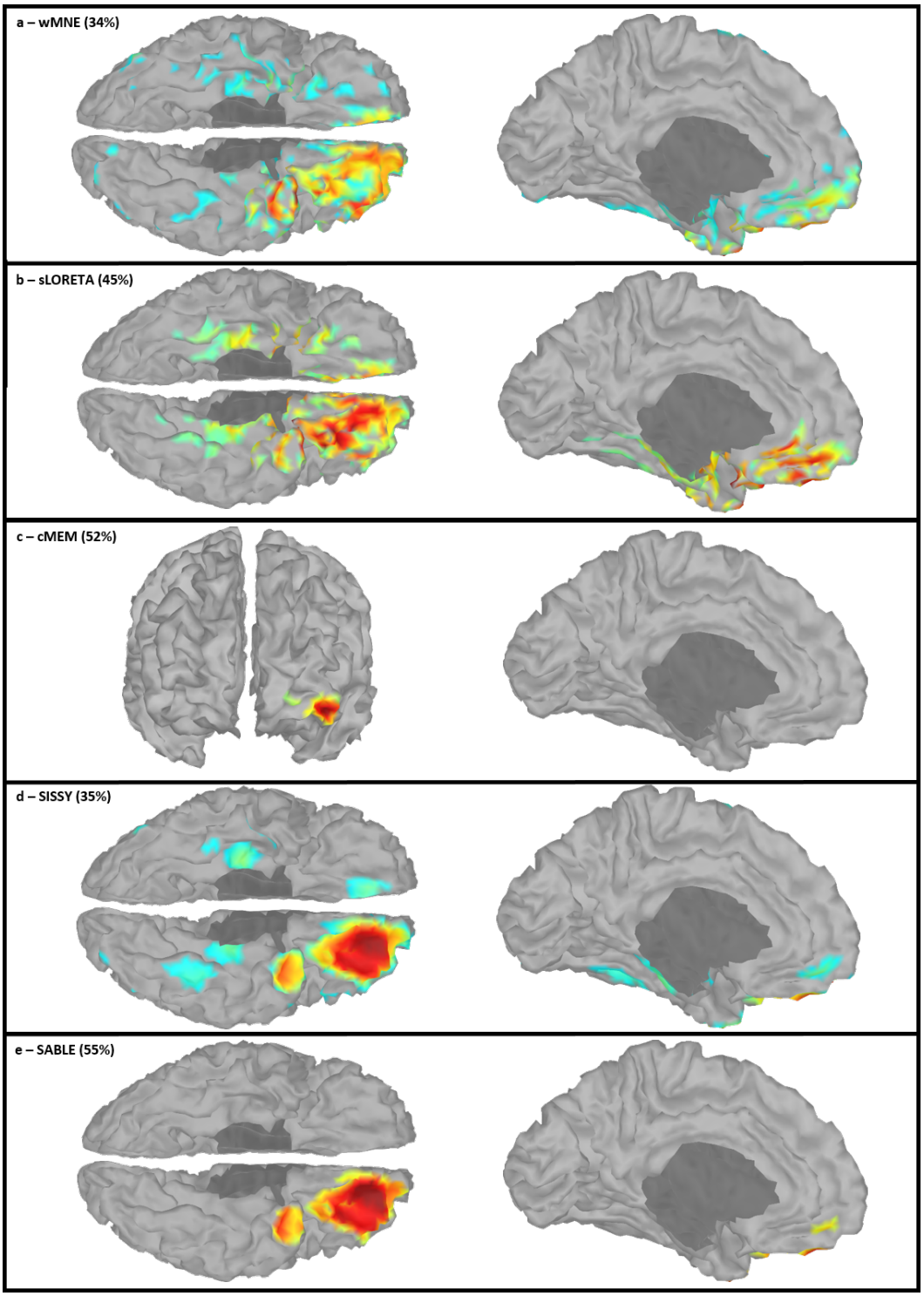


Figure 3.11: Thresholded source localization results at the output of wMNE (a), sLORETA (b), cMEM (c), SISSY (d) and SABLE (e) on real data. The threshold value applied correspond to those estimated with the Patch amplitude ratio criterion as three ROIs can be identified on the raw solutions.

wMNE and sLORETA are the only ones to show a high but scattered activity on the estimated EZ. One explanation could be that they exploit information of the depth of the sources using a weighted matrix. The same operation can be associated with

the array gain beamforming and may be adapted for SABLE and SISSY in future work. Even so, the preliminary results are deemed satisfactory and exciting for a new method and should be deployed to a more large set of real data to support these preliminary claims.

3.5 Conclusion

In this chapter, the main motivation is to propose and assess a new beamforming method for the distributed source model. Indeed, as stated in chapter 2, beamforming is used mainly for the equivalent source model partly on account of their sensitivity to correlated sources. However, the distributed source model is more suitable for epilepsy and, despite our intensive review, no beamforming method seems appropriate for this model. Another critical question addressed here concerns the practical application and the ability of this new distributed beamforming method to locate and reconstruct SOIs. Indeed, as a novel method, a pipeline must be provided in order to bring it out of the experimental phase. Moreover, its performance must be evaluated against other BSI methods to identify its benefits and drawbacks. For these reasons, we first propose the new beamforming method named SABLE (Source Imaging based on Structured Sparsity) by developing its theoretical aspect and by presenting a set of preliminary results. After that, an adequate pipeline is proposed to be used in the BSI-EEG context. More precisely, this pipeline is inspired by those already implemented in the software Brainstorm with some additional novelty. Finally, a quantitative/qualitative study comparing wMNE, sLORETA, cMEM, SISSY and SABLE is performed on simulated and real data.

In this thesis, we reviewed all beamforming methods in chapter 2 and identified the best beamforming pipeline for the equivalent source model. Despite promising simulation results, the application on real data shows its limitation attributable to the use of a source model not adapted for epilepsy. Convinced that these methods can improve confidence in the estimated source, we propose a new beamforming method suitable for the distributed source model. This novel method named SABLE merges both the reconstruction and the localization steps into one problem. SABLE proved to be more efficient than the traditional UG beamforming method and the regularized least-squares method SISSY methods with a high SNR value. However, like SISSY, SABLE requires a fine tuning of the hyper-parameter, making the trade-off between both cost functions for negative SNR values. Otherwise, the beamforming part takes the lead and results in solutions worst than SISSY. Compared to other methods used in the BSI-EEG field, SABLE depicts a better localization solution in the case of an SOI scenario distributed among several ROIs. Indeed, While other methods will present either a spatially overextended area or several low-energy areas, SABLE will present a set of relatively high energy regions. Although the SISSY method is more accurate

in estimating the spatial extent at the expense of energy distribution, SABLE provides a better trade-off between the estimate of amplitude and and that of spatial extension. Compared with the beamforming methods (see chapter 2) for real data application, it gives a faithful reconstruction of the source activity, and the localization results have been judged satisfactory by a neurologist expert.

Even if this new method gives auspicious results both in simulated data, the results on real data are not precisely consistent with the information provided by SEEG. However, it is worth remembering that the defined EZ is only an estimate and that the true ground truth in the real case is unknown. Even if the results are deemed satisfactory, future work will include a diversification of simulation scenarios and real data sets to rule on SABLE true capabilities. Moreover, with a few more adjustments, the SABLE pipeline can also be automated and thus made available to a more significant number of users. These adjustments include an effective estimation of the hyperparameters scheme and a more accurate rank estimation criterion based on random matrix theory. These are both last limiting factors for the departure of SABLE from the experimental setting to a clinical use.

Conclusion and perspectives

The present work had as main objective to study, the taxonomy, the assessment and the extension fo the beamforming methods to solve the inverse problem in BSI-EEG for epilepsy. Four main contributions can be highlighted from this work.

The first contribution is to perform a complete taxonomy of all existing beamforming methods, summarized by means of three tables whose combined content makes it possible to cover all practices in BSI-EEG. Furthermore, new approaches coming from the array processing field have also been introduced to supplement existing ones. This work has been completed with a comparative analysis performed on realistic epileptic EEG simulated data in order to identify the best pipeline. It shows that Ledoit's shrinkage method, coming from the RMT (Random Matrix Theory), is the best preprocessing for beamforming. The AOLS beamformer, which uses another form of shrinkage, depicts the best results for the one-source and uncorrelated sources scenarios. The LCMV beamformer could be considered as the most efficient for correlated sources, but we chose not to pursue its development further. Indeed, it features a localization bias which is very challenging to overcome and leads to unreliable results for real data. Finally, the newly introduced EEG similarity metric proves to be the best post-processing method. When applied to real data, this pipeline could efficiently locate the epileptogenic zone. Still, the reconstructed signals present some imperfections as they are contaminated by other source activities.

The second contribution is the proposition of a new regularized method called Temporal-Smoothing-based beamforming to improve the source reconstruction on real data. The mathematical development proves that it can be easily implemented as it only impacts the data covariance matrix. This constraint not only obliges the distribution of the estimated sources to be smooth but also improves the conditioning of the covariance matrix. As a result, the source estimation results of Temporal-Smoothing-based beamforming methods display a significant improvement on both simulated and real data.

The third contribution is the proposal of a new rank selection method based on RMT.

Indeed, as claimed throughout this thesis, knowing the number of SOI from the observation is challenging to obtain. In the cMEM pipeline, a strategy called inertia based on eigenvalues has been suggested to estimate this parameter (see section 3.3.2). However, we did not find any reference on such a method in the literature. On the other hand, the RMT field proposes a solution to characterize the support of the eigenvalues corresponding to random Gaussian activities. More precisely, it allows us to estimate the support boundaries values. The rank of the SOI corresponds to the number of eigenvalues over that range as they correspond to structured noise activities or epileptic activities. A judicious choice of variance allows us to distinguish between noise and epileptic activity as a first approximation, but other more complex methods will allow a more refined categorization.

The fourth contribution is the proposal of a new method that extends the traditional beamforming methods to the distributed source model suitable for epilepsy. Indeed, classical beamforming methods are more generally applied to the equivalent source model to localize and reconstruct uncorrelated source activities. Beyond the misuses, our review lists previous serious attempts to use beamforming to the extended source model. However, they require knowing the source position in advance, which appears to be a too strong assumption in practice. The novel method, named SABLE (Sparsity And Beamforming for brain source Localization and Estimation), overcomes this problem by using a latent variable and by exploiting both the UG and the SISSY methods. As a result, it localizes and estimates source activity at the same time. Preliminary results on simulations show that SABLE is more efficient than both methods to localize one distributed source for high SNR values. However, it requires a fine-tuning of the hyper-parameter between each cost function for low SNR values.

The fifth contribution of this thesis is the formulation of the SABLE pipeline and its assessment against three mature methods from the literature, namely wMNE, sLORETA and cMEM. Although this pipeline is strongly inspired by those implemented in Brainstorm to provide a fair comparison, it also contains new practices such as cross-fading, the removal of the corpus callosum and the new rank selection criterion based on RMT. Note that the latter requires the choice of a variance value that can be considered arbitrarily. Quantitative results on simulations show that SABLE is, from a broad perspective, as effective as SISSY and outperforms the three traditional methods wMNE, sLORETA and cMEM, to localize three extended patches with synchronized activity. More precisely, by using the new criterion named Patch Amplitude Ratio, we show that the beamforming part of SABLE allows us a better distribution of energy between the patches, which SISSY does not provide. This feature makes it possible to highlight potential epileptogenic zones hidden by an algorithm flaw, as we have seen through real data. However, we must remain cautious about these preliminary results as we only have one real data set.

This thesis is at the interface of three domains: physic, mathematics and algorithmic. The work was carried out taking into account these three dimensions with their unique constraints. This conjoint work leads to the proposition of new methods and pipelines assessed for both simulated and real epileptic data. Therefore, the perspectives of this work can be distinguished between these two topics. Regarding the method perspectives, one of the first perspectives which could emerge from this thesis work is to propose a fusion between Temporal-Smoothing-based beamforming and SABLE. However, as the results for the real case showed, there is no interest in merging them since the reconstruction results of SABLE are already smooth. However, there is an interest to work with the AOLS beamformer instead of the UG one at the basis of SABLE leading to a double shrinkage scenario. A second perspective for SABLE would be to test its effect in combination with the array gain beamforming method. Indeed, if we take the wMNE and sLORETA methods, both introduced priors on the source depth by means of a weight matrix. This assumption may be at the origin of the results observed on real data and is similar to the array gain method. Since the addition of such a preprocessing method is not straightforward due to scaling issues, it should be interesting to propose and assess a new SABLE method. Finally, we need to extend our qualitative study to a more extensive set of real data. We will have to strengthen the pipeline to guarantee the stability of the results, which is the subject of the next paragraph.

Regarding the pipeline perspectives, further work is warranted on investigating new hyper-parameters tuning strategies. Indeed, the methodology proposed through the SISSY method is quite particular and the role of the hyper-parameter is different from that of traditional methods [1]. Due to the variables substitution in the L_1 norms, the hyper-parameter between the constraints and the cost function now plays a thresholding role in the proximal operator. Although this new role is interesting, it does not compensate for the loss of the ability to emphasize either data fidelity or constraint depending on the noise level. Consequently, the hyper-parameter strategy for SISSY needs to be revised and extended to SABLE. Furthermore, we would like to extend the proposed rank selection method proposed in the pipeline to categorize more finely the eigenvalues belonging to the sources of interest, a structured noise and a Gaussian noise. These assumptions are already used in the field of array processing and are promising in the field of BSI, especially on denoising issues [2]. Once these two topics are addressed, a fully automated pipeline will be available for SABLE.

Appendices

Appendix A

Supplementary methods and materials

The following chapter needs these three assumptions:

A.A.1 The source random process $\{\boldsymbol{s}(t)\}$ and the noise random process $\{\boldsymbol{\nu}(t)\}$ are statistically independent to the second-order for any fixed index t;

A.A.2 The SOSE noise random process $\{\boldsymbol{\nu}(t)\}$ is a Gaussian random process with zero mean and with a covariance matrix equals to $\widehat{\boldsymbol{C}_{\boldsymbol{\nu}}}$;

A.A.3 For any random vector $\boldsymbol{a}$ and non random matrix $\boldsymbol{B}$, the following equality holds: $\mathbb{E}\{\boldsymbol{a}^{\mathsf{T}}\boldsymbol{B}\boldsymbol{a}\} = \mathrm{Tr}\{\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{a}}\} + \bar{a}^{\mathsf{T}}\boldsymbol{B}\bar{a}$.

The assumptions (**A.A.1**) and (**A.A.2**) are generally well-accepted or can be met using whitening. For assumption (**A.A.3**), we provide the following proof originating from [204]:

$$
\begin{aligned}
\mathbb{E}\{\boldsymbol{a}^{\mathsf{T}}\boldsymbol{B}\boldsymbol{a}\} &= \mathbb{E}\{\sum_{i=1}^{n}\sum_{j=1}^{n} a_i^{\mathsf{T}} b_{i,j} a_j\} \\
&= \mathbb{E}\{\sum_{i=1}^{n}\sum_{j=1}^{n} b_{i,j} a_i^{\mathsf{T}} a_j\} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} b_{i,j}\mathbb{E}\{a_i^{\mathsf{T}} a_j\} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} b_{i,j}(\boldsymbol{C}_{\boldsymbol{a}_{i,j}} + \bar{\boldsymbol{a}}_i\bar{\boldsymbol{a}}_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} b_{i,j}\boldsymbol{C}_{\boldsymbol{a}_{i,j}} + \sum_{i=1}^{n}\sum_{j=1}^{n} b_{i,j}\bar{\boldsymbol{a}}_i\bar{\boldsymbol{a}}_j \\
&= \sum_{i=1}^{n}\left[\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{a}}\right]_{i,i} + \bar{\boldsymbol{a}}^{\mathsf{T}}\boldsymbol{B}\bar{\boldsymbol{a}} = \mathrm{Tr}\{\boldsymbol{B}\boldsymbol{C}_{\boldsymbol{a}}\} + \bar{\boldsymbol{a}}^{\mathsf{T}}\boldsymbol{B}\bar{\boldsymbol{a}}
\end{aligned}
\tag{A.1}
$$

A.1 Mathematical link between EEG data SCM and source SCM

The EEG recordings at time t are assumed to be a realization of an N-dimensional SOSE random process $\{\boldsymbol{x}(t)\}$ given by:

$$\boldsymbol{x}(t) = \boldsymbol{G}\boldsymbol{s}(t) + \boldsymbol{\nu}(t) \tag{A.2}$$

where $\{\boldsymbol{s}(t)\}$ is an L-dimensional SOSE source random process and where $\{\boldsymbol{\nu}(t)\}$ is a SOSE noise random process. $\boldsymbol{G}$ is the so-called leadfield matrix, which models the electrical propagation of the sources to the scalp sensors. Using assumptions **A.A.1** and **A.A.2**, we can link the statistical properties of the EEG recordings and the sources with the following equations:

$$\begin{aligned}
\overline{\boldsymbol{x}} &= \mathbb{E}[\boldsymbol{x}(t)] \\
&\overset{(A.2)}{=} \mathbb{E}[\boldsymbol{G}\boldsymbol{s}(t) + \boldsymbol{\nu}(t)] \\
&= \mathbb{E}[\boldsymbol{G}\boldsymbol{s}(t)] + \mathbb{E}[\boldsymbol{\nu}(t)] \\
&= \boldsymbol{G}\overline{\boldsymbol{s}} + \overline{\boldsymbol{\nu}}
\end{aligned} \tag{A.3}$$

$$\begin{aligned}
\widehat{\boldsymbol{C}_x} &= \mathbb{E}[(\boldsymbol{x}(t) - \overline{\boldsymbol{x}})(\boldsymbol{x}(t) - \overline{\boldsymbol{x}})^\top] \\
&= \mathbb{E}[(\boldsymbol{G}\boldsymbol{s}(t) + \boldsymbol{\nu}(t) - \boldsymbol{G}\overline{\boldsymbol{s}} - \overline{\boldsymbol{\nu}})(\boldsymbol{G}\boldsymbol{s}(t) + \boldsymbol{\nu}(t) - \boldsymbol{G}\overline{\boldsymbol{s}} - \overline{\boldsymbol{\nu}})^\top] \\
&= \mathbb{E}[\boldsymbol{G}\boldsymbol{s}(t)\boldsymbol{s}(t)^\top\boldsymbol{G}^\top] + \mathbb{E}[\boldsymbol{G}\boldsymbol{s}(t)\boldsymbol{\nu}(t)^\top] - \mathbb{E}[\boldsymbol{G}\boldsymbol{s}(t)\overline{\boldsymbol{s}}^\top\boldsymbol{G}^\top] - \mathbb{E}[\boldsymbol{G}\boldsymbol{s}(t)\overline{\boldsymbol{\nu}}^\top] \\
&\quad + \mathbb{E}[\boldsymbol{\nu}(t)\boldsymbol{s}(t)^\top\boldsymbol{G}^\top] + \mathbb{E}[\boldsymbol{\nu}(t)\boldsymbol{\nu}(t)^\top] - \mathbb{E}[\boldsymbol{\nu}(t)\overline{\boldsymbol{s}}^\top\boldsymbol{G}^\top] - \mathbb{E}[\boldsymbol{\nu}(t)\overline{\boldsymbol{\nu}}^\top] \\
&\quad - \mathbb{E}[\boldsymbol{G}\overline{\boldsymbol{s}}\boldsymbol{s}(t)^\top\boldsymbol{G}^\top] - \mathbb{E}[\boldsymbol{G}\overline{\boldsymbol{s}}\boldsymbol{\nu}(t)^\top] + \mathbb{E}[\boldsymbol{G}\overline{\boldsymbol{s}}\overline{\boldsymbol{s}}^\top\boldsymbol{G}^\top] + \mathbb{E}[\boldsymbol{G}\overline{\boldsymbol{\nu}}^\top] \\
&\quad - \mathbb{E}[\overline{\boldsymbol{\nu}}\boldsymbol{s}(t)^\top\boldsymbol{G}^\top] - \mathbb{E}[\overline{\boldsymbol{\nu}}\boldsymbol{\nu}(t)^\top] + \mathbb{E}[\overline{\boldsymbol{\nu}}\overline{\boldsymbol{s}}^\top\boldsymbol{G}^\top] + \mathbb{E}[\overline{\boldsymbol{\nu}}\overline{\boldsymbol{\nu}}^\top] \\
&= \boldsymbol{G}\mathbb{E}[\boldsymbol{s}(t)\boldsymbol{s}(t)^\top]\boldsymbol{G}^\top - 2\boldsymbol{G}\mathbb{E}[\overline{\boldsymbol{s}}\boldsymbol{s}(t)^\top]\boldsymbol{G}^\top + \boldsymbol{G}\mathbb{E}[\overline{\boldsymbol{s}}\overline{\boldsymbol{s}}^\top]\boldsymbol{G}^\top \\
&\quad + \mathbb{E}[\boldsymbol{\nu}(t)\boldsymbol{\nu}(t)^\top] - 2\mathbb{E}[\boldsymbol{\nu}(t)\overline{\boldsymbol{\nu}}^\top] + \mathbb{E}[\overline{\boldsymbol{\nu}}\overline{\boldsymbol{\nu}}^\top] \\
&\quad + 2\boldsymbol{G}(\mathbb{E}[\boldsymbol{s}(t)\boldsymbol{\nu}(t)^\top] - \mathbb{E}[\boldsymbol{s}(t)\overline{\boldsymbol{\nu}}^\top] - \mathbb{E}[\overline{\boldsymbol{s}}\boldsymbol{\nu}(t)^\top] + \mathbb{E}[\overline{\boldsymbol{s}}\overline{\boldsymbol{\nu}}^\top]) \\
&= \boldsymbol{G}\mathbb{E}[(\boldsymbol{s}(t) - \overline{\boldsymbol{s}})(\boldsymbol{s}(t) - \overline{\boldsymbol{s}})^\top]\boldsymbol{G}^\top + \mathbb{E}[(\boldsymbol{\nu}(t) - \overline{\boldsymbol{\nu}})(\boldsymbol{\nu}(t) - \overline{\boldsymbol{\nu}})^\top] \\
&\quad + 2\boldsymbol{G}\mathbb{E}[(\boldsymbol{s}(t) - \overline{\boldsymbol{s}})(\boldsymbol{\nu}(t) - \overline{\boldsymbol{\nu}})^\top] \\
&\overset{(\mathbf{A.A.2})}{=} \boldsymbol{G}\widehat{\boldsymbol{C}_s}\boldsymbol{G}^\top + \widehat{\boldsymbol{C}_\nu} + 2\boldsymbol{G}\mathbb{E}[(\boldsymbol{s}(t) - \overline{\boldsymbol{s}})(\boldsymbol{\nu}(t) - \overline{\boldsymbol{\nu}})^\top] \\
&\overset{(\mathbf{A.A.1})}{=} \boldsymbol{G}\widehat{\boldsymbol{C}_s}\boldsymbol{G}^\top + \widehat{\boldsymbol{C}_\nu}
\end{aligned} \tag{A.4}$$

A.2 Beamforming optimization step development: proof of the optimality of beamforming

As defended in chapter 2, beamforming methods allow us to analytically and efficiently estimate the time course of the dipole. This efficiency is due to the optimization process that returns an unbiased and non-distorted signal estimator if sources are uncorrelated. In cases of correlated sources, the optimization step backfires as it will cancel their activities. Thus, if the sources of interest are correlated or in case of interferences, it will be impossible to find them. In the previous section of this manuscript, we studied this effect using simulations. In this appendix, we explain this phenomenon through mathematical development. Considering the estimation of the source activity $\hat{\mathbf{s}}$ at the position θ_{ℓ_i}, the beamforming can be express in terms of mean and the covariance as:

$$\begin{aligned}
\bar{\hat{s}}(\theta_{\ell_i}) &= \mathbb{E}[\boldsymbol{w}(\theta_{\ell_i})^\mathsf{T}\boldsymbol{x}(t)] \\
&= \boldsymbol{w}(\theta_{\ell_i})^\mathsf{T}\bar{\boldsymbol{x}} \\
\widehat{\sigma}(\theta_{\ell_i})^2 &= \mathbb{E}[(s(\theta_{\ell_i},t) - \bar{\hat{s}}(\theta_{\ell_i}))(s(\theta_{\ell_i},t) - \bar{\hat{s}}(\theta_{\ell_i})^\mathsf{T}] \\
&= \mathbb{E}[(\boldsymbol{w}(\theta_{\ell_i})^\mathsf{T}\boldsymbol{x}(t) - \boldsymbol{w}(\theta_{\ell_i})^\mathsf{T}\bar{\boldsymbol{x}})(\boldsymbol{x}(t)^\mathsf{T}\boldsymbol{w}(\theta_{\ell_i}) - \bar{\boldsymbol{x}}^\mathsf{T}\boldsymbol{w}(\theta_{\ell_i}))] \\
&= \boldsymbol{w}(\theta_{\ell_i})^\mathsf{T}\mathbb{E}[(\boldsymbol{x}(t) - \bar{\boldsymbol{x}})(\boldsymbol{x}(t)^\mathsf{T} - \bar{\boldsymbol{x}}^\mathsf{T})]\boldsymbol{w}(\theta_{\ell_i}) \\
&= \boldsymbol{w}(\theta_{\ell_i})^\mathsf{T}\boldsymbol{C}_{\boldsymbol{x}}\boldsymbol{w}(\theta_{\ell_i})
\end{aligned} \tag{A.5}$$

Injecting expression (A.4) of $\boldsymbol{C}_{\boldsymbol{x}}$ into (A.5) leads to the following expression of the variance:

$$\begin{aligned}
\widehat{\sigma}(\theta_{\ell_i})^2 &= \boldsymbol{w}(\theta_{\ell_i})^\mathsf{T}\boldsymbol{C}_{\boldsymbol{x}}\boldsymbol{w}(\theta_{\ell_i}) \\
&= \boldsymbol{w}(\theta_{\ell_i})^\mathsf{T}(\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{s}}\boldsymbol{G}^\mathsf{T} + \boldsymbol{C}_{\boldsymbol{\nu}})\boldsymbol{w}(\theta_{\ell_i}) \\
&= \boldsymbol{w}(\theta_{\ell_i})^\mathsf{T}\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{s}}\boldsymbol{G}^\mathsf{T}\boldsymbol{w}(\theta_{\ell_i}) + \boldsymbol{w}(\theta_{\ell_i})^\mathsf{T}\boldsymbol{C}_{\boldsymbol{\nu}}\boldsymbol{w}(\theta_{\ell_i})
\end{aligned} \tag{A.6}$$

The noise is considered to be null to simplify to demonstration, i.e. $\boldsymbol{w}(\theta_{\ell_i})^\mathsf{T}\boldsymbol{C}_{\boldsymbol{\nu}}\boldsymbol{w}(\theta_{\ell_i}) = 0$. We will consider that the SOI is at the position $i = 1$.

We must now distinguish several cases:

— The case where the dipoles are uncorrelated;

— The case where the dipoles are correlated;

Uncorrelated source scenario In the uncorrelated source scenario, the source covariance matrix $\boldsymbol{C}_{\boldsymbol{s}}$ is a diagonal matrix with the source variance θ_{ℓ_i} located at the i -th row and i -th column. The equation (A.6) can therefore be further developed as such:

$$\widehat{\sigma}(\theta_{\ell_1})^2 = \boldsymbol{w}(\theta_{\ell_1})^\mathsf{T}\boldsymbol{G}\boldsymbol{C}_{\boldsymbol{s}}\boldsymbol{G}^\mathsf{T}\boldsymbol{w}(\theta_{\ell_1})$$

$$
\begin{aligned}
&= \boldsymbol{w}(\theta_{\ell_1})^\intercal \sum_{i=1}^{L} \Big(\boldsymbol{g}(\theta_{\ell_i}) \sigma(\theta_{\ell_i})^2 \boldsymbol{g}(\theta_{\ell_i})^\intercal \Big) \boldsymbol{w}(\theta_{\ell_1})^\intercal \qquad \text{(A.7)} \\
&= \boldsymbol{w}(\theta_{\ell_1})^\intercal \boldsymbol{g}(\theta_{\ell_1}) \sigma(\theta_{\ell_1})^2 \boldsymbol{g}(\theta_{\ell_1})^\intercal \boldsymbol{w}(\theta_{\ell_1})^\intercal + \sum_{i=2}^{L} \boldsymbol{w}(\theta_{\ell_1})^\intercal \boldsymbol{g}(\theta_{\ell_i}) \sigma(\theta_{\ell_i})^2 \boldsymbol{g}(\theta_{\ell_i})^\intercal \boldsymbol{w}(\theta_{\ell_1})^\intercal \\
&= ||\boldsymbol{w}(\theta_{\ell_1})^\intercal \boldsymbol{g}(\theta_{\ell_1})||_2^2 \sigma(\theta_{\ell_1})^2 + \sum_{i=2}^{L} ||\boldsymbol{w}(\theta_{\ell_1})^\intercal \boldsymbol{g}(\theta_{\ell_i})||_2^2 \sigma(\theta_{\ell_i})^2
\end{aligned}
$$

When only the distortionless constraint is used (like in UG), it allows us to put the first term $||\boldsymbol{w}(\theta_{\ell_1})^\intercal \boldsymbol{g}(\theta_{\ell_1})||_2^2$ to the unit. The optimization process will ensure that the second term will be put to 0 since the minimum for a squared term is 0. Thus, the result of the whole optimization is the constant variance $\sigma(\theta_{\ell_i})^2$. We can conclude that the estimate $\widehat{\sigma}(\theta_{\ell_1})^2$ is equal to the true value of the source variance hence an optimal estimation. Note that the optimization is done correctly because all terms are positive.

$$
\begin{aligned}
\widehat{\sigma}(\theta_{\ell_1})^2 &= \underbrace{||\boldsymbol{w}(\theta_{\ell_1})^\intercal \boldsymbol{g}(\theta_{\ell_1})||_2^2}_{=1} \underbrace{\sigma(\theta_{\ell_1})^2}_{\text{cst}} + \sum_{i=2}^{\text{L}} \underbrace{||\boldsymbol{w}(\theta_{\ell_1})^\intercal \boldsymbol{g}(\theta_{\ell_i})||_2^2}_{\to 0} \underbrace{\sigma(\theta_{\ell_i})^2}_{\text{cst}} \qquad \text{(A.8)} \\
&= \sigma(\theta_{\ell_1})^2
\end{aligned}
$$

Correlated source scenario Assuming now that some sources are correlated, we repeat the set of computational steps used previously considering the cross-term element of the source covariance matrix $\sigma(\theta_{\ell_i})\sigma(\theta_{\ell_j})$ located at the i-th row and j-th column.

$$
\begin{aligned}
\widehat{\sigma}(\theta_{\ell_1})^2 &= \boldsymbol{w}(\theta_{\ell_1})^\intercal \boldsymbol{G} \boldsymbol{C}_{\boldsymbol{s}} \boldsymbol{G}^\intercal \boldsymbol{w}(\theta_{\ell_1}) \qquad \text{(A.9)} \\
&= \boldsymbol{w}(\theta_{\ell_1})^\intercal \sum_{i=1}^{L} \Big(\boldsymbol{g}(\theta_{\ell_i}) \sigma(\theta_{\ell_i})^2 \boldsymbol{g}(\theta_{\ell_i})^\intercal \Big) \boldsymbol{w}(\theta_{\ell_1})^\intercal \\
&+ \boldsymbol{w}(\theta_{\ell_1})^\intercal \sum_{i=1}^{L} \sum_{j=i+1}^{L} \Big(\boldsymbol{g}(\theta_{\ell_i}) \sigma(\theta_{\ell_i}) \sigma(\theta_{\ell_j}) \boldsymbol{g}(\theta_{\ell_j})^\intercal + \boldsymbol{g}(\theta_{\ell_j}) \sigma(\theta_{\ell_j}) \sigma(\theta_{\ell_i}) \boldsymbol{g}(\theta_{\ell_i})^\intercal \Big) \boldsymbol{w}(\theta_{\ell_1})^\intercal \\
&= \boldsymbol{w}(\theta_{\ell_1})^\intercal \boldsymbol{g}(\theta_{\ell_1}) \sigma(\theta_{\ell_1})^2 \boldsymbol{g}(\theta_{\ell_1})^\intercal \boldsymbol{w}(\theta_{\ell_1})^\intercal + \boldsymbol{w}(\theta_{\ell_1})^\intercal \sum_{i=2}^{L} \Big(\boldsymbol{g}(\theta_{\ell_i}) \sigma(\theta_{\ell_i})^2 \boldsymbol{g}(\theta_{\ell_i})^\intercal \Big) \boldsymbol{w}(\theta_{\ell_1})^\intercal \\
&+ \boldsymbol{w}(\theta_{\ell_1})^\intercal \sum_{j=2}^{L} \Big(\boldsymbol{g}(\theta_{\ell_1}) \sigma(\theta_{\ell_1}) \sigma(\theta_{\ell_j}) \boldsymbol{g}(\theta_{\ell_j})^\intercal + \boldsymbol{g}(\theta_{\ell_j}) \sigma(\theta_{\ell_j}) \sigma(\theta_{\ell_1}) \boldsymbol{g}(\theta_{\ell_1})^\intercal \Big) \boldsymbol{w}(\theta_{\ell_1}) \\
&+ \boldsymbol{w}(\theta_{\ell_1})^\intercal \sum_{i=2}^{L} \sum_{j=i+1}^{L} \Big(\boldsymbol{g}(\theta_{\ell_i}) \sigma(\theta_{\ell_i}) \sigma(\theta_{\ell_j}) \boldsymbol{g}(\theta_{\ell_j})^\intercal + \boldsymbol{g}(\theta_{\ell_j}) \sigma(\theta_{\ell_j}) \sigma(\theta_{\ell_i}) \boldsymbol{g}(\theta_{\ell_i})^\intercal \Big) \boldsymbol{w}(\theta_{\ell_1}) \\
&= \boldsymbol{w}(\theta_{\ell_1})^\intercal \boldsymbol{g}(\theta_{\ell_1}) \sigma(\theta_{\ell_1})^2 \boldsymbol{g}(\theta_{\ell_1})^\intercal \boldsymbol{w}(\theta_{\ell_1})^\intercal + \boldsymbol{w}(\theta_{\ell_1})^\intercal \sum_{i=2}^{L} \Big(\boldsymbol{g}(\theta_{\ell_i}) \sigma(\theta_{\ell_i})^2 \boldsymbol{g}(\theta_{\ell_i})^\intercal \Big) \boldsymbol{w}(\theta_{\ell_1})^\intercal
\end{aligned}
$$

$$
\begin{aligned}
&+ 2\boldsymbol{w}(\theta_{\ell_1})^{\mathsf{T}} \sum_{j=2}^{L} \left(\boldsymbol{g}(\theta_{\ell_1})\sigma(\theta_{\ell_1})\sigma(\theta_{\ell_j})\boldsymbol{g}(\theta_{\ell_j})^{\mathsf{T}} \right) \boldsymbol{w}(\theta_{\ell_1}) \\
&+ 2\boldsymbol{w}(\theta_{\ell_1})^{\mathsf{T}} \sum_{i=2}^{L} \sum_{j=i+1}^{L} \left(\boldsymbol{g}(\theta_{\ell_i})\sigma(\theta_{\ell_i})\sigma(\theta_{\ell_j})\boldsymbol{g}(\theta_{\ell_j})^{\mathsf{T}} \right) \boldsymbol{w}(\theta_{\ell_1}) \\
&= ||\boldsymbol{w}(\theta_{\ell_1})^{\mathsf{T}}\boldsymbol{g}(\theta_{\ell_1})||_2^2 \sigma(\theta_{\ell_1})^2 + \sum_{i=2}^{L} ||\boldsymbol{w}(\theta_{\ell_1})^{\mathsf{T}}\boldsymbol{g}(\theta_{\ell_i})||_2^2 \sigma(\theta_{\ell_i})^2 \\
&+ 2\boldsymbol{w}(\theta_{\ell_1})^{\mathsf{T}} \sum_{j=2}^{L} \left(\boldsymbol{g}(\theta_{\ell_1})\sigma(\theta_{\ell_1})\sigma(\theta_{\ell_i})\boldsymbol{g}(\theta_{\ell_j})^{\mathsf{T}} \right) \boldsymbol{w}(\theta_{\ell_1}) \\
&+ 2\boldsymbol{w}(\theta_{\ell_1})^{\mathsf{T}} \sum_{i=2}^{L} \sum_{j=i+1}^{L} \left(\boldsymbol{g}(\theta_{\ell_i})\sigma(\theta_{\ell_i})\sigma(\theta_{\ell_j})\boldsymbol{g}(\theta_{\ell_j})^{\mathsf{T}} \right) \boldsymbol{w}(\theta_{\ell_1})
\end{aligned}
$$

As we can see, if only the distortionless constraint is applied, the positivity is not guaranteed for all elements. Indeed, the 3rd and 4th terms are not null and not squared. Thus, nothing prevents the optimization from assigning a negative value to these terms and decreases the total to zero.

$$
\begin{aligned}
\widehat{\sigma}(\theta_{\ell_1})^2 = &\underbrace{||\boldsymbol{w}(\theta_{\ell_1})^{\mathsf{T}}\boldsymbol{g}(\theta_{\ell_1})||_2^2}_{=1} \underbrace{\sigma(\theta_{\ell_1})^2}_{\to \text{cst}} + \sum_{i=2}^{L} \underbrace{||\boldsymbol{w}(\theta_{\ell_1})^{\mathsf{T}}\boldsymbol{g}(\theta_{\ell_i})||_2^2}_{\to 0} \underbrace{\sigma(\theta_{\ell_i})^2}_{\to \text{cst}} \\
&+ 2\sum_{i=2}^{L} \underbrace{\boldsymbol{w}(\theta_{\ell_1})^{\mathsf{T}}\boldsymbol{g}(\theta_{\ell_1})}_{=1} \underbrace{\sigma(\theta_{\ell_1})\sigma(\theta_{\ell_i})}_{\to \text{cst}} \underbrace{\boldsymbol{g}(\theta_{\ell_i})^{\mathsf{T}}\boldsymbol{w}(\theta_{\ell_1})}_{\to <0} \\
&+ 2\sum_{i=2}^{L} \sum_{j=i+1}^{L} \underbrace{\boldsymbol{w}(\theta_{\ell_1})^{\mathsf{T}}\boldsymbol{g}(\theta_{\ell_i})}_{\to <0} \underbrace{\sigma(\theta_{\ell_i})\sigma(\theta_{\ell_j})}_{\to \text{cst}} \underbrace{\boldsymbol{g}(\theta_{\ell_j})^{\mathsf{T}}\boldsymbol{w}(\theta_{\ell_1})}_{\to <0} \\
&\leq \sigma(\theta_{\ell_1})^2
\end{aligned}
\tag{A.10}
$$

This behavior leads to a decrease or even a cancellation of $\sigma(\theta_{\ell_1})^2$. Our estimation as well as the source localization and reconstruction are consequently altered. Note that, we generalize the notion of covariance. Indeed, there are no differences done between:

— The correlation between the epileptic activities;

— The correlation between the epileptic activities and the background activities;

— The correlation within the background activities;

Constraint added by the nulling beamforming: The correlated sources are forced to be null by means of an additional constraint with nulling beamforming. If there are only two sources located at the known positions θ_{ℓ_1} and θ_{ℓ_2}, the constraint is written as such: $\boldsymbol{w}(\theta_{\ell_1})\boldsymbol{g}(\theta_{\ell_2}) = 0$. The proposition for the filter $\boldsymbol{w}(\theta_{\ell_1})$ becomes:

$$\begin{aligned}\widehat{\sigma}(\theta_{\ell_1})^2 &= \underbrace{||\boldsymbol{w}(\theta_{\ell_1})^\mathsf{T}\boldsymbol{g}(\theta_{\ell_1})||_2^2}_{=1}\underbrace{\sigma(\theta_{\ell_1})^2}_{\to\text{cst}} + \sum_{i=2}^{L}\underbrace{||\boldsymbol{w}(\theta_{\ell_1})^\mathsf{T}\boldsymbol{g}(\theta_{\ell_i})||_2^2}_{\to 0}\underbrace{\sigma(\theta_{\ell_i})^2}_{\to\text{cst}} \\ &+ 2\underbrace{\boldsymbol{w}(\theta_{\ell_1})^\mathsf{T}\boldsymbol{g}(\theta_{\ell_2})}_{=0}\underbrace{\sigma(\theta_{\ell_2})\sigma(\theta_{\ell_1})}_{\to\text{cst}}\underbrace{\boldsymbol{g}(\theta_{\ell_1})^\mathsf{T}\boldsymbol{w}(\theta_{\ell_1})}_{=1} \\ &= \sigma(\theta_{\ell_1})^2\end{aligned} \tag{A.11}$$

Again, we have an optimal estimator. However, the assumption of two unique correlated sources is, in practice, unachievable. Moreover, the knowledge of the exact position of these correlated sources is also impossible to have. Hence, the proposition of the LCMV beamformer that imposes a cancellation constraint on every location except the one studied:

$$\begin{aligned}\widehat{\sigma}(\theta_{\ell_1})^2 &= \underbrace{||\boldsymbol{w}(\theta_{\ell_1})^\mathsf{T}\boldsymbol{g}(\theta_{\ell_1})||_2^2}_{=1}\underbrace{\sigma(\theta_{\ell_1})^2}_{\to\text{cst}} + \sum_{i=2}^{L}\underbrace{||\boldsymbol{w}(\theta_{\ell_1})^\mathsf{T}\boldsymbol{g}(\theta_{\ell_i})||_2^2}_{=0}\underbrace{\sigma(\theta_{\ell_i})^2}_{\to\text{cst}} \\ &+ 2\sum_{i=2}^{L}\underbrace{\boldsymbol{w}(\theta_{\ell_1})^\mathsf{T}\boldsymbol{g}(\theta_{\ell_1})}_{=1}\underbrace{\sigma(\theta_{\ell_1})\sigma(\theta_{\ell_i})}_{\to\text{cst}}\underbrace{\boldsymbol{g}(\theta_{\ell_i})^\mathsf{T}\boldsymbol{w}(\theta_{\ell_1})}_{=0} \\ &+ 2\sum_{i=2}^{L}\sum_{j=i+1}^{L}\underbrace{\boldsymbol{w}(\theta_{\ell_1})^\mathsf{T}\boldsymbol{g}(\theta_{\ell_i})}_{=0}\underbrace{\sigma(\theta_{\ell_i})\sigma(\theta_{\ell_j})}_{\to\text{cst}}\underbrace{\boldsymbol{g}(\theta_{\ell_j})^\mathsf{T}\boldsymbol{w}(\theta_{\ell_1})}_{=0} \\ &= \sigma(\theta_{\ell_1})^2\end{aligned} \tag{A.12}$$

The problem with this method is the dimension of the constraints, which poses a second ill-posed inverse problem in the Hadamard sense. Hence the proposition of solving techniques that may bring distortion or bias (see section 2.3.3).

A.3 Development MMSE

The objective of MMSE beamforming methods is to find a filter that will decrease the expectation of the estimation error $||\boldsymbol{s}(t) - \hat{\boldsymbol{s}}(t)||_2^2$. To do so, we need to find an estimator that minimizes both the variance and the bias.

Then, the MMSE can be simplified as follows:

$$\begin{aligned}\mathbb{E}[||\boldsymbol{s}(t) - \hat{\boldsymbol{s}}(t)||_2^2] &\overset{(2.2)}{=} \mathbb{E}[||\boldsymbol{s}(t) - \boldsymbol{W}^\mathsf{T}\boldsymbol{x}(t)||_2^2] \\ &\overset{(1.1)}{=} \mathbb{E}[||\boldsymbol{s}(t) - \boldsymbol{W}^\mathsf{T}(\boldsymbol{G}\boldsymbol{s}(t) + \boldsymbol{n}(t))||_2^2] \\ &= \mathbb{E}[||\boldsymbol{s}(t) - \boldsymbol{W}^\mathsf{T}\boldsymbol{G}\boldsymbol{s}(t) + \boldsymbol{W}^\mathsf{T}\boldsymbol{n}(t)||_2^2] \\ &= \mathbb{E}[(\boldsymbol{s}(t) - \boldsymbol{W}^\mathsf{T}\boldsymbol{G}\boldsymbol{s}(t) + \boldsymbol{W}^\mathsf{T}\boldsymbol{n}(t))^\mathsf{T}(\boldsymbol{s}(t) - \boldsymbol{W}^\mathsf{T}\boldsymbol{G}\boldsymbol{s}(t) + \boldsymbol{W}^\mathsf{T}\boldsymbol{n}(t))] \\ &= \mathbb{E}[\boldsymbol{s}(t)^\mathsf{T}\boldsymbol{s}(t) - \boldsymbol{s}(t)^\mathsf{T}\boldsymbol{W}^\mathsf{T}\boldsymbol{G}\boldsymbol{s}(t) + \boldsymbol{s}(t)^\mathsf{T}\boldsymbol{W}^\mathsf{T}\boldsymbol{n}(t)\end{aligned} \tag{A.13}$$

$$
\begin{aligned}
&\quad - \boldsymbol{s}(t)^\intercal \boldsymbol{G}^\intercal \boldsymbol{s}(t) + \boldsymbol{s}(t)^\intercal \boldsymbol{G}^\intercal \boldsymbol{W}^\intercal \boldsymbol{G}\boldsymbol{s}(t) \\
&\quad - \boldsymbol{s}(t)^\intercal \boldsymbol{G}^\intercal \boldsymbol{W}^\intercal \boldsymbol{n}(t) + \boldsymbol{n}(t)^\intercal \boldsymbol{W}\boldsymbol{s}(t) \\
&\quad - \boldsymbol{n}(t)^\intercal \boldsymbol{W}\boldsymbol{W}^\intercal \boldsymbol{G}\boldsymbol{s}(t) + \boldsymbol{n}(t)^\intercal \boldsymbol{W}\boldsymbol{W}^\intercal \boldsymbol{n}(t)] \\
&= \mathbb{E}[\boldsymbol{s}(t)^\intercal \boldsymbol{s}(t) - \boldsymbol{s}(t)^\intercal \boldsymbol{W}^\intercal \boldsymbol{G}\boldsymbol{s}(t) - \boldsymbol{s}(t)^\intercal \boldsymbol{G}^\intercal \boldsymbol{s}(t) \\
&\quad + \boldsymbol{s}(t)^\intercal \boldsymbol{G}^\intercal \boldsymbol{W}^\intercal \boldsymbol{G}\boldsymbol{s}(t)] + \mathbb{E}[\boldsymbol{s}(t)^\intercal \boldsymbol{W}^\intercal \boldsymbol{n}(t)] \\
&\quad - \mathbb{E}[\boldsymbol{s}(t)^\intercal \boldsymbol{G}^\intercal \boldsymbol{W}^\intercal \boldsymbol{n}(t)] + \mathbb{E}[\boldsymbol{n}(t)^\intercal \boldsymbol{W}\boldsymbol{s}(t)] \\
&\quad - \mathbb{E}[\boldsymbol{n}(t)^\intercal \boldsymbol{W}\boldsymbol{W}^\intercal \boldsymbol{G}\boldsymbol{s}(t)] + \mathbb{E}[\boldsymbol{n}(t)^\intercal \boldsymbol{W}\boldsymbol{W}^\intercal \boldsymbol{n}(t)] \\
&= \mathbb{E}[||\boldsymbol{s}(t) - \boldsymbol{W}^\intercal \boldsymbol{G}\boldsymbol{s}(t)||_2^2] + \mathbb{E}[\boldsymbol{s}(t)^\intercal \boldsymbol{W}^\intercal \boldsymbol{n}(t)] \\
&\quad - \mathbb{E}[\boldsymbol{s}(t)^\intercal \boldsymbol{G}^\intercal \boldsymbol{W}^\intercal \boldsymbol{n}(t)] + \mathbb{E}[\boldsymbol{n}(t)^\intercal \boldsymbol{W}\boldsymbol{s}(t)] \\
&\quad - \mathbb{E}[\boldsymbol{n}(t)^\intercal \boldsymbol{W}\boldsymbol{W}^\intercal \boldsymbol{G}\boldsymbol{s}(t)] + \mathbb{E}[\boldsymbol{n}(t)^\intercal \boldsymbol{W}\boldsymbol{W}^\intercal \boldsymbol{n}(t)] \\
&= \mathbb{E}[||\boldsymbol{s}(t) - \boldsymbol{W}^\intercal \boldsymbol{G}\boldsymbol{s}(t)||_2^2] + \mathbb{E}[\boldsymbol{n}(t)^\intercal \boldsymbol{W}\boldsymbol{W}^\intercal \boldsymbol{n}(t)] \\
&= \mathbb{E}[||\boldsymbol{s}(t) - \boldsymbol{W}^\intercal \boldsymbol{G}\boldsymbol{s}(t)||_2^2] + \mathbb{E}[\boldsymbol{n}(t)^\intercal \boldsymbol{W} I_L \boldsymbol{W}^\intercal \boldsymbol{n}(t)] \\
&\overset{(\mathbf{A.A.3})}{=} \mathbb{E}[||\boldsymbol{s}(t) - \boldsymbol{W}^\intercal \boldsymbol{G}\boldsymbol{s}(t)||_2^2] + \mathrm{Tr}\{I_L \boldsymbol{C}_{(\boldsymbol{W}^\intercal \boldsymbol{n}(t))}\} \\
&= \mathbb{E}[||\boldsymbol{s}(t) - \boldsymbol{W}^\intercal \boldsymbol{G}\boldsymbol{s}(t)||_2^2] + \mathrm{Tr}\{\mathbb{E}[\boldsymbol{W}^\intercal \boldsymbol{n}(t)\boldsymbol{n}(t)^\intercal \boldsymbol{W}]\} \\
&= \mathbb{E}[||(I_L - \boldsymbol{W}^\intercal \boldsymbol{G})\boldsymbol{s}(t)||_2^2] + \mathrm{Tr}\{\boldsymbol{W}^\intercal \boldsymbol{C}_\nu \boldsymbol{W}\}
\end{aligned}
$$

This last equality contained two parts that will be the localization bias and the variance of the output beamformer [126, 127]:

$$
MMSE = \underbrace{\mathbb{E}[||(I_L - \boldsymbol{W}^\intercal \boldsymbol{G})\boldsymbol{s}(t)||_2^2]}_{\text{Bias}} + \underbrace{\mathrm{Tr}\{\boldsymbol{W}^\intercal \boldsymbol{C}_\nu \boldsymbol{W}\}}_{\text{Variance}} \tag{A.14}
$$

A good estimator $\boldsymbol{W}^\intercal$ must minimize the mean square error, but above all, must be unbiased. In order to eliminate the bias, the constraint $\boldsymbol{W}^\intercal = I_L$ must be respected [205]. This can be formulated as a MMSE beamformer as we saw in section 2.2.2:

$$
\begin{aligned}
\underset{\boldsymbol{W}}{\text{minimize}} \quad & \mathrm{Tr}\{\boldsymbol{W}^\intercal \boldsymbol{C}_\nu \boldsymbol{W}\} \\
\text{subject to} \quad & \boldsymbol{W}^\intercal \boldsymbol{G} = I_L
\end{aligned} \tag{A.15}
$$

The main problem is that the solution is not easy to obtain, and often optimization strategies will sacrifice the bias constraint to reach a lower MMSE value. Thus we get a solution whose localization is not reliable, as we have seen in the results of simulations and on real data with the proposed LCMV (see section 2.4).

A.4 Full mathematical development of SABLE algorithm

In chapter 3, we presented a new method named SABLE. In this appendix, we offer the full mathematical development of this method. Let us recall the optimization

problem posed in SABLE:

$$\begin{aligned} \min_{\boldsymbol{w},\boldsymbol{\phi}} \quad & \boldsymbol{w}^\intercal \boldsymbol{C_x} \boldsymbol{w} + \lambda_1 ||\boldsymbol{X} - \boldsymbol{GS}||_\mathsf{F}^2 + \lambda_2 (||\boldsymbol{T\phi}||_1 + \lambda_3 ||\boldsymbol{\phi}||_1) \\ \text{s. t.} \quad & \boldsymbol{S} = \boldsymbol{\phi}\bar{\boldsymbol{s}},\ \bar{\boldsymbol{s}} = \boldsymbol{w}^\intercal \boldsymbol{X} \text{ and } \boldsymbol{w}^\intercal \boldsymbol{G\phi} = 1 \end{aligned} \tag{A.16}$$

This problem can be simplified by substituting parameters as such:

$$\begin{aligned} \min_{\boldsymbol{w},\boldsymbol{\phi}} \quad & \boldsymbol{w}^\intercal \boldsymbol{C_x} \boldsymbol{w} + \lambda_1 ||\boldsymbol{X} - \boldsymbol{G\phi w}^\intercal \boldsymbol{X}||_\mathsf{F}^2 + \lambda_2 (||\boldsymbol{T\phi}||_1 + \lambda_3 ||\boldsymbol{\phi}||_1) \\ \text{s. t.} \quad & \boldsymbol{w}^\intercal \boldsymbol{G\phi} = 1 \end{aligned} \tag{A.17}$$

Variable substitution to handle L_1 norm The L_1 cannot be derived directly and require the substitution by latent variables as suggested in the SISSY method:

$$\begin{aligned} \min_{\boldsymbol{w},\boldsymbol{\phi},\boldsymbol{\psi},\boldsymbol{\gamma}} \quad & \boldsymbol{w}^\intercal \boldsymbol{C_x} \boldsymbol{w} + \lambda_1 ||\boldsymbol{X} - \boldsymbol{G\phi w}^\intercal \boldsymbol{X}||_\mathsf{F}^2 + \lambda_2 (||\boldsymbol{\psi}||_1 + \lambda_3 ||\boldsymbol{\gamma}||_1) \\ \text{s.t.} \quad & \boldsymbol{w}^\intercal \boldsymbol{G\phi} = 1 \\ & \boldsymbol{\psi} = \mathbf{T}\boldsymbol{\phi} \\ & \boldsymbol{\gamma} = \boldsymbol{\phi} \end{aligned} \tag{A.18}$$

Formulation of the augmented Lagrangian According to [167], the minimization problem above can be solved using the ADMM algorithm. It is an algorithm intended to blend multiple dual ascent sub-problem with a superior convergence property brought by the multipliers method. The main idea is to successively update primal and dual variables of the augmented Lagrangian until we reach convergence. The augmented Lagrangian of (A.18) can be written as:

$$\begin{aligned} \mathcal{L}_{\rho,\lambda}(\boldsymbol{w},\boldsymbol{\phi},\boldsymbol{\psi},\boldsymbol{\gamma},u,\boldsymbol{v},\boldsymbol{z}) = \boldsymbol{w}^\intercal \boldsymbol{C_x} \boldsymbol{w} &+ \lambda_1 ||\boldsymbol{X} - \bar{\boldsymbol{g}} \boldsymbol{w}^\intercal \boldsymbol{X}||_\mathsf{F}^2 \\ &+ \lambda_2 ||\boldsymbol{\psi}||_1 + \lambda_2 \lambda_3 ||\boldsymbol{\gamma}||_1 + u_n (\boldsymbol{w}^\intercal \bar{\boldsymbol{g}} - 1) + \boldsymbol{v}^\intercal (\boldsymbol{T\phi} - \boldsymbol{\psi}) \\ &+ \boldsymbol{z}^\intercal (\boldsymbol{\phi} - \boldsymbol{\gamma}) + \frac{\rho}{2} (||\boldsymbol{w}^\intercal \bar{\boldsymbol{g}} - 1||_2^2 + ||\boldsymbol{T\phi} - \boldsymbol{\psi}||_2^2 + ||\boldsymbol{\phi} - \boldsymbol{\gamma}||_2^2) \end{aligned} \tag{A.19}$$

The ADMM algorithm of this augmented Lagragian is formulated as follows:

initialization;
while n<$itMax$ **do**

$$\boldsymbol{w}_{n+1} = \underset{\boldsymbol{w}}{\operatorname{argmin}} \{\mathcal{L}(\boldsymbol{w}, \boldsymbol{\phi}_n, \boldsymbol{\psi}_n, \boldsymbol{\gamma}_n, u_n, \boldsymbol{v}_n, \boldsymbol{z}_n)\};$$
$$\boldsymbol{\phi}_{n+1} = \underset{\boldsymbol{\phi}}{\operatorname{argmin}} \{\mathcal{L}(\boldsymbol{w}_{n+1}, \boldsymbol{\phi}, \boldsymbol{\psi}_n, \boldsymbol{\gamma}_n, u_n, \boldsymbol{v}_n, \boldsymbol{z}_n)\};$$
$$\boldsymbol{\psi}_{n+1} = \underset{\boldsymbol{\psi}}{\operatorname{argmin}} \{\mathcal{L}(\boldsymbol{w}_{n+1}, \boldsymbol{\phi}_{n+1}, \boldsymbol{\psi}, \boldsymbol{\gamma}_n, u_n, \boldsymbol{v}_n, \boldsymbol{z}_n)\};$$
$$\boldsymbol{\gamma}_{n+1} = \underset{\boldsymbol{\gamma}}{\operatorname{argmin}} \{\mathcal{L}(\boldsymbol{w}_{n+1}, \boldsymbol{\phi}_{n+1}, \boldsymbol{\psi}_{n+1}, \boldsymbol{\gamma}, u_n, \boldsymbol{v}_n, \boldsymbol{z}_n)\};$$
$$u_{n+1} = u_n + \rho(\boldsymbol{w}_{n+1}\boldsymbol{G}\boldsymbol{\phi}_{n+1} - 1);$$
$$\boldsymbol{v}_{n+1} = \boldsymbol{v}_n + \rho(\boldsymbol{\psi}_{n+1} - \mathbf{T}\boldsymbol{\phi}_{n+1});$$
$$\boldsymbol{z}_{n+1} = \boldsymbol{z}_n + \rho(\boldsymbol{\gamma}_{n+1} - \boldsymbol{\phi}_{n+1});$$

end

Algorithm 1: Pseudo-algorithm of ADMM for SABLE

The objective now is then to provide an analytical solution for $\boldsymbol{w}_{n+1}$, $\boldsymbol{\phi}_{n+1}$, $\boldsymbol{\psi}_{n+1}$ and $\boldsymbol{\gamma}_{n+1}$.

Analytical solution for $\boldsymbol{w}_{n+1}$: The notation $\bar{\boldsymbol{g}}_n = \boldsymbol{G}\boldsymbol{\phi}_n$ is used to simplify the writing since we do not optimize according to $\boldsymbol{\phi}$. This variable represents the mean lead-field vector of the distributed source.

$$\begin{aligned}
\boldsymbol{w}_{n+1} &= \underset{\boldsymbol{w}}{\operatorname{argmin}} \{\mathcal{L}(\boldsymbol{w}, \boldsymbol{\phi}_n, \boldsymbol{\psi}_n, \boldsymbol{\gamma}_n, u_n, \boldsymbol{v}_n, \boldsymbol{z}_n)\} \\
&= \underset{\boldsymbol{w}}{\operatorname{argmin}} \{\boldsymbol{w}^\mathsf{T}\boldsymbol{C}_x\boldsymbol{w} + \lambda_1 ||\boldsymbol{X} - \bar{\boldsymbol{g}}_n\boldsymbol{w}^\mathsf{T}\boldsymbol{X}||_\mathsf{F}^2 + u_n(\boldsymbol{w}^\mathsf{T}\bar{\boldsymbol{g}}_n - 1) + \frac{\rho}{2}||\boldsymbol{w}^\mathsf{T}\bar{\boldsymbol{g}}_n - 1||_2^2\} \quad \text{(A.20)} \\
&= \underset{\boldsymbol{w}}{\operatorname{argmin}} \{\boldsymbol{w}^\mathsf{T}\boldsymbol{C}_x\boldsymbol{w} + \lambda_1 \operatorname{Tr}\{\boldsymbol{X}^\mathsf{T}\boldsymbol{w}\bar{\boldsymbol{g}}_n^\mathsf{T}\bar{\boldsymbol{g}}_n\boldsymbol{w}^\mathsf{T}\boldsymbol{X}\} - 2\operatorname{Tr}\{\boldsymbol{X}^\mathsf{T}\bar{\boldsymbol{g}}_n\boldsymbol{w}^\mathsf{T}\boldsymbol{X}\} + u_n(\boldsymbol{w}^\mathsf{T}\bar{\boldsymbol{g}}_n - 1) \\
&\qquad + \frac{\rho}{2}||\boldsymbol{w}^\mathsf{T}\bar{\boldsymbol{g}}_n - 1||_2^2\}
\end{aligned}$$

For the second and third terms, we use the derivative solution (118) in [206] and the property of the trace presented in section 2.5.1, respectively. It goes as follows:

$$\begin{aligned}
\frac{\delta}{\delta\boldsymbol{w}}\{\operatorname{Tr}\{\boldsymbol{X}^\mathsf{T}\boldsymbol{w}\bar{\boldsymbol{g}}_n^\mathsf{T}\bar{\boldsymbol{g}}_n\boldsymbol{w}^\mathsf{T}\boldsymbol{X}\}\} = \frac{\delta}{\delta\boldsymbol{w}}\{\operatorname{Tr}\{\boldsymbol{A}\boldsymbol{w}\mathbf{B}\boldsymbol{w}^\mathsf{T}\boldsymbol{C}\}\} &= \boldsymbol{A}^\mathsf{T}\boldsymbol{C}^\mathsf{T}\boldsymbol{w}\mathbf{B}^\mathsf{T} + \boldsymbol{C}\boldsymbol{A}\boldsymbol{w}\mathbf{B} \quad \text{(A.21)} \\
&= 2\boldsymbol{X}\boldsymbol{X}^\mathsf{T}\boldsymbol{w}\bar{\boldsymbol{g}}_n^\mathsf{T}\bar{\boldsymbol{g}}_n
\end{aligned}$$

$$\frac{\delta}{\delta\boldsymbol{w}}\{\operatorname{Tr}\{\boldsymbol{X}^\mathsf{T}\bar{\boldsymbol{g}}_n\boldsymbol{w}^\mathsf{T}\boldsymbol{X}\}\} = \frac{\delta}{\delta\boldsymbol{w}}\{\operatorname{Tr}\{\boldsymbol{X}\boldsymbol{X}^\mathsf{T}\bar{\boldsymbol{g}}_n\boldsymbol{w}^\mathsf{T}\}\} = \boldsymbol{X}\boldsymbol{X}^\mathsf{T}\bar{\boldsymbol{g}}_n \quad \text{(A.22)}$$

With these definitions and by using the vectorization operator, the following derivation is valid to find the minimum argument of $\boldsymbol{w}$:

$$
\begin{aligned}
0 &= 2\boldsymbol{C_x}\boldsymbol{w}_{n+1} + (2\lambda_1 \bar{\boldsymbol{g}}_n^\intercal \bar{\boldsymbol{g}}_n \otimes \boldsymbol{X}\boldsymbol{X}^\intercal)\boldsymbol{w}_{n+1} - 2\boldsymbol{X}\boldsymbol{X}^\intercal \bar{\boldsymbol{g}}_n + u_n \bar{\boldsymbol{g}}_n + \rho \bar{\boldsymbol{g}}_n \bar{\boldsymbol{g}}_n^\intercal \boldsymbol{w}_{n+1} - \rho \bar{\boldsymbol{g}}_n \\
[2\boldsymbol{C_x} + 2\lambda_1 \bar{\boldsymbol{g}}_n^\intercal \bar{\boldsymbol{g}}_n \otimes \boldsymbol{X}\boldsymbol{X}^\intercal + \rho \bar{\boldsymbol{g}}_n \bar{\boldsymbol{g}}_n^\intercal] \boldsymbol{w}_{n+1} &= 2\boldsymbol{X}\boldsymbol{X}^\intercal \bar{\boldsymbol{g}}_n + \bar{\boldsymbol{g}}_n(u_n - \rho) \qquad \text{(A.23)} \\
\boldsymbol{w}_{n+1} &= [2\boldsymbol{C_x} + 2\lambda_1 \bar{\boldsymbol{g}}_n^\intercal \bar{\boldsymbol{g}}_n \otimes \boldsymbol{X}\boldsymbol{X}^\intercal + 2\rho \bar{\boldsymbol{g}}_n \bar{\boldsymbol{g}}_n^\intercal]^{-1} [2\lambda_1 \boldsymbol{X}\boldsymbol{X}^\intercal \bar{\boldsymbol{g}}_n + \bar{\boldsymbol{g}}_n(u_n - \rho)]
\end{aligned}
$$

Analytical solution for $\boldsymbol{\phi}_{n+1}$:

$$
\begin{aligned}
\boldsymbol{\phi}_{n+1} &= \underset{\boldsymbol{\phi}}{\operatorname{argmin}} \{\mathcal{L}(\boldsymbol{w}_{n+1}, \boldsymbol{\phi}, \boldsymbol{\psi}_n, \boldsymbol{\gamma}_n, u_n, \boldsymbol{v}_n, \boldsymbol{z}_n)\} \qquad \text{(A.24)} \\
&= \underset{\boldsymbol{\phi}}{\operatorname{argmin}} \{\lambda_1 ||\boldsymbol{X} - \boldsymbol{G}\boldsymbol{\phi}\boldsymbol{w}_{n+1}^\intercal \boldsymbol{X}||_{\mathsf{F}}^2 + u_n(\boldsymbol{w}_{n+1}^\intercal \boldsymbol{G}\boldsymbol{\phi} - 1) + \boldsymbol{v}_n^\intercal(\boldsymbol{T}\boldsymbol{\phi} - \boldsymbol{\psi}_n) + \boldsymbol{z}_n^\intercal(\boldsymbol{\phi} - \boldsymbol{\gamma}_n) \\
&\qquad + \frac{\rho}{2}(||\boldsymbol{w}_{n+1}^\intercal \boldsymbol{G}\boldsymbol{\phi} - 1||_2^2 + ||\boldsymbol{T}\boldsymbol{\phi} - \boldsymbol{\psi}_n||_2^2 + ||\boldsymbol{\phi} - \boldsymbol{\gamma}_n||_2^2)\}
\end{aligned}
$$

Using the properties used to compute $\boldsymbol{w}_{n+1}$, we can easily obtain the following equality:

$$
\begin{aligned}
0 &= 2\lambda_1 \boldsymbol{w}_{n+1}^\intercal \boldsymbol{X}\boldsymbol{X}^\intercal \boldsymbol{w}_{n+1} \boldsymbol{\phi}_{n+1} \boldsymbol{G}^\intercal \boldsymbol{G} - 2\lambda_1 \boldsymbol{G}^\intercal \boldsymbol{X}\boldsymbol{X}^\intercal \boldsymbol{w}_{n+1} + u_n \boldsymbol{G}^\intercal \boldsymbol{w}_{n+1} \\
&\quad + \rho \boldsymbol{G}^\intercal \boldsymbol{w}_{n+1}(\boldsymbol{w}_{n+1}^\intercal \boldsymbol{G}\boldsymbol{\phi}_{n+1} - 1) + \boldsymbol{T}^\intercal \boldsymbol{v}_n + \boldsymbol{z}_n + \rho \boldsymbol{T}^\intercal(\boldsymbol{T}\boldsymbol{\phi}_{n+1} - \boldsymbol{\psi}_n) + \rho(\boldsymbol{\phi}_{n+1} - \boldsymbol{\gamma}_n) \\
[2\lambda_1 \boldsymbol{G}^\intercal \boldsymbol{G} \otimes \boldsymbol{w}_{n+1}^\intercal \boldsymbol{X}\boldsymbol{X}^\intercal \boldsymbol{w}_{n+1} &+ \rho(\boldsymbol{T}^\intercal \boldsymbol{T} + I_d + \boldsymbol{G}^\intercal \boldsymbol{w}_{n+1} \boldsymbol{w}_{n+1}^\intercal \boldsymbol{G})] \boldsymbol{\phi}_{n+1} = \\
&\quad 2\lambda_1 \boldsymbol{G}^\intercal \boldsymbol{X}\boldsymbol{X}^\intercal \boldsymbol{w}_{n+1} + \boldsymbol{G}^\intercal \boldsymbol{w}_{n+1}(\rho - u_n) + \boldsymbol{T}^\intercal(\rho \boldsymbol{\psi}_n - \boldsymbol{v}_n) + (\rho \boldsymbol{\gamma}_n - \boldsymbol{z}_n) \qquad \text{(A.25)} \\
\boldsymbol{\phi}_{n+1} &= [2\lambda_1 \boldsymbol{G}^\intercal \boldsymbol{G} \otimes \boldsymbol{w}_{n+1}^\intercal \boldsymbol{X}\boldsymbol{X}^\intercal \boldsymbol{w}_{n+1} + \rho(\boldsymbol{T}^\intercal \boldsymbol{T} + I_d + \boldsymbol{G}^\intercal \boldsymbol{w}_{n+1} \boldsymbol{w}_{n+1}^\intercal \boldsymbol{G})]^{-1} \\
&\quad [\lambda_1 \boldsymbol{G}^\intercal \boldsymbol{X}\boldsymbol{X}^\intercal \boldsymbol{w}_{n+1} - \boldsymbol{G}^\intercal \boldsymbol{w}_{n+1}(\rho - u_n) + \boldsymbol{T}^\intercal(\rho \boldsymbol{\psi}_n - \boldsymbol{v}_n) + (\rho \boldsymbol{\gamma}_n - \boldsymbol{z}_n)]
\end{aligned}
$$

Analytical solution for $\boldsymbol{\psi}_{n+1}$: Finding the global minimum of a L_1 norm is not straightforward. We need to reformulate the augmented Lagrangian in a form that can be compatible with the proximal operator see (3.8) in section 3.2.2.

$$
\begin{aligned}
\boldsymbol{\psi}_{n+1} &= \underset{\boldsymbol{\psi}}{\operatorname{argmin}} \{\mathcal{L}(\boldsymbol{w}_{n+1}, \boldsymbol{\phi}_{n+1}, \boldsymbol{\psi}, \boldsymbol{\gamma}_n, u_n, \boldsymbol{v}_n, \boldsymbol{z}_n)\} \\
&= \underset{\boldsymbol{\psi}}{\operatorname{argmin}} \{\lambda_2 ||\boldsymbol{\psi}||_1 - \boldsymbol{v}_n^\intercal \boldsymbol{\psi} + \frac{\rho}{2} ||\boldsymbol{T}\boldsymbol{\phi}_{n+1} - \boldsymbol{\psi}||_2^2\} \\
&= \underset{\boldsymbol{\psi}}{\operatorname{argmin}} \{\lambda_2 ||\boldsymbol{\psi}||_1 - \boldsymbol{v}_n^\intercal \boldsymbol{\psi} + \frac{\rho}{2}(\boldsymbol{\phi}_{n+1}^\intercal \boldsymbol{T}^\intercal \boldsymbol{T}\boldsymbol{\phi}_{n+1} - 2\boldsymbol{\psi}^\intercal \boldsymbol{T}\boldsymbol{\phi}_{n+1} + \boldsymbol{\psi}^\intercal \boldsymbol{\psi})\} \qquad \text{(A.26)} \\
&= \underset{\boldsymbol{\psi}}{\operatorname{argmin}} \{\lambda_2 ||\boldsymbol{\psi}||_1 + \frac{\rho}{2}(\boldsymbol{\phi}_{n+1}^\intercal \boldsymbol{T}^\intercal \boldsymbol{T}\boldsymbol{\phi}_{n+1} - 2\boldsymbol{\psi}^\intercal \boldsymbol{T}\boldsymbol{\phi}_{n+1} + \boldsymbol{\psi}^\intercal \boldsymbol{\psi} - \frac{2}{\rho} \boldsymbol{v}_n^\intercal \boldsymbol{\psi})\} \\
&= \underset{\boldsymbol{\psi}}{\operatorname{argmin}} \{\lambda_2 ||\boldsymbol{\psi}||_1 + \frac{\rho}{2}(\boldsymbol{\phi}_{n+1}^\intercal \boldsymbol{T}^\intercal \boldsymbol{T}\boldsymbol{\phi}_{n+1} - 2\boldsymbol{\psi}^\intercal \boldsymbol{T}\boldsymbol{\phi}_{n+1} + \boldsymbol{\psi}^\intercal \boldsymbol{\psi} - \frac{2}{\rho} \boldsymbol{v}_n^\intercal \boldsymbol{\psi}) + \frac{2}{\rho} \boldsymbol{v}_n \boldsymbol{T}\boldsymbol{\phi}_{n+1} + \frac{1}{\rho^2} \boldsymbol{v}_n^\intercal \boldsymbol{v}_n)\} \\
&= \underset{\boldsymbol{\psi}}{\operatorname{argmin}} \{\lambda_2 ||\boldsymbol{\psi}||_1 + \frac{\rho}{2} ||\boldsymbol{\psi} - (\boldsymbol{T}\boldsymbol{\phi}_{n+1} + \frac{1}{\rho} \boldsymbol{v}_n^\intercal)||_2^2\}
\end{aligned}
$$

$$= \underset{\psi}{\operatorname{argmin}} \{||\psi||_1 + \frac{\rho}{2\lambda_2}||\psi - (\boldsymbol{T}\phi_{n+1} + \frac{1}{\rho}\boldsymbol{v}_n^{\mathsf{T}})||_2^2\}$$
$$= \operatorname{prox}_{\frac{\lambda_2}{\rho}}\left(\boldsymbol{T}\phi_{n+1} + \frac{1}{\rho}\boldsymbol{v}_n^{\mathsf{T}}\right)$$

Analytical solution for γ_{n+1}:

$$\begin{aligned}\gamma_{n+1} &= \underset{\gamma}{\operatorname{argmin}} \{\mathcal{L}(\boldsymbol{w}_{n+1}, \phi_{n+1}, \psi_{n+1}, \gamma, u_n, \boldsymbol{v}_n, \boldsymbol{z}_n)\} \\ &= \underset{\gamma}{\operatorname{argmin}} \{\lambda_2\lambda_3||\gamma||_1 - \boldsymbol{z}^{\mathsf{T}}\gamma + ||\phi - \gamma||_2^2)\} \\ &= \operatorname{prox}_{\frac{\lambda_2\lambda_3}{\rho}}\left(\phi_{n+1} + \frac{1}{\rho}\boldsymbol{z}_n\right)\end{aligned} \tag{A.27}$$

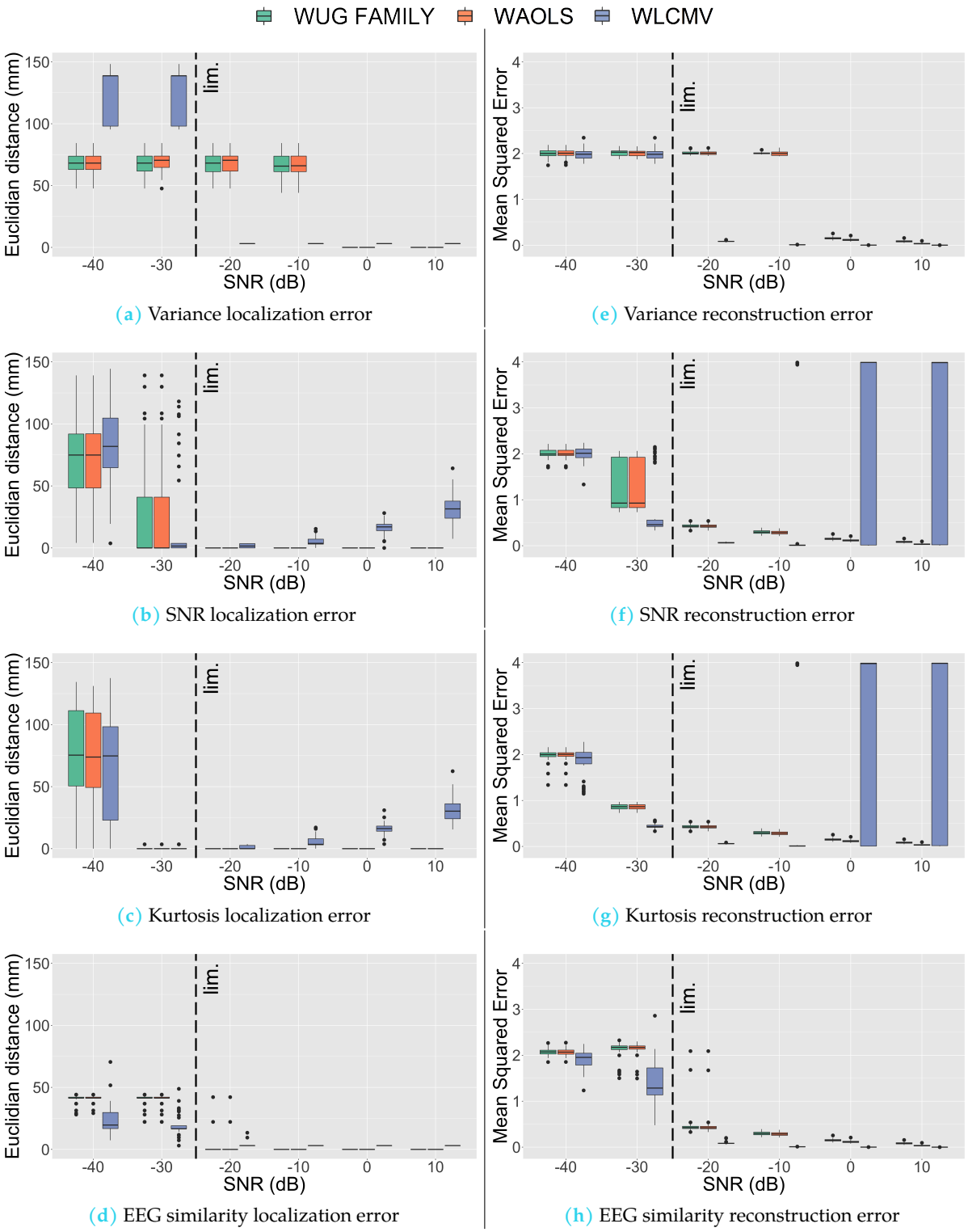
Appendix B

Supplementary results

This appendix gathers all the additional results that were not presented in the main part.

B.1 Additional results for classic beamforming

This part contains additional results to the beamforming algorithms efficiency study provided in section 2.4.2.



(a) Variance localization error

(e) Variance reconstruction error

(b) SNR localization error

(f) SNR reconstruction error

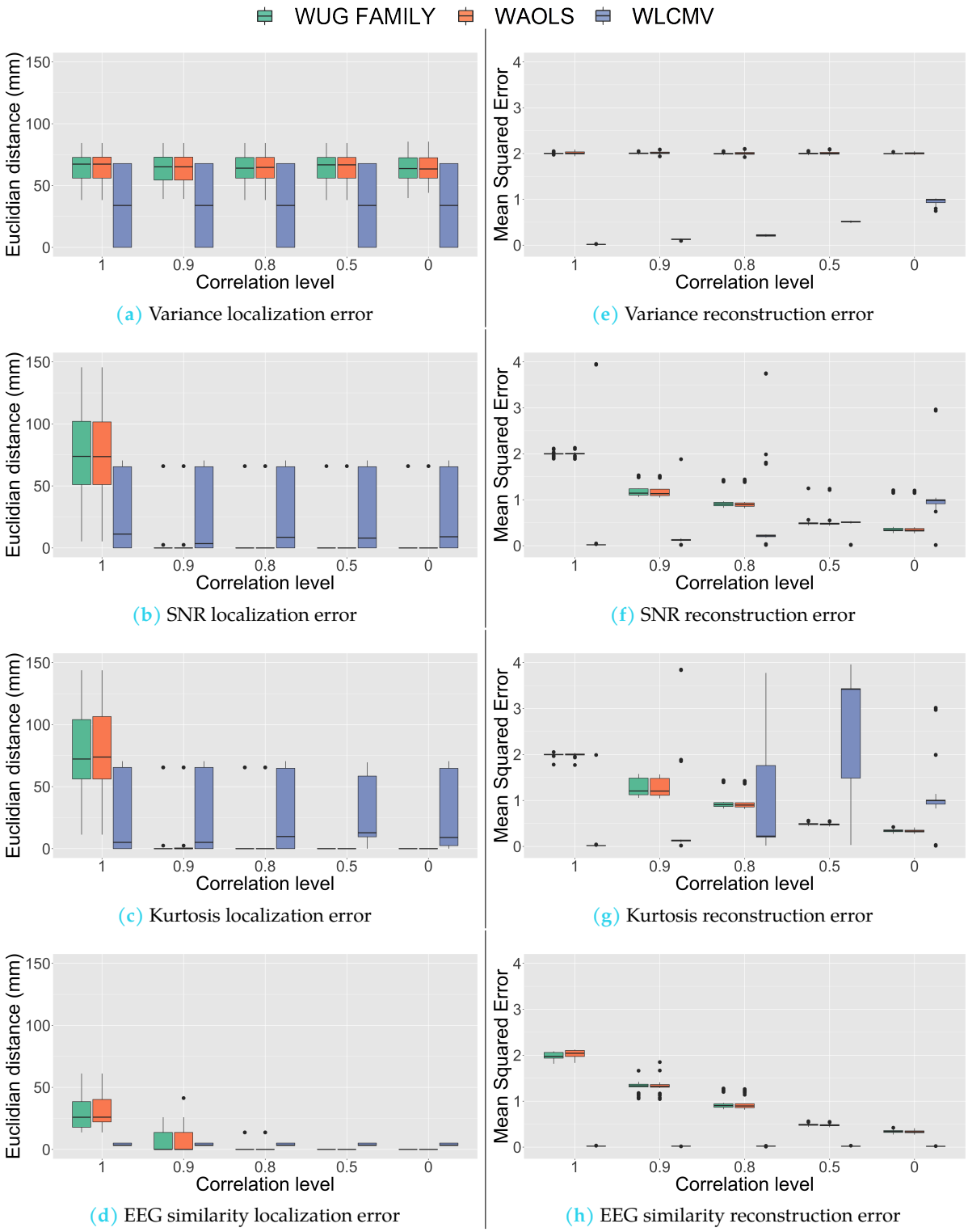
(c) Kurtosis localization error

(g) Kurtosis reconstruction error

(d) EEG similarity localization error

(h) EEG similarity reconstruction error

Figure B.1: Effect of different beamforming methods with the Shrinkage preprocessing method in the single source scenario: localization results are on the left and signal reconstruction on the right. The dashed line represents the limit of baseline identification.



(a) Variance localization error

(e) Variance reconstruction error

(b) SNR localization error

(f) SNR reconstruction error

(c) Kurtosis localization error

(g) Kurtosis reconstruction error

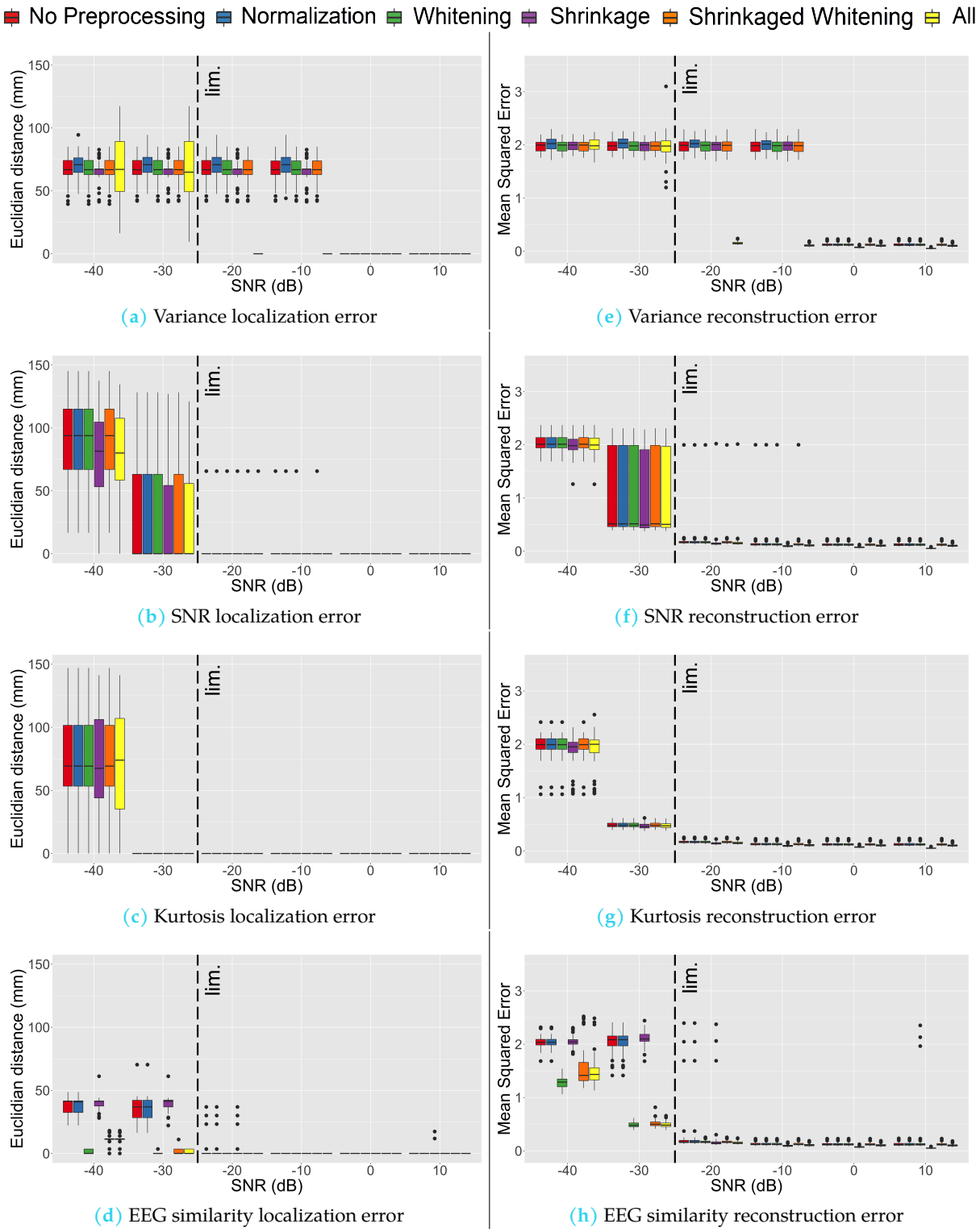
(d) EEG similarity localization error

(h) EEG similarity reconstruction error

Figure B.2: Effect of different beamforming methods with the Shrinkage preprocessing method in the two sources scenario: localization results are on the left and signal reconstruction on the right.

B.2 Additional results for Temporal-Smoothing-based beamforming

This part contains all results for the Temporal-Smoothing-based beamforming on the data of experiment 1 and 2 in section 2.4.2 and organized in the same manner.



(a) Variance localization error

(e) Variance reconstruction error

(b) SNR localization error

(f) SNR reconstruction error

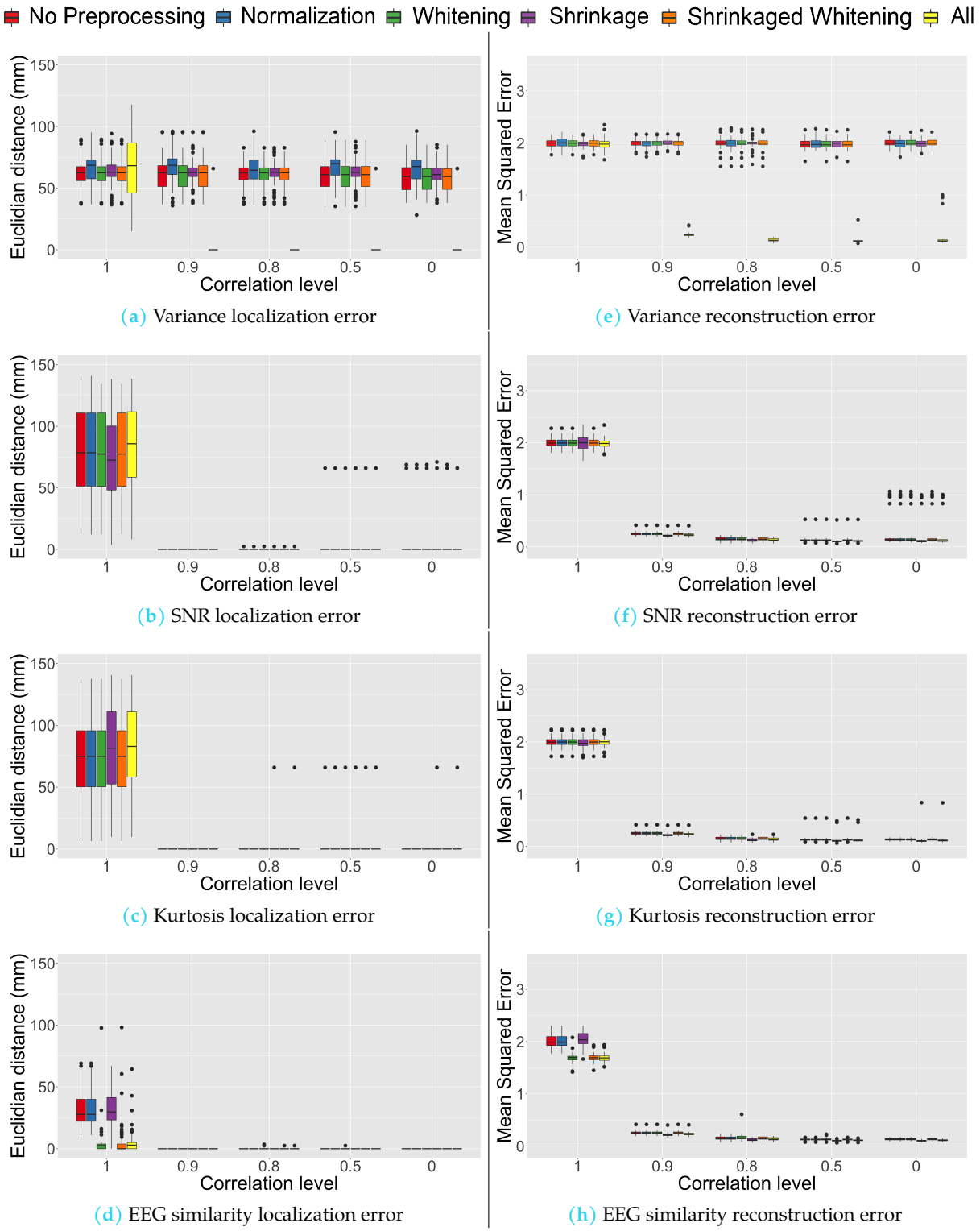
(c) Kurtosis localization error

(g) Kurtosis reconstruction error

(d) EEG similarity localization error

(h) EEG similarity reconstruction error

Figure B.3: Effect of different preprocessing methods on the Temporal-Smoothing-based UG beamformer for one source: localization results are on the left and signal reconstruction on the right. The dashed line represents the limit of baseline identification.



(a) Variance localization error

(e) Variance reconstruction error

(b) SNR localization error

(f) SNR reconstruction error

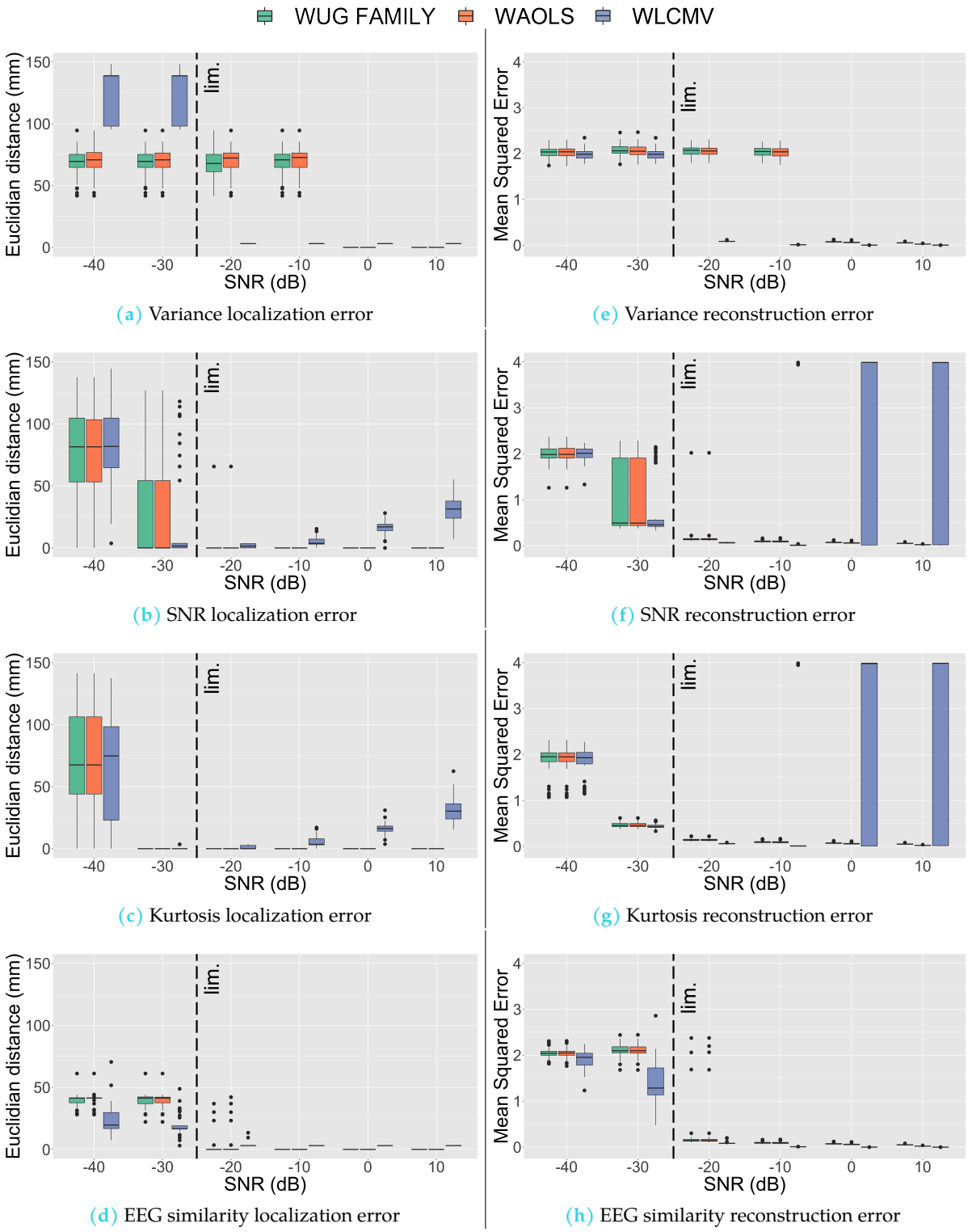
(c) Kurtosis localization error

(g) Kurtosis reconstruction error

(d) EEG similarity localization error

(h) EEG similarity reconstruction error

Figure B.4: Effect of different preprocessing methods on the Temporal-Smoothing-based UG beamformer for two sources depending on the correlation level: localization results are on the left and signal reconstruction on the right.



(a) Variance localization error

(e) Variance reconstruction error

(b) SNR localization error

(f) SNR reconstruction error

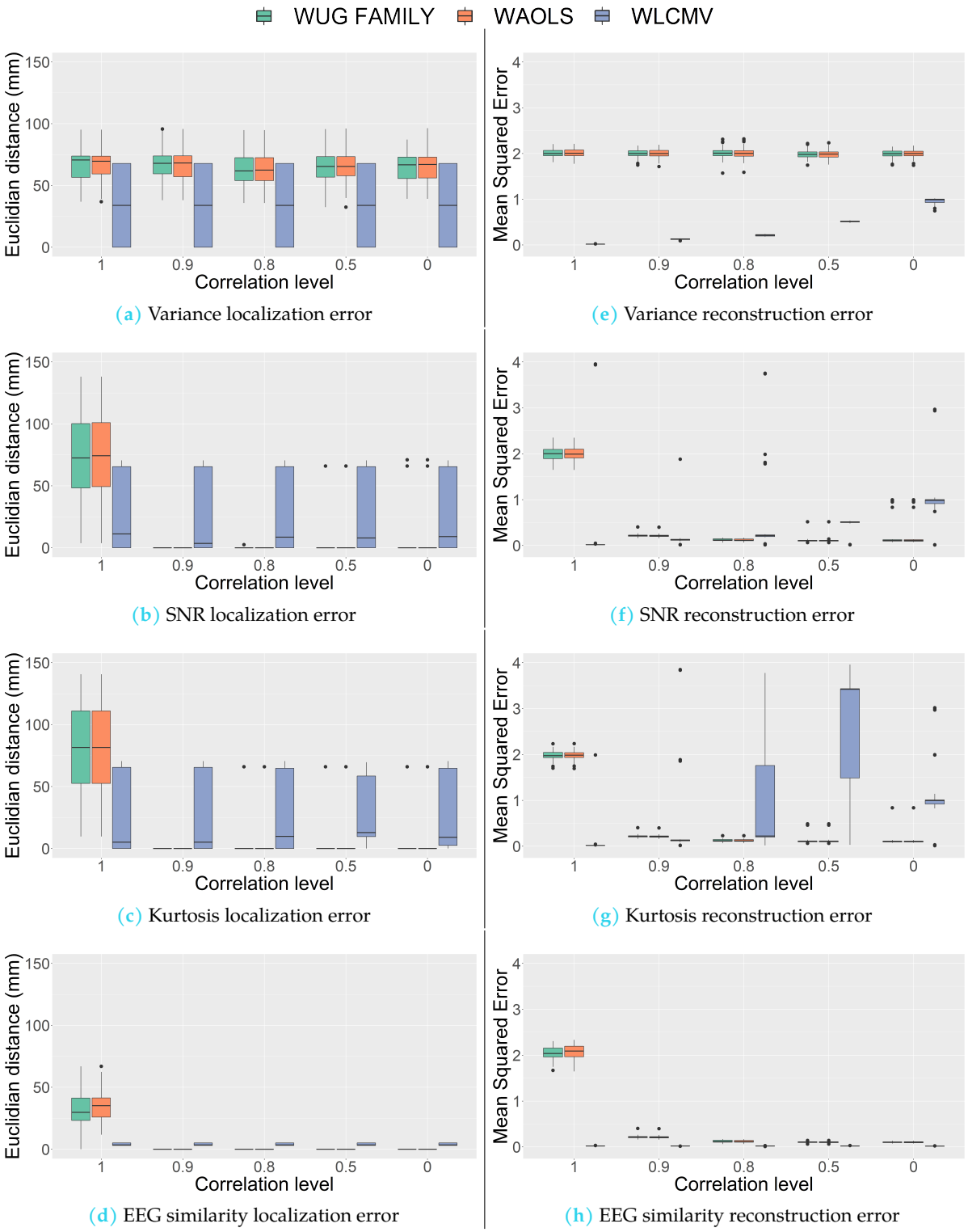
(c) Kurtosis localization error

(g) Kurtosis reconstruction error

(d) EEG similarity localization error

(h) EEG similarity reconstruction error

Figure B.5: Effect of different Temporal-Smoothing-based beamforming methods with the Shrinkage preprocessing method in the single source scenario: localization results are on the left and signal reconstruction on the right. The dashed line represents the limit of baseline identification.



(a) Variance localization error

(e) Variance reconstruction error

(b) SNR localization error

(f) SNR reconstruction error

(c) Kurtosis localization error

(g) Kurtosis reconstruction error

(d) EEG similarity localization error

(h) EEG similarity reconstruction error

Figure B.6: Effect of different Temporal-Smoothing-based beamforming methods with the Shrinkage preprocessing method in the two sources scenario: localization results are on the left and signal reconstruction on the right.

B.3 Additional results for comparative study

This part contains additional simulation results to the comparative study of chapter 3.

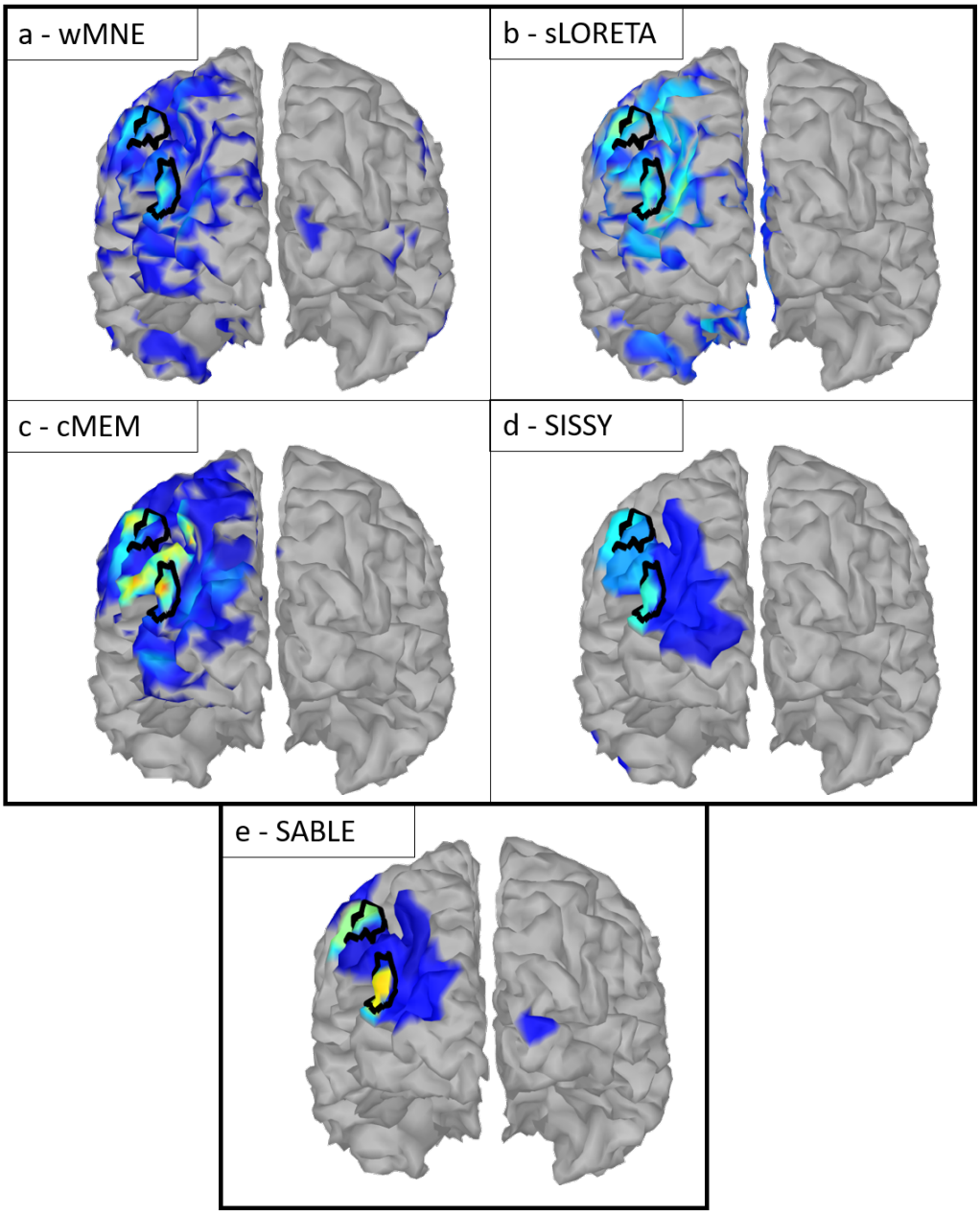


Figure B.7: Source localization results thresholded with the best ratio at the output of wMNE, sLORETA, cMEM, SISSY and SABLE for an SNR value of approximately -10dB (back). The patches outline is represented by a plain black line.

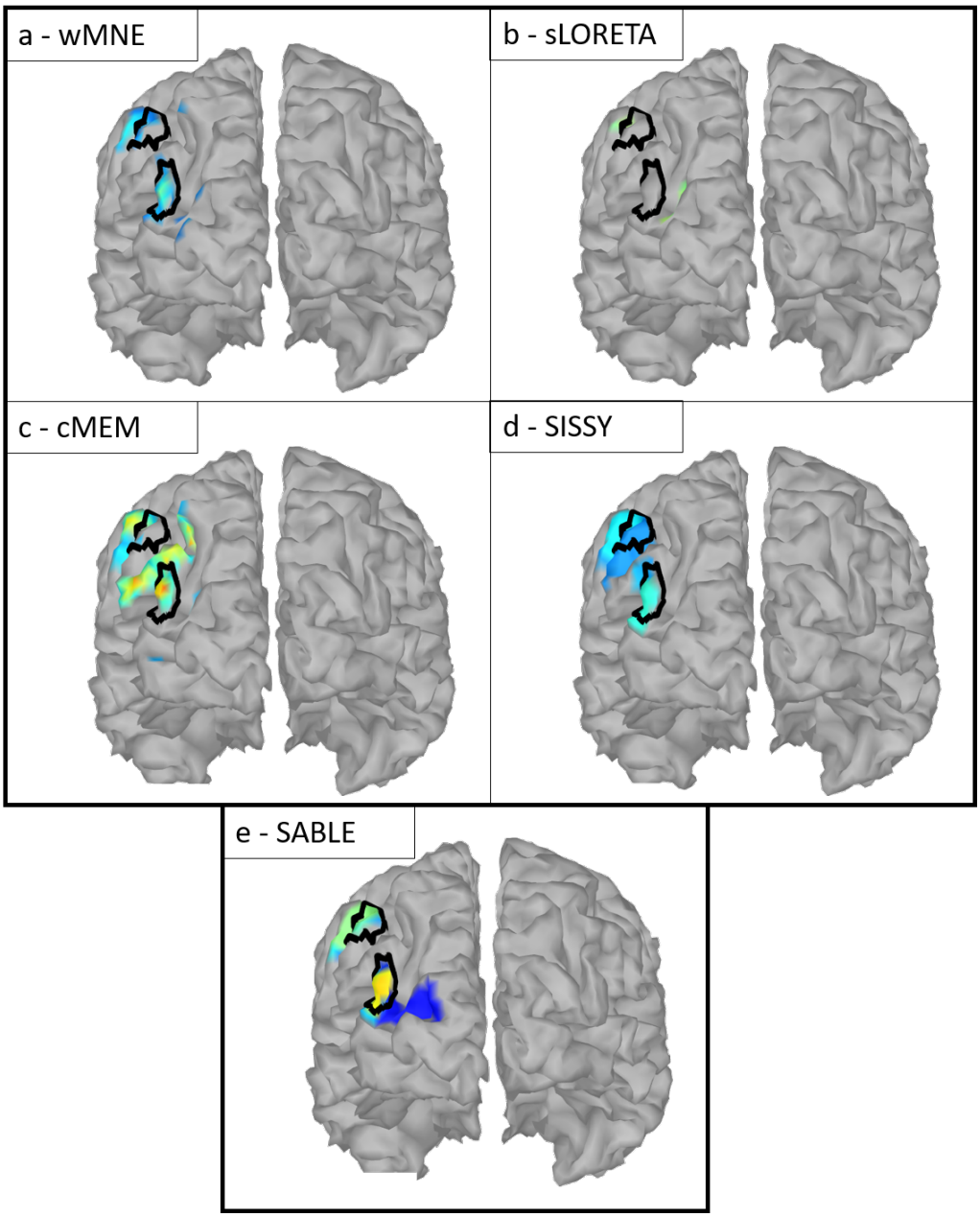


Figure B.8: Source localization results thresholded with the first 150 indexes at the output of wMNE, sLORETA, cMEM, SISSY and SABLE for an SNR value of approximately -10dB (back). The patches outline is represented by a plain black line.

List of publications

Parts of the work presented in this thesis can be associated with the following publications:

International conference abstract

— **Berraute, P.**, Albera, L., Kachenoura, A., Rochette, M., & Merlet, I., "How data preprocessing can improve brain source imaging based on beamforming." *Virtual Physiological Human* (pp. 1-2), 2020.

International conference papers

— El Houari, K., Kachenoura, A., **Berraute, P.**, Bensaid, S., Karfoul, A., Boichon-Grivot, C., Hernandez, A., Albera, L., Rochette, M., "Investigating Transmembrane Current Source Formulation for Solving the ECG Inverse Problem." *IEEE 10th Sensor Array and Multichannel Signal Processing Workshop* (*SAM*), 2018.

— **Berraute, P.**, Albera, L., Rochette, M., Merlet, I., & Kachenoura, A., "How beamforming can solve the brain source imaging problem ?" *IEEE Statistical Signal Processing Workshop 2021*, 2021.

Bibliography

[1] N. Galatsanos and A. K. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," *IEEE Transactions on Image Processing*, vol. 1, pp. 322–336, jul 1992.

[2] P. Vallet, *Random matrices and applications to statistical signal processing*. PhD thesis, Université Paris-Est, 2012.

[3] P. Berraute, L. Albera, A. Kachenoura *et al.*, "How data preprocessing can improve brain source imaging based on beamforming," in *Virtual Physiological Human*, (Paris), pp. 1–2, aug 2020.

[4] P. Berraute, L. Albera, M. Rochette *et al.*, "How beamforming can solve the brain source imaging problem ?," in *IEEE Statistical Signal Processing Workshop 2021*, (Rio de Janeiro), IEEE, 2021.

[5] J. E. Mendoza and A. L. Foundas, *Clinical neuroanatomy: A neurobehavioral approach*. Elsevier, 2008.

[6] A. B. Butler and W. Hodos, *Comparative Vertebrate Neuroanatomy*. Hoboken, NJ, USA: John Wiley & Sons, Inc., aug 2005.

[7] R. Carpenter and B. Reddi, *Neurophysiology: A conceptual approach*. Hodder Arnold, 2012.

[8] R. G. Northcutt, "Understanding Vertebrate Brain Evolution," *Integrative and Comparative Biology*, vol. 42, pp. 743–756, aug 2002.

[9] J. D. Newman, "Affective Neuroscience: The Foundations of Human and Animal Emotions.," *The Journal of Nervous & Mental Disease*, vol. 188, p. 59, jan 2000.

[10] F. Netter, *Atlas of Human Anatomy*, vol. 7. Elsevier, 2019.

[11] K. Brodmann, *Vergleichende Lokalisationslehre der Großhirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig: visual library, 1909.

[12] R. Baxter, N. Hastings, A. Law *et al.*, *Brodmann's Localisation in the Cerebral Cortex*, vol. 39. Boston, MA: Springer US, 2005.

[13] R. S. Petersen and M. E. Diamond, "Topographic Maps in the Brain," in *Encyclopedia of Life Sciences*, pp. 1–8, Chichester, UK: John Wiley & Sons, Ltd, mar 2002.

[14] L. Bonilha, E. Gleichgerrcht, T. Nesland *et al.*, "Gray matter axonal connectivity maps," *Frontiers in Psychiatry*, vol. 6, pp. 1–8, mar 2015.

[15] S. Herculano-Houzel, "The human brain in numbers: A linearly scaled-up primate brain," *Frontiers in Human Neuroscience*, vol. 3, pp. 1–11, nov 2009.

[16] H. Zeng and J. R. Sanes, "Neuronal cell-type classification: Challenges, opportunities and the path forward," *Nature Reviews Neuroscience*, vol. 18, no. 9, pp. 530–546, 2017.

[17] M. Raghavan, D. Fee, and P. E. Barkhaus, "Chapter 1 - Generation and propagation of the action potential," in *Clinical Neurophysiology: Basis and Technical Aspects* (K. H. Levin and P. Chauvel, eds.), vol. 160 of *Handbook of Clinical Neurology*, pp. 3–22, Elsevier, 2019.

[18] D. Džaja, A. Hladnik, I. Bičanić *et al.*, "Neocortical calretinin neurons in primates: Increase in proportion and microcircuitry structure," *Frontiers in Neuroanatomy*, vol. 8, pp. 1–6, sep 2014.

[19] F. Lopes da Silva, "EEG and MEG: Relevance to neuroscience," *Neuron*, vol. 80, no. 5, pp. 1112–1128, 2013.

[20] A. M. Feyissa and W. O. Tatum, "Chapter 7 - adult eeg," in *Clinical Neurophysiology: Basis and Technical Aspects* (K. H. Levin and P. Chauvel, eds.), vol. 160 of *Handbook of Clinical Neurology*, pp. 103–124, Elsevier, 2019.

[21] S. M. Stufflebeam and B. R. Rosen, "Mapping Cognitive Function," *Neuroimaging Clinics of North America*, vol. 17, pp. 469–484, nov 2007.

[22] N. K. Al-qazzaz, S. Hamid, B. Ali *et al.*, "Role of EEG as biomarker in the early detection and classification of dementia.," *Scientific World Journal*, 2014.

[23] S. Noachtar and J. Rémi, "The role of EEG in epilepsy: A critical review," *Epilepsy and Behavior*, vol. 15, no. 1, pp. 22–33, 2009.

[24] World Health Organization, "WHO | Epilepsy: a public health imperative," tech. rep., World Health Organization, 2019.

[25] H. Lüders, G. F. B. Vaca, N. Akamatsu *et al.*, "Classification of paroxysmal events and the four-dimensional epilepsy classification system," *Epileptic Disorders*, vol. 21, no. 1, pp. 1–29, 2019.

[26] A. Crespel, P. Gélisse, M. Bureau *et al.*, *Atlas d'électroencéphalographie : Tome 2, Les épilepsies, de l'EEG aux syndromes*. John Libbey Eurotext, 2006.

[27] D. Safieddine, *Apport des méthodes statistiques et déterministes à la réduction d ' artefacts et de bruits dans les signaux EEG*. PhD thesis, Université de Rennes 1, 2012.

[28] H. Haute Autorité de Santé, "Ald n°9 - epilepsies graves," tech. rep., Haute Autorité de Santé, H.A.S., Saint-Denis La Plaine, 2016.

[29] C. M. Michel, "Chapter 12 - High-resolution EEG," in *Clinical Neurophysiology: Basis and Technical Aspects* (K. H. Levin and P. Chauvel, eds.), vol. 160 of *Handbook of Clinical Neurology*, pp. 185–201, Elsevier, 2019.

[30] H. Hallez, B. Vanrumste, R. Grech *et al.*, "Review on solving the forward problem in EEG source analysis," *Journal of NeuroEngineering and Rehabilitation*, vol. 4, p. 46, dec 2007.

[31] F. Wendling, A. Hernandez, J. J. Bellanger *et al.*, "Interictal to ictal transition in human temporal lobe epilepsy: Insights from a computational model of intracerebral EEG," *Journal of Clinical Neurophysiology*, vol. 22, no. 5, pp. 343–356, 2005.

[32] F. Wendling, P. Benquet, F. Bartolomei *et al.*, "Computational models of epileptiform activity," *Journal of Neuroscience Methods*, vol. 260, pp. 233–251, 2016.

[33] F. Wendling, J. J. Bellanger, F. Bartolomei *et al.*, "Relevance of nonlinear lumped-parameter models in the analysis of depth-EEG epileptic signals," *Biological Cybernetics*, vol. 83, no. 4, pp. 367–378, 2000.

[34] D. Cosandier-Rimélé, I. Merlet, J. M. Badier *et al.*, "The neuronal sources of EEG: Modeling of simultaneous scalp and intracerebral recordings in epilepsy," *NeuroImage*, vol. 42, no. 1, pp. 135–146, 2008.

[35] F. Neugebauer, G. Möddel, S. Rampp *et al.*, "The Effect of Head Model Simplification on Beamformer Source Localization," *Frontiers in Neuroscience*, vol. 11, p. 625, nov 2017.

[36] A. Dale, B. Fischl, and M. I. Sereno, "Cortical surface-based analysis: I. segmentation and surface reconstruction," *NeuroImage*, vol. 9, no. 2, pp. 179 – 194, 1999.

[37] B. Fischl, M. I. Sereno, and A. Dale, "Cortical surface-based analysis: Ii: Inflation, flattening, and a surface-based coordinate system," *NeuroImage*, vol. 9, no. 2, pp. 195 – 207, 1999.

[38] B. Fischl, A. Liu, and A. M. Dale, "Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex," *IEEE Medical Imaging*, vol. 20, pp. 70–80, jan 2001.

[39] H. Becker, "Denoising , separation and localization of EEG sources in the context of epilepsy," *HAL Archives* (*PhD thesis*), 2014.

[40] F. Tadel, S. Baillet, J. C. Mosher *et al.*, "Brainstorm: A user-friendly application for MEG/EEG analysis," *Computational Intelligence and Neuroscience*, 2011.

[41] A. Gramfort, *Mapping, timing and tracking cortical activations with MEG and EEG: Methods and application to human vision*. PhD thesis, Ecole nationale superieure des telecommunications - ENST, oct 2010.

[42] A. Gramfort, T. Papadopoulo, E. Olivi *et al.*, "OpenMEEG: Opensource software for quasistatic bioelectromagnetics," *BioMedical Engineering Online*, vol. 9, no. 1, p. 45, 2010.

[43] A. Ferreira, W. C. Celeste, F. A. Cheein *et al.*, "Human-machine interfaces based on EMG and EEG applied to robotic systems.," *Journal of neuroengineering and rehabilitation*, vol. 5, p. 10, mar 2008.

[44] J. Zhang, B. Wang, C. Zhang *et al.*, "An EEG/EMG/EOG-based multimodal human-machine interface to real-time control of a soft robot hand," *Frontiers in Neurorobotics*, vol. 13, p. 7, 2019.

[45] H. Becker, L. Albera, P. Comon *et al.*, "Brain-source imaging: From sparse to tensor models," *IEEE Signal Processing Magazine*, vol. 32, no. 6, pp. 100–112, 2015.

[46] M. Scherg and D. Von Cramon, "Evoked dipole source potentials of the human auditory cortex," *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, vol. 65, pp. 344–360, sep 1986.

[47] K. Uutela, M. Hämäläinen, and R. Salmelin, "Global optimization in the localization of neuromagnetic sources," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 6, pp. 716–723, 1998.

[48] C. M. Michel, M. M. Murray, G. Lantz *et al.*, "EEG source imaging," 2004.

[49] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, pp. 276–280, mar 1986.

[50] G. Birot, L. Albera, and P. Chevalier, "Sequential High-Resolution Direction Finding From Higher Order Statistics," *IEEE Transactions on Signal Processing*, vol. 58, pp. 4144–4155, aug 2010.

[51] R. Grech, T. Cassar, J. Muscat *et al.*, "Review on solving the inverse problem in EEG source analysis," *Journal of NeuroEngineering and Rehabilitation*, vol. 5, no. 1, p. 25, 2008.

[52] J. C. Mosher and R. M. Leahy, "Source localization using recursively applied and projected (rap) music," *IEEE Transactions on Signal Processing*, vol. 47, no. 2, pp. 332–340, 1999.

[53] M. S. Hämäläinen and R. J. Ilmoniemi, "Interpreting magnetic fields of the brain: minimum norm estimates," *Medical & Biological Engineering & Computing*, vol. 32, no. 1, pp. 35–42, 1994.

[54] R. Pascual-Marqui, C. Michel, and D. Lehmann, "Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain," *International Journal of Psychophysiology*, vol. 18, pp. 49–65, oct 1994.

[55] R. D. Pascual-Marqui, "Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details.," *Methods and findings in experimental and clinical pharmacology*, vol. 24 Suppl D, pp. 5–12, 2002.

[56] M. Wagner, M. Fuchs, H.-A. Wischmann *et al.*, "Smooth reconstruction of cortical sources from EEG or MEG recordings," *NeuroImage*, vol. 3, no. 3, p. S168, 1996.

[57] R. G. D. P. Menendez, S. G. Andino, G. Lantz *et al.*, "Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations," *Brain Topography*, vol. 14, no. 2, pp. 131–137, 2001.

[58] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Royal Statistical Society*, vol. 58, no. 1, pp. 267–288, 1996.

[59] L. Ding, "Reconstructing cortical current density by exploring sparseness in the transform domain," *Physics in Medicine and Biology*, vol. 54, pp. 2683–2697, may 2009.

[60] W. Ou, P. Golland, and M. Hämäläinen, "A distributed spatio-temporal EEG/MEG inverse solver.," *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 11, pp. 26–34, 2008.

[61] J. Montoya-Martínez, A. Artés-Rodríguez, M. Pontil *et al.*, "A regularized matrix factorization approach to induce structured sparse-low-rank solutions in the EEG inverse problem," *EURASIP Journal on Advances in Signal Processing*, p. 97, dec 2014.

[62] H. Becker, L. Albera, P. Comon *et al.*, "SISSY: An efficient and automatic algorithm for the analysis of EEG sources based on structured sparsity," *NeuroImage*, vol. 157, pp. 157–172, aug 2017.

[63] F. Cong, Q. H. Lin, L. D. Kuang *et al.*, "Tensor decomposition of EEG signals: A brief review," *Journal of Neuroscience Methods*, vol. 248, pp. 59–69, 2015.

[64] J. M. Zumer, H. T. Attias, K. Sekihara *et al.*, "A probabilistic algorithm integrating source localization and noise suppression for MEG and EEG data," *NeuroImage*, vol. 37, pp. 102–115, aug 2007.

[65] D. Wipf and S. Nagarajan, "A unified Bayesian framework for MEG/EEG source imaging," *NeuroImage*, vol. 44, pp. 947–966, feb 2009.

[66] F. h. Costa, *Bayesian M/EEG source localization with possible joint skull conductivity estimation*. PhD thesis, Université de Toulouse, 2017. Thèse de doctorat dirigée par Tourneret, Jean-Yves et Batatia, Hadj Image, Information et Hypermédia Toulouse, INPT 2017.

[67] K. Friston, L. Harrison, J. Daunizeau *et al.*, "Multiple sparse priors for the M/EEG inverse problem," *NeuroImage*, vol. 39, pp. 1104–1120, feb 2008.

[68] G. Birot, L. Albera, F. Wendling *et al.*, "Localization of extended brain sources from EEG/MEG: The ExSo-MUSIC approach," *NeuroImage*, vol. 56, pp. 102–113, may 2011.

[69] P. Chevalier, A. Ferreol, and L. Albera, "High-Resolution Direction Finding From Higher Order Statistics: The 2q-MUSIC Algorithm," *IEEE Transactions on Signal Processing*, vol. 54, pp. 2986–2997, aug 2006.

[70] W. Van Drongelen, M. Yuchtman, B. D. Van Veen *et al.*, "A spatial filtering technique to detect and localize multiple sources in the brain," *Brain Topography*, vol. 9, no. 1, pp. 39–49, 1996.

[71] B. D. Van Veen, W. Van Drongelen, M. Yuchtman *et al.*, "Localization of brain electrical activity via linearly constrained minimum variance spatial filtering," *IEEE Transactions on Biomedical Engineering*, vol. 44, pp. 867–880, sep 1997.

[72] N. van Klink, A. Mol, C. Ferrier *et al.*, "Beamforming applied to surface EEG improves ripple visibility," *Clinical Neurophysiology*, vol. 129, pp. 101–111, jan 2018.

[73] O. Hauk, M. Stenroos, and M. Treder, "EEG/MEG Source Estimation and Spatial Filtering: The Linear Toolkit," in *Magnetoencephalography*, pp. 1–37, Cham: Springer International Publishing, 2019.

[74] H. R. Mohseni, F. Ghaderi, E. E. Wilding *et al.*, "A beamforming particle filter for EEG dipole source localization," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 337–340, IEEE, apr 2009.

[75] I. Kumihashi and K. Sekihara, "Array-gain constraint minimum-norm spatial filter with recursively updated gram matrix for biomagnetic source imaging," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 6, pp. 1358–1365, 2010.

[76] P. Georgieva, N. Bouaynaya, F. Silva *et al.*, "A Beamformer-Particle Filter Framework for Localization of Correlated EEG Sources," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 3, pp. 880–892, 2016.

[77] K. Sekihara and S. S. Nagarajan, *Adaptive Spatial Filters for Electromagnetic Brain Imaging*. Series in Biomedical Engineering, Berlin, Heidelberg: Springer Berlin Heidelberg, 2008.

[78] A. Moiseev, J. M. Gaspar, J. A. Schneider *et al.*, "Application of multi-source minimum variance beamformers for reconstruction of correlated neural activity," *NeuroImage*, vol. 58, pp. 481–496, sep 2011.

[79] Y. Jonmohamadi, G. Poudel, C. Innes *et al.*, "Comparison of beamformers for EEG source signal reconstruction," *Biomedical Signal Processing and Control*, vol. 14, no. 1, pp. 175–188, 2014.

[80] T. Piotrowski, J. Nikadon, and D. Gutiérrez, "MV-PURE Spatial Filters With Application to EEG/MEG Source Reconstruction," *IEEE Transactions on Signal Processing*, vol. 67, pp. 553–567, feb 2019.

[81] A. Gramfort, *Mapping, timing and tracking cortical activations with MEG and EEG: Methods and application to human vision*. PhD thesis, Ecole nationale superieure des telecommunications - ENST, oct 2010.

[82] S. Robinson, "Functional neuroimaging by synthetic aperture magnetometry (sam)," *Recent Advances in Biomagnetism*, pp. 302–305, 1999.

[83] J. Vrba and S. E. Robinson, "Linearly constrained minimum variance beamformers, synthetic aperture magnetometry, and MUSIC in MEG applications," in *Conference Record of the Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 313–317, IEEE, 2000.

[84] A. Hillebrand and G. R. Barnes, "The use of anatomical constraints with MEG beamformers," *NeuroImage*, vol. 20, pp. 2302–2313, dec 2003.

[85] R. Greenblatt, A. Ossadtchi, and M. Pflieger, "Local linear estimators for the bioelectromagnetic inverse problem," *IEEE Transactions on Signal Processing*, vol. 53, pp. 3403–3412, sep 2005.

[86] T. Imada, "A method for MEG data that obtains linearly-constrained minimum-variance beamformer solution by minimum-norm least-squares method," *IFMBE Proceedings*, vol. 28, pp. 152–154, 2010.

[87] D. J. McFarland, L. M. McCane, S. V. David *et al.*, "Spatial filter selection for EEG-based communication," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 3, pp. 386–394, 1997.

[88] V. Murzin, A. Fuchs, and J. Scott Kelso, "Detection of correlated sources in EEG using combination of beamforming and surface Laplacian methods," *Journal of Neuroscience Methods*, vol. 218, pp. 96–102, aug 2013.

[89] A. Kachenoura, L. Albera, L. Senhadji *et al.*, "ICA: A potential tool for BCI systems," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 57–68, 2008.

[90] Y. Jonmohamadi, G. Poudel, C. Innes *et al.*, "Source-space ICA for EEG source separation, localization, and time-course reconstruction," *NeuroImage*, vol. 101, pp. 720–737, 2014.

[91] A. Jaiswal, J. Nenonen, M. Stenroos *et al.*, "Comparison of beamformer implementations for MEG source localization," *bioRxiv*, vol. 5, 2019.

[92] J. Vrba, S. Taulu, J. Nenonen *et al.*, "Signal Space Separation Beamformer," *Brain Topography*, vol. 23, pp. 128–133, jun 2010.

[93] N. van Klink, A. Hillebrand, and M. Zijlmans, "Identification of epileptic high frequency oscillations in the time domain by using MEG beamformer-based virtual sensors," *Clinical Neurophysiology*, vol. 127, no. 1, pp. 197–208, 2016.

[94] N. van Klink, F. van Rosmalen, J. Nenonen *et al.*, "Automatic detection and visualisation of MEG ripple oscillations in epilepsy," *NeuroImage: Clinical*, vol. 15, pp. 689–701, dec 2017.

[95] Y. Hu, Y. Lin, B. Yang *et al.*, "Deep source localization with magnetoencephalography based on sensor array decomposition and beamforming," *Sensors (Switzerland)*, vol. 17, no. 8, 2017.

[96] Y. Hu, C. Yin, J. Zhang *et al.*, "Partial least square aided beamforming algorithm in magnetoencephalography source imaging," *Frontiers in Neuroscience*, vol. 12, pp. 1–11, sep 2018.

[97] T. Kimura, M. Kako, H. Kamiyama *et al.*, "Inverse solution for time-correlated multiple sources using Beamformer method," *International Congress Series*, vol. 1300, pp. 417–420, 2007.

[98] Z. Fatima, M. A. Quraan, N. Kovacevic *et al.*, "ICA-based artifact correction improves spatial localization of adaptive spatial filters in MEG," *NeuroImage*, vol. 78, pp. 284–294, 2013.

[99] J. Gross, J. Kujala, M. Hamalainen *et al.*, "Dynamic imaging of coherent sources: Studying neural interactions in the human brain," *Proceedings of the National Academy of Sciences*, vol. 98, pp. 694–699, jan 2001.

[100] M. Treder and G. Nolte, "Source reconstruction of broadband EEG/MEG data using the frequency-adaptive broadband (FAB) beamformer," *bioRxiv*, p. 502690, 2018.

[101] S. S. Dalal, A. G. Guggisberg, E. Edwards *et al.*, "Five-dimensional neuroimaging: Localization of the time-frequency dynamics of cortical activity," *NeuroImage*, vol. 40, no. 4, pp. 1686–1700, 2008.

[102] R. Ilmoniemi and J. Sarvas, *Brain Signals: Physics and Mathematics of MEG and EEG*. The MIT Press, 2019.

[103] K. Sekihara, S. Nagarajan, D. Poeppel *et al.*, "Reconstructing spatio-temporal activities of neural sources using an MEG vector beamformer technique," *IEEE Transactions on Biomedical Engineering*, vol. 48, pp. 760–771, jul 2001.

[104] M. Woolrich, L. Hunt, A. Groves *et al.*, "MEG beamforming using Bayesian PCA for adaptive data covariance matrix regularization," *NeuroImage*, vol. 57, no. 4, pp. 1466–1479, 2011.

[105] M. Van Vliet, N. Chumerin, S. De Deyne *et al.*, "Single-trial ERP component analysis using a spatiotemporal LCMV beamformer," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 1, pp. 55–66, 2016.

[106] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis*, vol. 88, no. 2, pp. 365–411, 2004.

[107] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. USA: Prentice-Hall, Inc., 1993.

[108] J. Serra and M. Najar, "Asymptotically optimal linear Shrinkage of sample LMMSE and MVDR filters," *IEEE Transactions on Signal Processing*, vol. 62, no. 14, pp. 3552–3564, 2014.

[109] T. Piotrowski, C. C. Zaragoza-Martinez, D. Gutierrez *et al.*, "MV-PURE estimator of dipole source signals in EEG," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 968–972, IEEE, may 2013.

[110] M. Grosse-Wentrup, C. Liefhold, K. Gramann *et al.*, "Beamforming in noninvasive brain-computer interfaces," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1209–1219, 2009.

[111] A. Barachant, *Commande robuste d'un effecteur par une interface cerveau machine EEG asynchrone*. PhD thesis, Grenobles, 2012.

[112] L. Spyrou and S. Sanei, "A new beamforming approach for the localisation ofevent related potentials," *European Signal Processing Conference*, pp. 2489–2493, 2007.

[113] L. Spyrou, *Adaptive techniques for the detection and localization of event related potentials from EEGs using reference signals* . PhD thesis, Cardiff University, 2008.

[114] L. Spyrou and S. Sanei, "Source localization of event-related potentials incorporating spatial notch filters," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 9, pp. 2232–2239, 2008.

[115] S. S. Dalal, K. Sekihara, and S. S. Nagarajan, "Modified beamformers for coherent source region suppression," *IEEE Transactions on Biomedical Engineering*, vol. 53, pp. 1357–1363, jul 2006.

[116] H. B. Hui and R. M. Leahy, "Linearly constrained MEG beamformers for MVAR modeling of cortical interactions," *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006.*, no. 1, pp. 237–240, 2006.

[117] H. B. Hui, D. Pantazis, S. L. Bressler *et al.*, "Identifying true cortical interactions in MEG using the nulling beamformer," *NeuroImage*, vol. 49, pp. 3161–3174, feb 2010.

[118] K. D. Rana, M. S. Hämäläinen, and L. M. Vaina, "Improving the Nulling Beamformer Using Subspace Suppression," *Frontiers in Computational Neuroscience*, vol. 12, p. 35, jun 2018.

[119] M. J. Brookes, C. M. Stevenson, G. R. Barnes *et al.*, "Beamformer reconstruction of correlated sources using a modified source model," *NeuroImage*, vol. 34, pp. 1454–1465, feb 2007.

[120] M. Diwakar, M. X. Huang, R. Srinivasan *et al.*, "Dual-Core Beamformer for obtaining highly correlated neuronal networks in MEG," *NeuroImage*, vol. 54, no. 1, pp. 253–263, 2011.

[121] H. Van Trees, *Detection, Estimation, and Modulation Theory, Optimum Array Processing*, vol. IV. New York, USA: John Wiley & Sons, Inc., 2002.

[122] G. Borgiotti and L. Kaplan, "Superresolution of uncorrelated interference sources by using adaptive array techniques," *IEEE Transactions on Antennas and Propagation*, vol. 27, no. 6, pp. 842–845, 1979.

[123] A. M. Dale, A. K. Liu, B. R. Fischl *et al.*, "Dynamic Statistical Parametric Mapping," *Neuron*, vol. 26, pp. 55–67, apr 2000.

[124] O. Hauk, D. G. Wakeman, and R. Henson, "Comparison of noise-normalized minimum norm estimates for MEG analysis using multiple resolution metrics," *NeuroImage*, vol. 54, no. 3, pp. 1966–1974, 2011.

[125] P. S. Unnikrishna and C. S. Burrus, *Array Signal Processing*. Signal Processing and Digital Filtering, New York, NY: Springer New York, 1989.

[126] I. Yamada and J. Elbadraoui, "Minimum-Variance Pseudo-Unbiased Low-Rank Estimator for Ill-Conditioned Inverse Problems," in *2006 IEEE International Conference on Acoustics Speed and Signal Processing Proceedings*, vol. 3, pp. III–325–III–328, IEEE, 2006.

[127] T. Piotrowski and I. Yamada, "Convex formulation of the stochastic MV-PURE estimator and its relation to the reduced rank Wiener filter," in *2008 International Conference on Signals and Electronic Systems*, pp. 397–400, IEEE, 2008.

[128] T. Piotrowski, R. L. Cavalcante, and I. Yamada, "Stochastic MV-PURE estimator - Robust reduced-rank estimator for stochastic linear model," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1293–1303, 2009.

[129] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, pp. 441–446, dec 2000.

[130] M. S. Treder, A. K. Porbadnigk, F. Shahbazi Avarvand *et al.*, "The LDA beamformer: Optimal estimation of ERP source time series using linear discriminant analysis," *NeuroImage*, vol. 129, pp. 279–291, apr 2016.

[131] B. Wittevrongel, *Brain-Computer Interfacing based on Time Domain EEG Response Detection combined with Beamforming*. PhD thesis, KU Leuven Biomedical Sciences Group Faculty of Medicine, oct 2018.

[132] H. R. Mohseni, M. L. Kringelbach, P. P. Smith *et al.*, "Application of a null-beamformer to source localisation in MEG data of deep brain stimulation," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pp. 4120–4123, IEEE, 2010.

[133] M. Popescu, E. A. Popescu, T. Chan *et al.*, "Spatio-temporal reconstruction of bilateral auditory steady-state responses using MEG beamformers," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 3, pp. 1092–1102, 2008.

[134] D. Gutierrez and C. C. Zaragoza-Martinez, "EEG source localization using beamforming in energy-constrained regions," *Conference Record - Asilomar Conference on Signals, Systems and Computers*, pp. 858–861, 2012.

[135] T. Piotrowski and J. Nikadon, "Localization of Brain Activity from EEG/MEG Using MV-PURE Framework," *arXiv*, pp. 1–13, 2018.

[136] L. Griffiths and C. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Transactions on Antennas and Propagation*, vol. 30, no. 1, pp. 27–34, 1982.

[137] B. D. Van Veen and K. M. Buckley, "Beamforming: A Versatile Approach to Spatial Filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, 1988.

[138] D. Gutiérrez, A. Nehorai, and A. Dogandžić, "Performance analysis of reduced-rank beamformers for estimating dipole source signals using EEG/MEG," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 5, pp. 840–844, 2006.

[139] P. Vergallo, A. Lay-Ekuakille, N. I. Giannoccaro *et al.*, "Processing EEG signals through Beamforming techniques for seizure diagnosis," *Proceedings of the International Conference on Sensing Technology, ICST*, pp. 497–501, 2012.

[140] S. A. Hossein Hosseini, A. Sohrabpour, M. Akcakaya *et al.*, "Electromagnetic brain source imaging by means of a robust minimum variance beamformer," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 10, pp. 2365–2374, 2018.

[141] A. Hillebrand and G. R. Barnes, "Practical constraints on estimation of source extent with MEG beamformers," *NeuroImage*, vol. 54, no. 4, pp. 2732–2740, 2011.

[142] T. Limpiti, B. D. Van Veen, and R. T. Wakai, "Cortical patch basis model for spatially extended neural activity," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 9, pp. 1740–1754, 2006.

[143] O. Hauk and M. Stenroos, "A framework for the design of flexible cross-talk functions for spatial filtering of EEG/MEG data: DeFleCT," *Human Brain Mapping*, vol. 35, no. 4, pp. 1642–1653, 2014.

[144] R. A. Monzingo, R. L. Haupt, and T. W. Miller, *Introduction to Adaptive Arrays*. Institution of Engineering and Technology, jan 2011.

[145] L. Chang and C. C. Yeh, "Performance of DMI and Eigenspace-Based Beamformers," *IEEE Transactions on Antennas and Propagation*, vol. 40, no. 11, pp. 1336–1347, 1992.

[146] Jung-Lang Yu and Chien-Chung Yeh, "Generalized eigenspace-based beamformers," *IEEE Transactions on Signal Processing*, vol. 43, no. 11, pp. 2453–2461, 1995.

[147] D. Cheyne, W. Gaetz, L. Garnero *et al.*, "Neuromagnetic imaging of cortical oscillations accompanying tactile stimulation," *Cognitive Brain Research*, vol. 17, no. 3, pp. 599–611, 2003.

[148] P. L. Furlong, A. R. Hobson, Q. Aziz *et al.*, "Dissociating the spatio-temporal characteristics of cortical neuronal activity associated with human volitional swallowing in the healthy adult brain," *NeuroImage*, vol. 22, no. 4, pp. 1447–1455, 2004.

[149] M. B. Hall, I. A. Nissen, E. C. van Straaten *et al.*, "An evaluation of kurtosis beamforming in magnetoencephalography to localize the epileptogenic zone in drug resistant epilepsy patients," *Clinical Neurophysiology*, vol. 129, pp. 1221–1229, jun 2018.

[150] S. E. Robinson, S. S. Nagarajan, M. Mantle *et al.*, "Localization of interictal spikes using SAM(g2) and dipole fit.," *Neurology & clinical neurophysiology : NCN*, p. 74, nov 2004.

[151] S. E. Robinson and J. Vrba, "Functional neuroimaging by syntheticaperture magnetometry (SAM)," in *Recent Advances in Biomagnetism* (T. Yoshimoto, M. Kotani, S. Kuriki *et al.*, eds.), (Sendai, Japan), pp. 302–305, Tohoku Univ. Press, 1999.

[152] D. Cheyne, A. C. Bostan, W. Gaetz *et al.*, "Event-related beamforming: A robust method for presurgical functional mapping using MEG," *Clinical Neurophysiology*, vol. 118, pp. 1691–1704, aug 2007.

[153] A. T. Herdman, A. Moiseev, and U. Ribary, "Localizing Event-Related Potentials Using Multi-source Minimum Variance Beamformers: A Validation Study," *Brain Topography*, vol. 31, no. 4, pp. 546–565, 2018.

[154] T. Piotrowski, D. Gutiérrez, I. Yamada *et al.*, "Reduced-rank neural activity index for EEG/MEG multi-source localization," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pp. 4708–4712, IEEE, may 2014.

[155] D. Rivière, D. Geffroy, I. Denghien *et al.*, "BrainVISA: An extensible software environment for sharing multimodal neuroimaging data and processing tools," in *15th HBM*, 2009.

[156] Q. Fang and D. A. Boas, "Tetrahedral mesh generation from volumetric binary and grayscale images," *Proceedings - 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009*, pp. 1142–1145, 2009.

[157] I. Unterberger, R. Bauer, G. Walser *et al.*, "Corpus callosum and epilepsies," *Seizure*, vol. 37, pp. 55–60, apr 2016.

[158] C. Studholme, D. J. Hawkes, and D. L. Hill, "Normalized entropy measure for multimodality image alignment," in *Medical Imaging 1998: Image Processing* (K. M. Hanson, ed.), vol. 3338, pp. 132–143, jun 1998.

[159] S. B. Eickhoff, K. E. Stephan, H. Mohlberg *et al.*, "A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data," *NeuroImage*, vol. 25, pp. 1325–1335, may 2005.

[160] S. Baillet and L. Garnero, "A Bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 5, pp. 374–385, 1997.

[161] U. Schmitt, A. K. Louis, F. Darvas *et al.*, "Numerical aspects of spatio-temporal current density reconstruction from EEG-/MEG-data," *IEEE Transactions on Medical Imaging*, vol. 20, no. 4, pp. 314–324, 2001.

[162] U. Schmitt and A. K. Louis, "Efficient algorithms for the regularization of dynamic inverse problems: I. Theory," *Inverse Problems*, vol. 18, no. 3, pp. 645–658, 2002.

[163] D. A. Harville, *Matrix Algebra From a Statistician's Perspective*. New York, NY: Springer New York, 1997.

[164] M. Vega-Hernández, E. Martínez-Montes, J. M. Sánchez-Bornot *et al.*, "Penalized Least Squares methods for solving the EEG Inverse Problem," *Statistica Sinica*, vol. 18, no. 4, pp. 1535–1551, 2008.

[165] H. Becker, L. Albera, P. Comon *et al.*, "Fast, variation-based methods for the analysis of extended brain sources," *European Signal Processing Conference*, pp. 41–45, 2014.

[166] R. Tibshirani, M. Saunders, S. Rosset *et al.*, "Sparsity and smoothness via the fused lasso," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, pp. 91–108, feb 2005.

[167] S. Boyd, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[168] N. Parikh and S. Boyd, "Proximal Algorithms," *Foundations and Trends R in Optimization*, vol. 1, no. 3, pp. 123–231, 2013.

[169] C. M. Michel and D. Brunet, "EEG source imaging: A practical review of the analysis steps," *Frontiers in Neurology*, vol. 10, apr 2019.

[170] F. Tadel, E. Bock, G. Niso *et al.*, "MEG/EEG group analysis with brainstorm," *Frontiers in Neuroscience*, vol. 13, pp. 1–21, feb 2019.

[171] aliobaibk, "Brain entropy in space and time (best) plug-in for brainstorm." `https://github.com/multifunkim/best-brainstorm`, 2019.

[172] A. Gramfort, M. Luessi, E. Larson *et al.*, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, no. 267, pp. 1–13, 2013.

[173] A. Gramfort, M. Luessi, E. Larson *et al.*, "MNE software for processing MEG and EEG data," *NeuroImage*, vol. 86, pp. 446–460, 2014.

[174] OpenMEEG, "Openmeeg: forward problems solver in the field of eeg and meg." `https://github.com/openmeeg/openmeeg`, 2018.

[175] C. Amblard, E. Lapalme, and J. M. Lina, "Biomagnetic Source Detection by Maximum Entropy and Graphical Models," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 3, pp. 427–442, 2004.

[176] C. Grova, J. Daunizeau, J. M. Lina *et al.*, "Evaluation of EEG localization methods using realistic simulations of interictal spikes," *NeuroImage*, vol. 29, no. 3, pp. 734–753, 2006.

[177] J. M. Lina, R. Chowdhury, E. Lemay *et al.*, "Wavelet-based localization of oscillatory sources from magnetoencephalography data," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 8, pp. 2350–2364, 2014.

[178] Y. Zerouali, C. L. Herry, B. Jemel *et al.*, "Localization of synchronous cortical neural sources," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 3, pp. 770–780, 2013.

[179] P. Renard, A. Alcolea, and D. Ginsbourger, "Stochastic versus Deterministic Approaches," in *Environmental Modelling*, pp. 133–149, Chichester, UK: John Wiley & Sons, Ltd, jan 2013.

[180] E. T. Jaynes, "Information Theory and Statistical Mechanics," *Physical Review*, vol. 106, pp. 620–630, may 1957.

[181] R. A. Chowdhury, J. M. Lina, E. Kobayashi *et al.*, "MEG Source Localization of Spatially Extended Generators of Epileptic Activity: Comparing Entropic and Hierarchical Bayesian Approaches," *PLoS ONE*, vol. 8, no. 2, 2013.

[182] J. Mattout, M. Pélégrini-Issac, L. Garnero *et al.*, "Multivariate source prelocalization (MSP): Use of functionally informed basis functions for better conditioning the MEG inverse problem," *NeuroImage*, vol. 26, no. 2, pp. 356–373, 2005.

[183] J. J. Kormylo and J. M. Mendel, "Maximum Likelihood Detection and Estimation of Bernoulli-Gaussian Processes," *IEEE Transactions on Information Theory*, vol. 28, no. 3, pp. 482–488, 1982.

[184] S. Bourguignon and H. Carfantan, "Bernoulli-Gaussian spectral analysis of unevenly spaced astrophysical data," in *IEEE/SP 13th Workshop on Statistical Signal Processing, 2005*, pp. 811–816, IEEE, 2005.

[185] S. Langford, *Digital Audio Editing: Correcting and Enhancing Audio in Pro Tools, Logic Pro, Cubase, and Studio One*. Routledge, 1st editio ed., 2013.

[186] M. Fink, M. Holters, and U. Zölzer, "Signal-Matched power-complementary cross-fading and dry-wet mixing," *DAFx 2016 - Proceedings of the 19th International Conference on Digital Audio Effects*, vol. 1, no. 1, pp. 109–112, 2016.

[187] S. Hu, Y. Lai, P. A. Valdes-Sosa *et al.*, "How do reference montage and electrodes setup affect the measured scalp EEG potentials?," *Journal of Neural Engineering*, vol. 15, no. 2, 2018.

[188] N. Alger, *Data-Scalable Hessian Preconditioning for Distributed Parameter PDE-Constrained Inverse Problems*. PhD thesis, University of Texas at Austin, 2019.

[189] E. P. Wigner, "Characteristic Vectors of Bordered Matrices with Infinite Dimensions II," in *The Collected Works of Eugene Paul Wigner*, vol. 62, pp. 541–545, Berlin, Heidelberg: Springer Berlin Heidelberg, 1993.

[190] J. Wishart and M. S. Bartlett, "The generalised product moment distribution in a normal system," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 29, pp. 260–270, may 1933.

[191] A. N. Bishop, P. Del Moral, and A. Niclas, "An introduction to Wishart matrix moments," *Foundations and Trends in Machine Learning*, vol. 11, no. 2, pp. 97–218, 2018.

[192] V. A. Marčenko and L. A. Pastur, "Distribution of Eigenvalues for some sets of random matrices," *Mathematics of the USSR-Sbornik*, vol. 1, pp. 457–483, apr 1967.

[193] G. Livan, M. Novaes, and P. Vivo, *Introduction to Random Matrices - Theory and Practice*. Springer New York, 2017.

[194] J. Baik and J. W. Silverstein, "Eigenvalues of large sample covariance matrices of spiked population models," *Journal of Multivariate Analysis*, vol. 97, no. 6, pp. 1382–1408, 2006.

[195] P. Šeba, "Random matrix analysis of human eeg data," *Physical Review Letters*, vol. 91, no. 19, pp. 1–4, 2003.

[196] E. Dolezal and P. Seba, "Spectral Density of Sample Covariance Matrices of Colored Noise," *arXiv*, pp. 1–9, jul 2008.

[197] M. Pal and P. M. Rao, "EEG Epilepsy Seizure Signal Characterization through theories of Random Matrix Analysis," in *8th International Conference on Latest Trends in Engineering and Technology (ICLTET-2016) May 5-6 2016 Dubai (UAE)*, pp. 8–11, International Institute of Engineers, may 2016.

[198] T. Sørenson, *A Method of Establishing Groups of Equal Amplitude in Plant Sociology Based on Similarity of Species Content and Its Application to Analyses of the Vegetation on Danish Commons*. Biologiske skrifter, I kommission hos E. Munksgaard, 1948.

[199] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, pp. 297–302, jul 1945.

[200] J. Yao and J. P. Dewald, "Evaluation of different cortical source localization methods using simulated and experimental eeg data," *NeuroImage*, vol. 25, no. 2, pp. 369–382, 2005.

[201] A. Molins, S. M. Stufflebeam, E. N. Brown *et al.*, "Quantification of the benefit from integrating MEG and EEG data in minimum l2-norm estimation," *NeuroImage*, vol. 42, no. 3, pp. 1069–1077, 2008.

[202] R. A. Chowdhury, I. Merlet, G. Birot *et al.*, "Complex patterns of spatially extended generators of epileptic activity: Comparison of source localization methods cMEM and 4-ExSo-MUSIC on high resolution EEG and MEG data," *NeuroImage*, vol. 143, pp. 175–195, 2016.

[203] T. Hedrich, G. Pellegrino, E. Kobayashi *et al.*, "Comparison of the spatial resolution of source imaging techniques in high-density eeg and meg," *NeuroImage*, vol. 157, pp. 531–544, dec 2017.

[204] S. Provost and A. Mathai, *Quadratic Forms in Random Variables: Theory and Applications / A.M. Mathai, Serge B. Provost*. Statistics : textbooks and monographs, Marcel Dek ker, 1992.

[205] D. G. Luenberger, *Optimization by Vector Space Methods*. John Wiley & Sons, Inc., apr 1970.

[206] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," 2012.

UNIVERSITE BRETAGNE LOIRE / MATHSTIC

UNIVERSITÉ DE RENNES 1

Titre: Identification des réseaux épileptogènes dans le cas des épilepsies partielles pharmaco-résistantes

Mot clés: Beamforming, imagerie de sources cérébrales, EEG-hd, localisation de sources, reconstruction de sources, SABLE.

Résumé: L'épilepsie est l'une des maladies neurologiques les plus courantes, touchant 2,4 millions de personnes dans le monde. 20 à 30% de ces patients sont résistants aux médicaments et nécessitent des procédures alternatives pour réduire la fréquence ou la sévérité des crises. Ces procédures nécessitent la localisation de la zone épileptogène qui s'obtient en résolvant un problème inverse mal posé au sens de Hadamard. Une pléthore de méthodes sont disponibles pour résoudre ce problème. Parmi elles, les méthodes de beamforming présentent des aprioris uniques mais sont souvent mal utilisées pour être compatibles avec le modèle de source distribuée connu pour être adapté à l'épilepsie. Dans cette thèse, nous proposons tout d'abord de réaliser une étude approfondie des méthodes de beamforming existantes afin d'identifier les problèmes actuels sur le sujet. En plus d'un travail littéraire et mathématiques, nous réalisons également une étude sur des données simulées et réelles qui conduira à la proposition du meilleur pipeline de beamforming pour la localisation de sources ponctuelles épileptiques ainsi qu'à l'introduction de nouvelles pratiques. La deuxième contribution de cette thèse est la proposition d'une nouvelle méthode appelée SABLE qui adapte le beamforming au cas des sources distribuées. Cette dernière exploite conjointement la méthode du beamforming UG (*Unit Gain*) et la méthode SISSY (*Source Imaging based on Structured Sparsity*) pour surmonter les défauts inhérents au beamforming. Les résultats sur des données simulées et réelles ont révélé un gain d'efficacité supérieur aux méthodes traditionnelles d'imagerie des sources cérébrales. La dernière contribution de ce manuscrit est la proposition d'un pipeline repensé qui amène SABLE à un niveau de maturité algorithmique équivalente aux méthodes traditionnelles.

Title: Identification of epileptogenic networks in drug-resistant partial epilepsy

Keywords: Beamforming, brain source imaging, hr-EEG, source localization, source reconstruction, SABLE

Abstract: Epilepsy is one of the most common neurological diseases, affecting 2.4 million people worldwide. 20-30% of these patients are drug-resistant and require alternative procedures to reduce the frequency or severity of seizures. These procedures require the localization of the epileptogenic zone which is obtained by solving an ill-posed inverse problem in Hadamard's sense. A plethora of methods are available to solve this problem. Among them, beamforming methods present unique prior but are often misused to be compatible with the distributed source model known to be suitable for epilepsy. In this thesis, we first propose carrying out an in-depth study of the existing beamforming methods to identify the current problems on the subject. In addition to literary and mathematical work, we also perform a study on simulated and real data that will lead to the proposal of the best beamforming pipeline for the localization of epileptic point sources as well as the introduction of new practices. The second contribution of this thesis is the proposal of a new method called SABLE which adapts beamforming to the case of distributed sources. The latter jointly exploits the UG (Unit Gain) beamforming method and the SISSY (Source Imaging based on Structured Sparsity) method to overcome the inherent defects of beamforming. The results on simulated and real data revealed a gain in efficiency beyond the traditional Brain Source Imaging methods. The last contribution of this manuscript is the proposal of a redesigned pipeline that brings SABLE to a level of algorithmic maturity equivalent to traditional methods.