Politecnico di Torino
Master Degree in Physics of Complex Systems

Temporal Regularization in EEG Source Localization

An Informed Approach to Neural Dynamics Reconstruction

Politecnico
di Torino

Supervisor:
Prof. Luca Mesin

Candidate:
Daniel Alexandru Cazacu

Academic Year 2025/2026

Contents

Abstract

Electroencephalography (EEG) source localization is a well-known ill-posed inverse problem: thousands of cortical dipole moments must be reconstructed from measurements at only a few scalp electrodes, making regularization essential. Standard approaches like Tikhonov regularization and LORETA, impose spatial constraints such as minimum energy or spatial smoothness, but they treat each time point independently. This assumption of temporal independence oversimplifies the underlying neural dynamics. Neural populations exhibit intrinsic smoothness arising from membrane time constants (10–50 ms), synaptic filtering, and recurrent network interactions. Since EEG reflects field potentials rather than spike trains, and given its typical frequency range (1–100 Hz), it is reasonable to assume that neural activations are temporally correlated rather than white. Here, is introduced a spatio-temporal regularization framework that incorporates a physiologically motivated temporal prior. The idea is to penalize the first temporal derivative of the dipole moments, encoding the belief that neural currents evolve smoothly. The resulting objective combines data fidelity with a spatial (Tikhonov) penalty and a temporal smoothness term $\|DJ^\top\|^2$, where D is the first-difference operator. Through vectorization and Kronecker products, the normal equations reduce to a block-tridiagonal system that couples each time point to its neighbors with strength λ_t. This structure permits efficient solution via Cholesky decomposition and chunk processing, circumventing the memory demands of naive Kronecker matrices. The Bayesian interpretation is straightforward: the prior covariance precision includes $\lambda_s I + \lambda_t (D^\top D \otimes I)$, explicitly modeling temporal correlations absent in conventional methods. The method is validated on synthetic ground truth data generated from smooth neural mass activations, as Brownian dynamics with friction and oscillatory bursts, using a leadfield model. Comparisons with Tikhonov and LORETA show that temporal regularization improves reconstruction accuracy across multiple metrics: dipole-wise correlation, parcel-wise correlation, and temporal waveform fidelity. The improvement is most pronounced under high noise (SNR below 5 dB), where purely spatial estimates become dominated by jitter, whereas the spatio-temporal approach leverages temporal correlations to suppress noise. Qualitatively, temporal regularization recovers physiologically plausible trajectories and eliminates the source switching that plagues time-independent solvers. The results indicate that temporal regularization is a necessary biophysical constraint. By aligning the inverse solution with the known temporal correlation structure of neural dynamics, the reconstructions become more accurate, stable, and interpretable. This work thus demonstrates that statistical inversion, combined with computational methods and informed priors, can improve EEG source localization by incorporating physiological plausibility. **Keywords:** EEG inverse problem, source localization, temporal regularization, ill-posed problems, Bayesian inference, neural dynamics, spatio-temporal priors

List of Abbreviations

AAL	Automated Anatomical Labeling atlas
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
BEM	Boundary Element Method
BOLD	Blood-Oxygen-Level-Dependent
CLT	Central Limit Theorem
CN	Condition Number
CSF	Cerebrospinal Fluid
DF	Degrees of Freedom
ECoG	Electrocorticography
EEG	Electroencephalography
eLORETA	exact Low Resolution Electromagnetic Tomography
ERP	Event-Related Potential
FEM	Finite Element Method
fMRI	functional Magnetic Resonance Imaging
GABA	γ-aminobutyric acid
HRF	Hemodynamic Response Function
i.i.d.	independent and identically distributed
LFP	Local Field Potential
LORETA	Low Resolution Electromagnetic Tomography
MAP	Maximum A Posteriori
MEG	Magnetoencephalography
MNE	Minimum Norm Estimation
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
NMDA	N-methyl-D-aspartate
PCA	Principal Component Analysis
PDF	Probability Density Function
PSD	Power Spectral Density
PSP	Postsynaptic Potential
RMSE	Root Mean Square Error
RMS	Root Mean Square
SNR	Signal-to-Noise Ratio
SSTR	Simultaneous Spatio-Temporal Regularization
SVD	Singular Value Decomposition
WMNE	Weighted Minimum Norm Estimation

Introduction

Electroencephalography (EEG) is a non-invasive technique that measures the brain's electrical activity from the scalp, offering millisecond temporal resolution essential for studying neural dynamics. Its utility is fundamentally constrained by an ill-posed inverse problem: the cortical sources of the measured EEG must be inferred from a limited set of scalp electrodes. As it will be established in Chapter 2, this problem is characterized by non-uniqueness and instability, where the number of unknown dipole sources far exceeds the number of measurements, and small amounts of noise can lead to wildly inaccurate reconstructions.

To render the problem tractable, regularization is required. Classical methods, such as Tikhonov regularization (minimum norm estimation) and LORETA, address ill-posedness by imposing spatial constraints, favoring solutions with either minimum energy or maximum spatial smoothness. As detailed in Chapter 3, these approaches treat each time point independently. While computationally efficient, this assumption of temporal independence is in stark contrast with the known physiology of neural populations.

The central motivation for this thesis is the physiological reality that neural activity is inherently smooth and temporally correlated. As argued in Chapter 1, neurons do not fire independently at each millisecond; rather, their dynamics are shaped by membrane time constants (10–50 ms), synaptic filtering, and recurrent network interactions. The macroscopic signals captured by EEG, which average over millions of neurons, reflect this ordered, population-level behavior. Therefore, a reconstruction that assumes temporal whiteness is both statistically inefficient and physically implausible.

The following addresses this gap by introducing a spatio-temporal regularization framework that incorporates a physiologically motivated temporal prior. The objective is to go beyond purely spatial constraints by explicitly modeling the temporal evolution of neural currents. The proposed method, called SSTR (Simultaneous Spatio-Temporal Regularization), penalizes the first temporal derivative of the dipole moments, encoding the belief that neural currents should change smoothly over time. This leads to a combined objective function that balances data fidelity with both spatial (Tikhonov) and temporal smoothness penalties.

The primary contributions are the following. First, it provides a mathematical formulation of the spatio-temporal inverse problem, showing how temporal regularization leads to a block-tridiagonal system that couples solutions across time points. Second, it presents an efficient computational implementation, leveraging Kronecker products, Sylvester equations, and dual formulations to make the solution of this large-scale problem feasible and memory-efficient. Third, it validates the method on a controlled synthetic ground truth dataset, comparing its performance against standard Tikhonov and LORETA approaches.

The SSTR method shows greater robustness to noise, particularly at low signal-to-noise ratios, and recovers more physiologically plausible temporal trajectories. This work

underscores that aligning the inverse solution with known biophysical constraints, in this case, the temporal correlation structure of neural dynamics, is a necessary step for accurate and interpretable EEG source localization.

The remainder of this thesis is organized as follows. Chapter 1 establishes the physiological basis for temporal smoothness in neural dynamics. Chapter 2 presents the mathematical formalism of the EEG forward and inverse problems. Chapter 3 reviews existing spatial priors, highlighting their limitations. Chapter 4 introduces the theory and efficient implementation of the proposed temporal regularization method. Chapter 5 details the empirical validation framework, including the construction of a synthetic ground truth. Chapter 6 presents the quantitative and qualitative results, comparing the proposed method against traditional approaches. Finally, Chapter 7 summarizes the findings and discusses directions for future research.

Chapter 1

Neural Dynamics and EEG

1.1 The Brain as a Complex Dynamical System

The human brain represents one of the most intricate examples of a complex dynamical system in nature. Comprising approximately 10^{11} neurons, each making thousands of synaptic connections, the brain exhibits emergent phenomena that cannot be predicted from the properties of individual neurons alone.

1.1.1 Microscopic Chaos, Macroscopic Order

At the microscopic level, individual neurons exhibit highly irregular, seemingly stochastic spiking behavior. The timing of action potentials can vary substantially across repeated presentations of identical stimuli, and the membrane potential fluctuates continuously due to synaptic bombardment from thousands of presynaptic neurons.

However, by observing populations of neurons collectively, a different picture emerges. The collective activity of large neuronal populations exhibits ordered, predictable dynamics, which is a property found in complex systems, where *the whole is different from the sum of its parts*.

At microscopic level, individual neurons spike irregularly and chaotically, with trial-to-trial variability. Predicting any single neuron's activation trajectory is impossible. At macroscopic level populations show coherent oscillations, reliable event-related potentials, and predictable state transitions that evolve according to deterministic laws . There is emergence of field potentials, oscillations, and network states from collective interactions that cannot be reduced to single-neuron properties, but arise from collective statistics.

This microscopic chaos with macroscopic order has many implications for EEG source localization. EEG electrodes, positioned centimeters away from neural sources, aggregate signals from millions of neurons. The fluctuations of individual neurons then average out, leaving only the coherent population-level dynamics. Thus, EEG measures the *macroscopic order* emergent from microscopic chaos.[2]

1.1.2 Neural Mass Models and Collective Dynamics

The recognition that population-level dynamics are the appropriate scale for modeling EEG has led to the development of *neural mass models*, that treat neuronal populations as continuous variables, typically mean firing rates and average synaptic activities, rather than simulating individual spiking neurons.[1] This is justified for a single EEG electrode

measures the summed activity of approximately 10^6—10^8 neurons. With such large numbers, is assumed that the average behavior dominates, while individual variations cancel. By the central limit theorem, is inferred that the population aggregate converges to a Gaussian-distributed mean field.

Physically, volume conduction through brain tissue, cerebrospinal fluid, skull, and scalp acts as a spatial low-pass filter. High-spatial-frequency components (corresponding to fine-grained spiking patterns) are attenuated, while low-spatial-frequency components (corresponding to synchronized population activity) propagate to the scalp.

The dynamical equations for a neural mass typically take the form:

$$\tau \frac{dE}{dt} = -E + S(wE - \theta) \tag{1.1}$$

where E represents the mean excitatory population activity, τ is a time constant (typically 5–20 ms), $S(\cdot)$ is a sigmoidal activation function representing the population firing rate response to input, w is the connection strength within or between populations, and θ is a threshold parameter.

The sigmoid function requires some kind of nonlinearity:

$$S(x) = \frac{1}{1 + e^{-x}} \tag{1.2}$$

or similar forms. This introduces saturation, so that populations cannot fire faster than their refractory periods allow, which creates the possibility of multiple stable states, oscillations, and bistability.

The collective behavior depends on coupling strength:
Weak coupling ($w < w_c$) where populations exhibit desynchronized, low-amplitude activity, the normal resting state.
Strong coupling ($w > w_c$) where pathological synchronization emerges, as seen in epileptic seizures.
The consequent **critical coupling** ($w \approx w_c$) is found when the system exhibits scale-invariant dynamics, $1/f$ spectra, and optimal information processing, the hypothesized operating point of healthy cortex.

The idea for EEG source localization is that the variables in neural mass models, like mean firing rates, average synaptic activities, are *continuous and smooth.* They evolve according to differential equations with characteristic time constants, not as independent point processes.

1.1.3 Timescales of Neural Processes

Neural activity spans a large range of timescales, from microseconds to minutes. The hierarchy is useful to give some intuition for why temporal smoothness is a physiologically motivated assumption.[4]

EEG occupies a specific band: Typical EEG recordings filter signals between 0.1–100 Hz, corresponding to periods of 10 ms to 10 s. This band is both too slow to capture individual action potentials (1 ms duration) and too fast to capture metabolic changes (minutes to hours), but matched to population synchronization dynamics

Timescale separation enables modeling: Processes at different timescales can often be treated separately: *Fast variables* (spikes, fast synaptic transmission) can be

Table 1.1: Hierarchy of neural timescales.

Timescale	Processes	Relevance to EEG
10^{-9} s	Ion channel gating	Below measurement limit
10^{-6} s	Channel opening kinetics	Below measurement limit
1–10 ms	Action potentials, AMPA synapses	Too fast for direct EEG
10–100 ms	NMDA receptors, dendritic integration	Gamma oscillations (25–100 Hz)
0.1–1 s	Neural adaptation, short-term plasticity	Alpha (8–13 Hz), beta (13–30 Hz)
1–10 s	Working memory, cognitive processes	Theta (4–8 Hz), delta (1–4 Hz)
10 s +	Learning, consolidation	Infra-slow oscillations (<0.1 Hz)

averaged or treated as noise; *Slow variables* (synaptic weights, neuromodulator concentrations) can be treated as constant parameters; *Intermediate variables* (population firing rates, field potentials) are the main interest for the work's discussion. Temporal smoothness emerges from biological filters, as multiple mechanisms present smooth neural responses, like Membrane time constants $\tau_m = R_m C_m \approx 10$–$50$ ms, where neurons integrate inputs over this window; or synaptic time constants like AMPA: 2–5 ms, NMDA: 50–150 ms, GABA: 5–10 ms. Dendrites act as distributed RC circuits, smoothing inputs in both space and time, and netwrek feedback coming when recurrent excitation sustains activity beyond input duration.

1.1.4 Spatial and Temporal Correlations in Neural Activity

Complex systems are characterized by correlations across multiple scales. In the brain, these are not measurement artifacts but empirical properties of neural organization.

Spatial Correlations:

They arise from anatomical connectivity. For instance **local connections** that come from within a cortical column (1 mm diameter), where neurons are densely interconnected, and activity in neighboring neurons is highly correlated. **Short-range connections**, instead, esist between nearby columns (few mm), where correlations remain significant due to horizontal fibers in layers 2/3; in **long-range connections**, instead, white matter tracts link to distant cortical regions, which creates correlations between functionally distant but related areas. There are also regimes and regions that have **scale-free properties**, where spatial correlations often follow power laws, which an indicator of criticality in the system.

Temporal Correlations:

Other type of correlations are temporal, and described as **autocorrelation**, which is the way activity at time t predicts activity at time $t+\tau$ for lags up to hundreds of milliseconds. This is quantified by the autocorrelation function:

$$R(\tau) = \frac{\mathbb{E}[(X_t - \mu)(X_{t+\tau} - \mu)]}{\sigma^2} \tag{1.3}$$

where R is the autocorrelation function of time τ, and $\mathbb{E}$ is the expectation value of the random variable X_t, μ is the avg and σ^2 is the variance. Another kind of correlation comes

from the **power spectrum** profile, which is a power spectral density of neural activity, and typically follows $P(f) \propto 1/f^{\alpha}$ with $\alpha \approx 1$, which is a "pink noise" that indicates correlations across all timescales (criticality).

There exists some physiological evidence for this, as LFP and MEG studies show significant temporal structures. Empytically, event-related potentials evolve smoothly over hundreds of milliseconds, and oscillatory bursts have smooth envelopes, not abrupt onsets and offsets.

Correlations carry information about brain function. **Phase synchrony** may exist, as synchronization between distant regions indexes information transfer and functional connectivity. **Cross-frequency coupling** may exist too, where the phase of slower oscillations modulates the amplitude of faster oscillations (e.g., theta–gamma coupling during memory tasks).

The Criticality Hypothesis

Experimental evidence suggests the brain operates near a critical point. The size of electrical avalanches show a power law distributions, and a long range temporal correlations, with noise scaling as $1/f$. The observed temporal correlations and temporal regularization seeks to preserve signatures of this critical state.[5]

1.2 Electromagnetic Foundations of EEG

1.2.1 Biophysics of Extracellular Fields

On why Post-Synaptic Potentials Dominate EEG

The EEG primarily reflects *post-synaptic potentials* (PSPs).[6]

Generally, action potentials are brief (1 ms), and even if thousands of neurons fire synchronously, their contributions do not sum effectively at the scalp. They propagate along axons, creating a quadrupolar current configuration that falls off as $1/r^3$(too rapidly to contribute significantly at scalp distances. PSPs, by contrast, last 10–100 ms, allowing substantial temporal summation. They arise from aligned dendritic trees, creating dipolar configurations that fall off as $1/r^2$.

Geometrically, neurons in the cortex have apical dendrites oriented perpendicular to the cortical surface. This alignment creates coherent dipole moments across populations of thousands of neurons.

EEG measures the *synchronized post-synaptic activity of aligned pyramidal neurons.* Each such neuron acts as a microscopic current dipole; the macroscopic EEG signal arises from the superposition of millions of these dipoles.

Primary and Volume Currents

To characterize the forward problem, one must distinguish two types of currents: the **Primary (or source) current $\mathbf{J}^p$** which is the current flowing across neuronal membranes, and is to be reconstruct in source localization, and the **volume (or return) current $\mathbf{J}^v$** which is the passive, ohmic current flowing through the extracellular space to complete the circuit, and is what EEG electrodes actually measure.

Current conservation (charge continuity) requires:

$$\nabla \cdot (\mathbf{J}^p + \mathbf{J}^v) = 0 \tag{1.4}$$

The volume current is related to the electric field $\mathbf{E}$ and electric potential Φ by Ohm's law in a conductive medium:

$$\mathbf{J}^v = \sigma \mathbf{E} = -\sigma \nabla \Phi \tag{1.5}$$

where σ is a conductivity.

Combining equations follows the relationship:

$$\nabla \cdot (\sigma \nabla \Phi) = \nabla \cdot \mathbf{J}^p \tag{1.6}$$

which is Poisson's equation for the electric potential, with the divergence of the primary current as the source term. The EEG inverse problem consists of estimating $\mathbf{J}^p$ from measurements of Φ on the scalp.

The Dipole Approximation

For source localization, one typically approximates each active neuronal population as a *current dipole*. This approximation is justified by a multipole expansion of the source current distribution:

$$\mathbf{J}^p(\mathbf{r}) \approx \mathbf{q}\delta(\mathbf{r} - \mathbf{r}_0) + \mathbf{Q} \cdot \nabla \delta(\mathbf{r} - \mathbf{r}_0) + \dots \tag{1.7}$$

The dipole term $\mathbf{q}$ (the first moment) dominates at distances large compared to the source extent, that is precisely the situation for EEG, where electrodes are centimeters away from cortical sources.

The dipole moment $\mathbf{q}$ has both magnitude and orientation. For a population of N aligned pyramidal neurons, the net dipole moment is approximately:

$$\mathbf{q} \approx \sum_{i=1}^{N} q_i \hat{\mathbf{n}}_i \approx N \bar{q} \hat{\mathbf{n}} \tag{1.8}$$

where $\hat{\mathbf{n}}$ is the average orientation (typically perpendicular to the cortical surface). Typical dipole moments in EEG are on the order of 10–100 nAm, corresponding to 10^5–10^6 synchronously active neurons .

The Forward Problem is Well-Posed

The forward problem is well-posed in the Hadamard sense: **there exists** a unique potential distribution for any physically realizable source configuration; Given the source $\mathbf{J}$, the potential $\mathbf{\Phi}$ **is uniquely determined**; Small changes in $\mathbf{J}$ produce small changes in $\mathbf{\Phi}$, as the condition number of $\mathbf{L}$ is moderate, 10–100, so **it's stable**.

Given the nature of the problem, EEG signals can be computed given a source model, and the problem is not hard.

1.2.2 The Ill-Posed Nature of EEG Source Localization

Hadamard's Definition of Well-Posed Problems

As previously supposed, a problem is Hadamard *well-posed* if it satisfies three conditions: A solution exists; The solution is unique; The solution depends continuously on the data.

If a problem violates any of these conditions, is called *ill-posed*. The EEG inverse problem violates conditions of uniqueness and stability.

Non-Uniqueness: More Unknowns Than Equations

Consider the dimensions of the EEG inverse problem:

$$\text{Measurements:} \quad N_c \times T \text{ (typically } 64 \times 1000 = 64,000) \tag{1.9}$$

$$\text{Unknowns:} \quad 3N_d \times T \text{ (typically } 3 \times 8000 \times 1000 = 24,000,000) \tag{1.10}$$

where N_c is the number of channels, T is the number of discrete time points and N_d is the number of dipoles. The system is underdetermined by a lot, and the linear system $\mathbf{\Phi} = \mathbf{LJ}$ has infinitely many solutions. Any solution can be written as:

$$\mathbf{J} = \mathbf{J}_0 + \mathbf{J}_{\text{null}} \tag{1.11}$$

where $\mathbf{J}_0$ is a particular solution and $\mathbf{J}_{\text{null}}$ is any vector in the null space of $\mathbf{L}$ (i.e., $\mathbf{L}\mathbf{J}_{\text{null}} = 0$). These *silent sources* produce no scalp potential and are invisible to EEG.

Therefore there exist some configurations of neural activity that EEG simply cannot see. This non-uniqueness limitation that different dipole arrangements can produce identical scalp potentials, is imposed by volume conduction.

Instability: Small Errors have Large Consequences

Even if one could resolve the non-uniqueness through additional assumptions, there is an intrinsic instability. The condition number of the leadfield matrix $\mathbf{L}$ quantifies how much errors in the data amplify when solving the inverse problem.

For realistic head models, this number can be so huge that a 1 µV measurement error could translate into a 1 V error in the estimated source amplitudes which is a completely meaningless results.[15]

The instability comes from the low condicutivity of the skull (about 100× less than brain), that acts as a spatial low-pass filter. High spatial frequency components of the source distribution are exponentially attenuated and aren't recovered stably from scalp measurements.

Regularization as Physical Necessity

Given non-uniqueness and instability, regularization is necessary. This addresses the afromentioned issues by injecting prior information, which are constraints that select one solution among the infinite possibilities to stabilize the inversion. The choice of regularization is the choice of what "reasonable" means for neural sources.

All regularization methods navigate some sort of trade-off:
Strong regularization gives solutions that are stable and unique but may be over-smoothed, missing fine spatial or temporal detail. Bias is high, variance low.

Weak regularization gives solutions that can capture more detail but are unstable and sensitive to noise. Bias is low, variance high.

Optimally, there should be a balance between these extremes, minimizing the total error (bias^2 + variance).

1.3 Physiological Justification for Temporal Smoothness

Given the physics of EEG generation and the mathematical structure of the inverse problem, one can now justify why temporal smoothness is a physiologically appropriate prior for neural source reconstruction.

1.3.1 Neural Integration and Temporal Filtering

Membrane Time Constants as Biological Low-Pass Filters

As previously stated, every neuron acts as a biological low-pass filter. The membrane has both resistance R_m and capacitance C_m, giving it a characteristic time constant:

$$\tau_m = R_m C_m \approx 10\text{–}50 \text{ ms} \tag{1.12}$$

which determines how quickly the membrane potential responds to synaptic inputs. An instantaneous current injection produces a voltage change that rises and decays with time constant τ_m. Inputs separated by less than τ_m integrate, while inputs separated by more than τ_m are filtered out.[10]

The consequence is that neurons cannot respond instantaneously to rapid input fluctuations. This membrane filter has a cutoff frequency $f_c = 1/(2\pi\tau_m) \approx 3\text{–}16$ Hz, while fluctuations above this frequency are attenuated.

Synaptic Dynamics Add Further Smoothing

Synaptic transmission introduces additional temporal filtering:

- **AMPA receptors** (fast excitation): $\tau \approx 2\text{–}5$ ms, cutoff 30–80 Hz
- **NMDA receptors** (slow excitation): $\tau \approx 50\text{–}150$ ms, cutoff 1–3 Hz
- **GABA$_A$ receptors** (fast inhibition): $\tau \approx 5\text{–}10$ ms, cutoff 15–30 Hz
- **GABA$_B$ receptors** (slow inhibition): $\tau \approx 100\text{–}200$ ms, cutoff 0.8–1.6 Hz

The net effect is a cascaded filtering, where synaptic transmission smooths the spike train input, and membrane integration smooths further.

Population-Level Smoothing

When considering populations of neurons, the smoothing becomes even more pronounced. Individual neurons may fire irregularly, but their collective activity is smoother, for there exists some **statistical averaging** (by the CLT, the sum of many independent irregular processes converges to a Gaussian process with reduced variance and smoother sample

paths), **asynchronous firing** (neurons in a population rarely fire precisely together, the temporal jitter smooths the population-level average) and **recurrent connectivity** (feedback loops create attractor dynamics, with smooth trajectories between attractors in state space).

EEG/MEG Temporal Evolution

Event-related potentials (ERPs) provide perhaps better evidence. The N100, P300, and other ERP components evolve smoothly over hundreds of milliseconds. Even for stimuli as brief as 1 ms, the evoked response is a smooth waveform.

Oscillatory activity also shows smooth envelopes. The amplitude of alpha, beta, and gamma oscillations modulates slowly relative to the carrier frequency This is not specific to EEG only. MEG shows identical temporal structure. The hemodynamic response measured by fMRI is a convolution of neural activity with the canonical HRF (time-to-peak 5 s) with the same kind of behaviour.

1.3.2 Timescales of Interest in Cognitive Neuroscience

Cognitive Processes Have Characteristic Timescales

The cognitive processes studied with EEG operate on specific timescales:

All of these timescales are substantially longer than the 1 ms duration of an action potential. The neural processes underlying cognition are inherently slow, involving the coordinated activity of large populations over hundreds of milliseconds.[11]

Oscillatory Bands Reflect Functional Timescales

The canonical EEG frequency bands correspond to different cognitive functions and timescales:

- **Gamma (30–100 Hz)**: Period 10–33 ms. Associated with local processing, feature binding, and attention.
- **Beta (13–30 Hz)**: Period 33–77 ms. Dominant during sensorimotor processing and motor preparation.
- **Alpha (8–13 Hz)**: Period 77–125 ms. Reflects cortical inhibition, attention, and working memory.
- **Theta (4–8 Hz)**: Period 125–250 ms. Critical for navigation, memory encoding, and retrieval.
- **Delta (1–4 Hz)**: Period 250–1000 ms. Prominent during deep sleep and certain cognitive tasks.

Even gamma, the fastest oscillation, has a period of 10–30 ms, which is far longer than the millisecond timescale of individual spikes. At the population level, activity varies on these timescales, which are those EEG is designed to measure anyways.

Chapter 2

Mathematical Formalism of EEG Inverse Problems

2.1 The Forward Problem as an Integral Equation

2.1.1 From Maxwell to Poisson

We begin with Maxwell's equations in their differential form:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.1}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \tag{2.2}$$

$$\nabla \cdot \mathbf{D} = \rho \tag{2.3}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.4}$$

$$\mathbf{D} = \epsilon \mathbf{E} \tag{2.5}$$

$$\mathbf{B} = \mu \mathbf{H} \tag{2.6}$$

where $\epsilon \approx \epsilon_0$ and $\mu \approx \mu_0$ (tissue is non-magnetic and has permittivity near vacuum in the considered regime). For EEG signals, the frequency content is typically below 100 Hz. Under these conditions the quasi-static approximation applies: the time derivatives in Equations (2.1) and (2.2) become negligible compared to the other terms. This approximation is used for biological tissue at these frequencies . Consequently:

$$\nabla \times \mathbf{E} \approx 0 \tag{2.7}$$

Equation (2.7) implies that the electric field is irrotational and can therefore be expressed as the gradient of a scalar potential:

$$\mathbf{E} = -\nabla \Phi \tag{2.8}$$

where Φ denotes the electric potential.

The total current density in the tissue consists of two components: the primary (source) current $\mathbf{J}^{(p)}$ generated by neural activity, and the Ohmic volume currents resulting from the electric field in a conductive medium:

$$\mathbf{J} = \mathbf{J}^{(p)} + \sigma \mathbf{E} = \mathbf{J}^{(p)} - \sigma \nabla \Phi \tag{2.9}$$

where σ is the tissue conductivity (typically piecewise constant across different tissue types: brain, cerebrospinal fluid, skull, scalp).

In the quasi-static regime, charge conservation requires $\nabla \cdot \mathbf{J} = 0$. Applying this to Equation (2.9) yields the Poisson equation for the potential:

$$\nabla \cdot (\sigma \nabla \Phi) = \nabla \cdot \mathbf{J}^{(p)} \tag{2.10}$$

This equation, with appropriate boundary conditions (typically Neumann conditions at the scalp-air interface), governs the relationship between neural current sources and the resulting electric potential.

2.1.2 Green's Function Solution and the Dipole Approximation

Equation (2.10) is a linear partial differential equation. Its solution can be expressed using Green's function formalism. For a point source at location $\mathbf{r}'$, the Green's function $G(\mathbf{r}, \mathbf{r}')$ satisfies:

$$\nabla \cdot (\sigma \nabla G(\mathbf{r}, \mathbf{r}')) = \delta(\mathbf{r} - \mathbf{r}') \tag{2.11}$$

with the same boundary conditions as Equation (2.10). The physical interpretation of $G(\mathbf{r}, \mathbf{r}')$ is the potential at observation point $\mathbf{r}$ due to a unit point source at $\mathbf{r}'$.

By linearity, the potential due to an arbitrary source distribution is:

$$\Phi(\mathbf{r}) = \int_{\Omega} G(\mathbf{r}, \mathbf{r}') \, \nabla' \cdot \mathbf{J}^{(p)}(\mathbf{r}') \, d^3\mathbf{r}' \tag{2.12}$$

where Ω denotes the source volume (the brain).

In neurophysiology, the sources are postsynaptic currents in pyramidal neurons, which can be modeled as current dipoles . A current dipole at location $\mathbf{r}_0$ with moment $\mathbf{q}$ is represented as:

$$\mathbf{J}^{(p)}(\mathbf{r}') = \mathbf{q} \, \delta(\mathbf{r}' - \mathbf{r}_0) \tag{2.13}$$

Substituting Equation (2.13) into (2.12) and integrating by parts yields:

$$\Phi(\mathbf{r}) = \mathbf{G}(\mathbf{r}, \mathbf{r}_0) \cdot \mathbf{q} \tag{2.14}$$

where $\mathbf{G}(\mathbf{r}, \mathbf{r}_0) = \nabla' G(\mathbf{r}, \mathbf{r}')|_{\mathbf{r}'=\mathbf{r}_0}$ is a vector-valued function. Equation (2.14) has a clear physical interpretation: the potential at $\mathbf{r}$ from a dipole at $\mathbf{r}_0$ is the dot product of the dipole moment with the gradient of the Green's function.

2.1.3 Discretization and the Leadfield Matrix

In practice, neural activity is modeled by a large number of dipoles distributed throughout the cortical volume. Discretizing the source space into N_d dipole locations $\{\mathbf{r}_j\}_{j=1}^{N_d}$, each with an associated dipole moment $\mathbf{q}_j \in \mathbb{R}^3$, the total potential at measurement electrode $\mathbf{e}_i$ is the superposition of contributions from all dipoles:

$$\Phi(\mathbf{e}_i, t) = \sum_{j=1}^{N_d} \mathbf{G}(\mathbf{e}_i, \mathbf{r}_j) \cdot \mathbf{q}_j(t) \tag{2.15}$$

For N_c electrodes, Equation (2.15) can be written in matrix form:

$$\underbrace{\begin{bmatrix} \Phi(\mathbf{e}_1, t) \\ \Phi(\mathbf{e}_2, t) \\ \vdots \\ \Phi(\mathbf{e}_{N_c}, t) \end{bmatrix}}_{\boldsymbol{\phi}(t)} = \underbrace{\begin{bmatrix} \mathbf{G}(\mathbf{e}_1, \mathbf{r}_1) & \mathbf{G}(\mathbf{e}_1, \mathbf{r}_2) & \cdots & \mathbf{G}(\mathbf{e}_1, \mathbf{r}_{N_d}) \\ \mathbf{G}(\mathbf{e}_2, \mathbf{r}_1) & \mathbf{G}(\mathbf{e}_2, \mathbf{r}_2) & \cdots & \mathbf{G}(\mathbf{e}_2, \mathbf{r}_{N_d}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{G}(\mathbf{e}_{N_c}, \mathbf{r}_1) & \mathbf{G}(\mathbf{e}_{N_c}, \mathbf{r}_2) & \cdots & \mathbf{G}(\mathbf{e}_{N_c}, \mathbf{r}_{N_d}) \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} \mathbf{q}_1(t) \\ \mathbf{q}_2(t) \\ \vdots \\ \mathbf{q}_{N_d}(t) \end{bmatrix}}_{\mathbf{j}(t)} \tag{2.16}$$

The matrix $\mathbf{L} \in \mathbb{R}^{N_c \times 3N_d}$ is the *leadfield matrix*. Each 1×3 block $\mathbf{G}(\mathbf{e}_i, \mathbf{r}_j)$ represents the sensitivity of electrode i to the three orthogonal components of the dipole at location j.

For a time series with T samples, we can stack the measurements and sources:

$$\underbrace{\begin{bmatrix} \boldsymbol{\phi}(1) & \boldsymbol{\phi}(2) & \cdots & \boldsymbol{\phi}(T) \end{bmatrix}}_{\boldsymbol{\Phi} \in \mathbb{R}^{N_c \times T}} = \mathbf{L} \underbrace{\begin{bmatrix} \mathbf{j}(1) & \mathbf{j}(2) & \cdots & \mathbf{j}(T) \end{bmatrix}}_{\mathbf{J} \in \mathbb{R}^{3N_d \times T}} \tag{2.17}$$

Equation (2.17) is the forward equation of EEG: the measured data $\boldsymbol{\Phi}$ is a linear transformation of the unknown source activity $\mathbf{J}$, mediated by the leadfield matrix $\mathbf{L}$.[12], [13]

2.1.4 Numerical Computation of the Leadfield

The computation of $\mathbf{L}$ requires solving Equation (2.10) for each dipole location and orientation. For spherical head models (typically 4 concentric spheres representing brain, CSF, skull, and scalp), analytic solutions exist using spherical harmonics expansions . However, realistic head geometries require numerical methods:

- **Boundary Element Method (BEM)**: Solves the integral formulation of Equation (2.10), discretizing only the interfaces between tissues of different conductivities. Computational cost scales as $O(N^3)$ where N is the number of boundary elements.
- **Finite Element Method (FEM)**: Discretizes the entire volume, allowing for anisotropic conductivities. It's more accurate but computationally expensive, with cost scaling as $O(N^3)$ where N is the number of volume elements.

The leadfield is computed once per subject (given MRI-derived geometry) and can be reused for all subsequent analyses. [14]

2.1.5 Dimensional Analysis

The dimensions of Equation (2.17) reveal the challenge of EEG source localization. Typical values are:

- $N_c \approx 64$ (standard clinical EEG caps) to 256 (high-density research systems)
- $N_d \approx 5000$ to 10000 (dipoles spaced ∼5–10 mm apart)
- Each dipole has 3 orientation components

Thus, $\mathbf{L} \in \mathbb{R}^{64\times 15000}$, a wide matrix with far more columns than rows. The information loss factor is big: approximatively

$$\frac{N_c}{3N_d} \approx \frac{64}{15000} \approx 0.004 \tag{2.18}$$

So only 0.4% of the information needed to uniquely determine the sources is held. This underdetermination is both a technical and physical limitation : EEG measures electric potentials far from their sources, through the blurring and attenuating effects of the resistive skull and scalp, and this necessitates the use of prior information to select plausible solutions from the infinite space of possibilities.

2.2 The Inverse Problem as Ill-Posed Matrix Inversion

The inverse problem consists of recovering the source activity $\mathbf{J}$ from the measurements $\mathbf{\Phi}$. In the presence of measurement noise, Equation (2.17) becomes:

$$\mathbf{\Phi} = \mathbf{L}\mathbf{J} + \boldsymbol{\varepsilon} \tag{2.19}$$

where $\boldsymbol{\varepsilon} \in \mathbb{R}^{N_c\times T}$ represents additive noise, typically assumed to be zero-mean Gaussian with covariance $\sigma^2\mathbf{I}$.

2.2.1 Formal Statement and Naïve Solution

A naïve approach to Equation (2.19) would be to seek the least-squares solution:

$$\hat{\mathbf{J}}_{\text{LS}} = \arg\min_{\mathbf{J}} \|\mathbf{\Phi} - \mathbf{L}\mathbf{J}\|_F^2 \tag{2.20}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. The solution is given by the pseudoinverse:

$$\hat{\mathbf{J}}_{\text{LS}} = \mathbf{L}^+\mathbf{\Phi} \tag{2.21}$$

where $\mathbf{L}^+$ is the Moore-Penrose pseudoinverse. However, this solution is physically meaningless due to the ill-posed nature of the problem.

2.2.2 Singular Value Decomposition Analysis

To understand the ill-posedness, we examine the singular value decomposition (SVD) of $\mathbf{L}$:

$$\mathbf{L} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top \tag{2.22}$$

where $\mathbf{U} \in \mathbb{R}^{N_c\times N_c}$ and $\mathbf{V} \in \mathbb{R}^{3N_d\times 3N_d}$ are orthogonal matrices, and $\mathbf{\Sigma} \in \mathbb{R}^{N_c\times 3N_d}$ is a diagonal matrix containing the singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{N_c} \geq 0$.

The pseudoinverse can be expressed in terms of the SVD:

$$\mathbf{L}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^\top, \quad \mathbf{\Sigma}^+ = \begin{bmatrix} \text{diag}(1/\sigma_1, 1/\sigma_2, \ldots, 1/\sigma_{N_c}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{2.23}$$

Now consider the true source activity expressed in the basis of the right singular vectors:

$$\mathbf{J}_{\text{true}} = \sum_{i=1}^{3N_d} \alpha_i \mathbf{v}_i \tag{2.24}$$

and the noise expressed in the basis of left singular vectors:

$$\boldsymbol{\varepsilon} = \sum_{i=1}^{N_c} \delta_i \mathbf{u}_i \tag{2.25}$$

The least-squares reconstruction becomes:

$$\hat{\mathbf{J}}_{\text{LS}} = \sum_{i=1}^{N_c} \left(\alpha_i + \frac{\delta_i}{\sigma_i} \right) \mathbf{v}_i + \sum_{i=N_c+1}^{3N_d} \alpha_i \mathbf{v}_i \tag{2.26}$$

2.2.3 The Two Faces of Ill-Posedness

Equation (2.26) reveals two problems:

Noise Amplification

For small singular values σ_i, the term δ_i/σ_i becomes enormous. Even modest noise δ_i can completely dominate the reconstruction. The condition number quantifies this sensitivity:

$$\kappa(\mathbf{L}) = \frac{\sigma_1}{\sigma_{N_c}} \gg 1 \tag{2.27}$$

For realistic EEG leadfields, $\kappa(\mathbf{L})$ is typically on the order of 10^6 to 10^8. This violates Hadamard's criterion for well-posed problems, as the solution does not depend continuously on the data.

Null Space Ambiguity

The second sum in Equation (2.26) (over $i > N_c$) represents components of the true source that lie in the null space of $\mathbf{L}$:

$$\text{Null}(\mathbf{L}) = \{\mathbf{J} \in \mathbb{R}^{3N_d} : \mathbf{L}\mathbf{J} = \mathbf{0}\} \tag{2.28}$$

These are *silent sources* that produce no measurable scalp potential. The dimension of the null space is:

$$\dim(\text{Null}(\mathbf{L})) = 3N_d - \text{rank}(\mathbf{L}) \approx 3N_d - N_c \tag{2.29}$$

For typical values, this is approximately $15000 - 64 = 14936$ which is an enormous space of indistinguishable source configurations. This, again, violates Hadamard's uniqueness criterion.

2.2.4 Regularization as Singular Value Filtering

The solution to both problems is to replace the naïve $1/\sigma_i$ factors with filtered versions:

$$\hat{\mathbf{J}}_{\text{reg}} = \sum_{i=1}^{N_c} f(\sigma_i) \left(\alpha_i + \frac{\delta_i}{\sigma_i} \right) \mathbf{v}_i \tag{2.30}$$

where $f(\sigma)$ is a filter function that tends to zero as $\sigma \to 0$. For Tikhonov regularization:

$$f_{\text{Tikhonov}}(\sigma) = \frac{\sigma^2}{\sigma^2 + \lambda} \tag{2.31}$$

where $\lambda > 0$ is the regularization parameter. This filter smoothly attenuates components corresponding to small singular values, suppressing noise amplification at the cost of introducing bias.

The physical interpretation is that one wants to trade resolution for stability. By filtering out components associated with small singular values,it is accepted that fine spatial details cannot be resolved, but robustness to noise is gained. The challenge lies in choosing the filter function and its parameters to balance these competing demands.

2.3 Bayesian Framework for Priors on Inverse Problems

The filtering perspective of Section 2.2 provides a mathematical justification for regularization but does not offer guidance on how to choose the filter function. The Bayesian framework provides a principled approach: regularization emerges naturally from the specification of prior beliefs about the sources.[16],[17]

2.3.1 Maximum a Posteriori Estimation

Both the measurements and the sources are treated as random variables. Bayes' theorem relates the posterior distribution of the sources given the data to the likelihood and prior:

$$p(\mathbf{J}|\mathbf{\Phi}) = \frac{p(\mathbf{\Phi}|\mathbf{J})\, p(\mathbf{J})}{p(\mathbf{\Phi})} \tag{2.32}$$

The maximum a posteriori (MAP) estimate seeks the source configuration that maximizes this posterior:

$$\hat{\mathbf{J}}_{\text{MAP}} = \arg\max_{\mathbf{J}} p(\mathbf{J}|\mathbf{\Phi}) = \arg\max_{\mathbf{J}} \left[\log p(\mathbf{\Phi}|\mathbf{J}) + \log p(\mathbf{J}) \right] \tag{2.33}$$

2.3.2 Gaussian Assumptions

The standard assumptions are Gaussian likelihood and Gaussian prior:

$$p(\mathbf{\Phi}|\mathbf{J}) \propto \exp\left(-\frac{1}{2\sigma^2} \|\mathbf{\Phi} - \mathbf{L}\mathbf{J}\|_F^2 \right) \tag{2.34}$$

$$p(\mathbf{J}) \propto \exp\left(-\frac{1}{2} \mathbf{j}^\top \mathbf{\Sigma}_J^{-1} \mathbf{j} \right) \tag{2.35}$$

where $\mathbf{j} = \text{vec}(\mathbf{J}) \in \mathbb{R}^{3N_dT}$ is the vectorized source activity, and $\mathbf{\Sigma}_J \in \mathbb{R}^{3N_dT \times 3N_dT}$ is the prior covariance matrix.

Substituting into Equation (2.33) and taking logarithms:

$$\hat{\mathbf{j}}_{\text{MAP}} = \arg\min_{\mathbf{j}} \left[\frac{1}{2\sigma^2} \|\boldsymbol{\phi} - \mathbf{A}\mathbf{j}\|_2^2 + \frac{1}{2} \mathbf{j}^\top \mathbf{\Sigma}_J^{-1} \mathbf{j} \right] \tag{2.36}$$

where $\boldsymbol{\phi} = \text{vec}(\mathbf{\Phi})$ and $\mathbf{A} = \mathbf{I}_T \otimes \mathbf{L}$ (Kronecker product).

2.3.3 From Prior Covariance to Regularization Operator

The solution to Equation (2.36) is given by the normal equations:

$$\left(\frac{1}{\sigma^2} \mathbf{A}^\top \mathbf{A} + \mathbf{\Sigma}_J^{-1} \right) \hat{\mathbf{j}}_{\text{MAP}} = \frac{1}{\sigma^2} \mathbf{A}^\top \boldsymbol{\phi} \tag{2.37}$$

Multiplying both sides by σ^2:

$$\left(\mathbf{A}^\top \mathbf{A} + \sigma^2 \mathbf{\Sigma}_J^{-1} \right) \hat{\mathbf{j}}_{\text{MAP}} = \mathbf{A}^\top \boldsymbol{\phi} \tag{2.38}$$

Now define the regularization operator $\mathbf{R} = \mathbf{\Sigma}_J^{-1}$ and the regularization parameter $\lambda = \sigma^2$. Then Equation (2.38) becomes:

$$\left(\mathbf{A}^\top \mathbf{A} + \lambda \mathbf{R} \right) \hat{\mathbf{j}}_{\text{MAP}} = \mathbf{A}^\top \boldsymbol{\phi} \tag{2.39}$$

This is the normal equations for a regularized inverse problem. The idea is that **the choice of regularization operator R encodes prior beliefs about the sources**.

2.3.4 Interpretation of Standard Methods as Priors

As is going to be developed in the next chapter, the work is going to focus on:

- **Tikhonov regularization**: $\mathbf{R} = \mathbf{I}$ corresponds to a prior covariance $\mathbf{\Sigma}_J = (1/\lambda)\mathbf{I}$. This encodes the belief that all dipole components are independent and identically distributed with equal variance (a spatially and temporally white prior).
- **LORETA** (Low Resolution Electromagnetic Tomography): $\mathbf{R} = \mathbf{W}^\top \mathbf{W}$ where $\mathbf{W}$ is a discrete spatial Laplacian operator. This corresponds to a prior covariance $\mathbf{\Sigma}_J = (1/\lambda)(\mathbf{W}^\top \mathbf{W})^{-1}$, encoding the belief that nearby dipoles should have similar strengths (spatial smoothness).

2.3.5 Hyperparameters as Prior Strength

The regularization parameter λ has a clear interpretation in the Bayesian framework:

$$\lambda = \sigma^2 = \frac{\text{noise variance}}{\text{prior variance}} \tag{2.40}$$

Large λ corresponds to high noise or weak prior confidence, leading to strong regularization (smooth, stable solutions). Small λ corresponds to low noise or strong prior confidence, with more detailed reconstructions, again, at the cost of stability.

For multiple regularization terms, one can have multiple hyperparameters:

$$\mathbf{R} = \lambda_s \mathbf{R}_s + \lambda_t \mathbf{R}_t \tag{2.41}$$

corresponding to separate spatial and temporal priors with their own strength parameters.

2.3.6 Evidence Maximization for Parameter Selection

The hyperparameters λ (and, in this case, λ_s and λ_t) can be chosen by maximizing the evidence (marginal likelihood):

$$p(\mathbf{\Phi}|\lambda) = \int p(\mathbf{\Phi}|\mathbf{J}, \lambda)\, p(\mathbf{J}|\lambda)\, d\mathbf{J} \tag{2.42}$$

For Gaussian models, this integral can be evaluated analytically, providing a principled way to determine regularization strength from the data itself (empirical Bayes approach).

2.3.7 Connection to Temporal Regularization

The Bayesian framework reveals the natural extension to incorporate temporal information. From Chapter 11, neural activity exhibits significant temporal correlations, and activations behave in a not independent way from one moment to the next. This physiological fact should be encoded in the prior.

A prior that captures temporal smoothness takes the form:

$$p(\mathbf{J}) \propto \exp\left(-\frac{1}{2}\left[\lambda_s \|\mathbf{J}\|_F^2 + \lambda_t \|\mathbf{D}\mathbf{J}^\top\|_F^2\right]\right) \tag{2.43}$$

where $\mathbf{D}$ is a discrete first-difference operator. The term $\|\mathbf{D}\mathbf{J}^\top\|_F^2$ penalizes rapid temporal changes, encoding the belief that sources should evolve smoothly.

Chapter 3

Spatial Priors in Already Existing Methods

3.1 Tikhonov Regularization: Minimum Norm Estimation

Tikhonov regularization, also known as minimum norm estimation (MNE), represents the most basic approach to regularizing the ill-posed EEG inverse problem. Despite its simplicity, it embodies important physical principles while revealing critical limitations that motivate more sophisticated approaches.[18]

3.1.1 Mathematical Formulation and Spectral Analysis

Formal Problem Statement

Given the discrete forward model:

$$\Phi = LJ + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma_n^2 I), \tag{3.1}$$

where $\Phi \in \mathbb{R}^{N_c \times T}$ contains scalp potentials, $L \in \mathbb{R}^{N_c \times 3N_d}$ is the leadfield matrix, and $J \in \mathbb{R}^{3N_d \times T}$ represents dipole moments, the Tikhonov-regularized estimate minimizes the following objective functional:

$$\mathcal{F}_{\text{Tik}}(J) = \|\Phi - LJ\|_F^2 + \lambda \|J\|_F^2, \tag{3.2}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, and $\lambda > 0$ is the regularization parameter balancing data fidelity against solution norm.

Time-Dimensional Separability

A crucial observation is that the objective in (3.2) separates completely across time dimensions:

$$\mathcal{F}_{\text{Tik}}(J) = \sum_{t=1}^{T} \left[\|\phi_t - L j_t\|_2^2 + \lambda \|j_t\|_2^2 \right], \tag{3.3}$$

where $\phi_t = \Phi(:, t)$ and $j_t = J(:, t)$. This separability leads to *independent* solutions for each time point:

$$\hat{j}_t = \underset{j \in \mathbb{R}^{3N_d}}{\arg\min} \left[\|\phi_t - Lj\|_2^2 + \lambda \|j\|_2^2 \right], \quad t = 1, \ldots, T. \tag{3.4}$$

Closed-Form Solution and Spectral Filtering

The minimization in (3.4) yields the well-known normal equations:

$$(L^\top L + \lambda I)\hat{j}_t = L^\top \phi_t, \quad t = 1, \ldots, T. \tag{3.5}$$

To understand the filtering properties, consider the singular value decomposition (SVD) of the leadfield matrix:

$$L = U \Sigma V^\top = \sum_{i=1}^{r} \sigma_i u_i v_i^\top, \tag{3.6}$$

where $r = \operatorname{rank}(L) \leq \min(N_c, 3N_d)$, $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, and $U \in \mathbb{R}^{N_c \times r}$, $V \in \mathbb{R}^{3N_d \times r}$ have orthonormal columns.

The Tikhonov solution admits the spectral representation:

$$\hat{j}_t = \sum_{i=1}^{r} f_i(\lambda) \frac{u_i^\top \phi_t}{\sigma_i} v_i, \quad \text{where} \quad f_i(\lambda) = \frac{\sigma_i^2}{\sigma_i^2 + \lambda}. \tag{3.7}$$

The filter factors $f_i(\lambda)$ implement a *Wiener-like* spectral filtering:

$$f_i(\lambda) \approx 1 \quad \text{for} \quad \sigma_i^2 \gg \lambda \quad \text{(well-conditioned modes preserved)}, \tag{3.8}$$

$$f_i(\lambda) \approx \frac{\sigma_i^2}{\lambda} \quad \text{for} \quad \sigma_i^2 \ll \lambda \quad \text{(ill-conditioned modes suppressed)}. \tag{3.9}$$

Computational Aspects

The temporal separability confers significant computational advantages. The regularized inverse operator:

$$M_\lambda = (L^\top L + \lambda I)^{-1} L^\top \tag{3.10}$$

needs to be computed only once, at cost $\mathcal{O}((3N_d)^3)$ for the matrix inversion. Subsequent application to each time point requires only matrix-vector multiplication:

$$\hat{j}_t = M_\lambda \phi_t, \quad \text{cost: } \mathcal{O}((3N_d)^2) \text{ per time point.} \tag{3.11}$$

For T time points, the total complexity is $\mathcal{O}((3N_d)^3 + T(3N_d)^2)$, which is linear in T.

3.1.2 Physical Interpretation: Minimum Energy Principle

Energy Minimization Perspective

The regularization term $\lambda \|J\|_F^2$ embodies the principle that among all source configurations consistent with the measurements (within the noise tolerance), select the one with minimum total "activation energy".

Define the total neural current energy as:

$$\mathcal{E}_{\text{current}} = \sum_{t=1}^{T} \|j_t\|_2^2 = \sum_{t=1}^{T} \sum_{i=1}^{3N_d} |j_i(t)|^2. \tag{3.12}$$

Tikhonov regularization thus implements a sort of *minimum energy principle* which is analogous to selecting the charge distribution with minimal electrostatic energy.

Neurobiological Justification

This principle makes possible the following considerations:

Neural activity is metabolically expensive. The brain likely minimizes unnecessary activation, consistent with efficient coding principles. Since simpler explanations (requiring less total current) are preferred, one alignes with parsimony principles in biological inference. Large current magnitudes are less probable under a Gaussian prior with zero mean and isotropic covariance so thus follows the following.

Spatial Bias and Depth Compensation

A limitation emerges from the isotropic penalty. Consider a dipole at depth d with moment q. The resulting scalp potential scales approximately as $\phi \sim q/d^2$ for a spherical conductor. To produce a given scalp potential ϕ_0, a deep dipole requires larger moment.

$$q_{\text{deep}} \approx \left(\frac{d_{\text{deep}}}{d_{\text{superficial}}} \right)^2 q_{\text{superficial}}. \tag{3.13}$$

The uniform penalty $\|j\|_2^2$ biases solutions toward superficial sources, as these need less current to explain the measurements.

Weighted minimum norm approaches attempt to compensate by introducing depth-dependent weights:

$$\mathcal{F}_{\text{WMNE}}(J) = \|\Phi - LJ\|_F^2 + \lambda \|WJ\|_F^2, \tag{3.14}$$

where typically $W = \text{diag}(\|L_i\|_2^{-1})$ normalizes by leadfield column norms. However, this adjustment remains largely ad hoc rather than physiologically grounded.

Temporal Independence Assumption

The limitation of standard Tikhonov regularization is its implicit assumption of temporal independence. This actually corresponds to the prior:

$$p(J) = \prod_{t=1}^{T} p(j_t), \quad \text{with} \quad p(j_t) = \mathcal{N}(0, \lambda^{-1} I). \tag{3.15}$$

This factorization assumes *white temporal spectrum* (complete absence of temporal correlations) which fundamentally contradicts established neurophysiological evidence presented in Chapter 1.

Information-Theoretic Consequences

Consider a signal-plus-noise model where the true neural current J_{true} exhibits temporal correlations characterized by covariance Σ_t, while the measurement noise ε is white with variance σ_n^2. The signal-to-noise ratio (SNR) in the kth temporal frequency component is:

$$\text{SNR}(f_k) = \frac{S_J(f_k)}{S_\varepsilon(f_k)}, \tag{3.16}$$

where $S_J(f)$ and $S_\varepsilon(f)$ are power spectral densities.

By solving each time point independently, Tikhonov regularization ignores the colored nature of the signal spectrum $S_J(f)$, fails to exploit frequency-dependent SNR variations and treats all temporal frequencies equally, despite differing information content

Spatial Resolution and Temporal Jitter

The temporal independence assumption also degrades spatial resolution. Consider the following two scenarios: in case of **smoothly moving source**, a dipole moving continuously through cortex over time. Independent estimation would produce a noisy, jittery trajectory rather than a smooth path. For a **sustained activation** instead, a dipole would maintain approximately constant moment over several time samples. Independent estimates scatter around the true value, reducing precision.

The variance of the Tikhonov estimate at each time point is:

$$\text{Var}(\hat{j}_t) = \sigma_n^2 (L^\top L + \lambda I)^{-1} L^\top L (L^\top L + \lambda I)^{-1}. \tag{3.17}$$

Nothing in the method tries to reduce this variance by leveraging temporal smoothness.

Ideal Characterization of Limitations

Consider the Tikhonov estimator $\hat{J}_{\text{Tik}}$ and an oracle estimator $\hat{J}_{\text{oracle}}$ that knows the true temporal covariance Σ_t. For a Gaussian source with temporal covariance Σ_t, the mean squared error ratio satisfies:

$$\frac{\mathbb{E}\,\|\hat{J}_{\text{Tik}} - J_{\text{true}}\|_F^2}{\mathbb{E}\,\|\hat{J}_{\text{oracle}} - J_{\text{true}}\|_F^2} \geq 1 + \frac{\text{tr}(\Sigma_t - \bar{\sigma}_t^2 I)^2}{T \bar{\sigma}_t^4}, \tag{3.18}$$

where $\bar{\sigma}_t^2 = \text{tr}(\Sigma_t)/T$. Equality holds only when $\Sigma_t = \bar{\sigma}_t^2 I$ (temporally white).

Proof. The proof follows from the Gauss-Markov theorem and the fact that Tikhonov is the best linear unbiased estimator only under temporal whiteness. The additional term quantifies the efficiency loss due to ignoring temporal correlations. □

Connection to Broader Inverse Problem Theory

Tikhonov regularization represents a special case of the general regularization framework:

$$\hat{J} = \arg\min_J \left[\|\Phi - LJ\|_F^2 + \lambda \langle J, RJ \rangle \right], \tag{3.19}$$

where R is a regularization operator. As previously defined, standard Tikhonov uses $R = I$, implementing an isotropic prior.

The idea is that the choice $R = I$ corresponds to assuming: Spatial independence between dipoles Equal variance across all locations Temporal independence across samples No preferred orientation

Each of these assumptions is physiologically questionable, motivating more structured regularization operators.

3.2 LORETA: Spatial Smoothness via Laplacian Operator

Low Resolution Electromagnetic Tomography (LORETA) represents an advancement over Tikhonov regularization by incorporating spatial smoothness constraints. It addresses important limitations of minimum norm estimation, it retains the assumption of temporal independence, motivating the need for more comprehensive spatio-temporal regularization.[19]

3.2.1 Mathematical Formulation: From Tikhonov to Spatial Smoothness

Continuous Formulation

The LORETA approach originates from the continuous regularization framework:

$$\mathcal{F}_{\text{LORETA}}[J] = \int_0^T \left[\|\phi(t) - \mathcal{L}J(\cdot, t)\|^2 + \lambda \int_\Omega \| \nabla^2 J(\mathbf{r}, t)\|^2 d^3\mathbf{r} \right] dt, \tag{3.20}$$

where $\mathcal{L}$ is the continuous forward operator and ∇^2 denotes the spatial Laplacian operator. The key innovation is the replacement of the L^2 norm penalty $\|J\|^2$ with a Sobolev-type penalty $\| \nabla^2 J\|^2$, enforcing spatial smoothness.

Discrete Laplacian Operator Construction

For a discrete set of dipole locations $\{\mathbf{r}_i\}_{i=1}^{N_d}$, we construct the graph Laplacian matrix $L_s \in \mathbb{R}^{N_d \times N_d}$. Consider the weighted adjacency matrix A with elements:

$$A_{ij} = \begin{cases} w_{ij}, & \text{if } \|\mathbf{r}_i - \mathbf{r}_j\| \leq \rho \\ 0, & \text{otherwise} \end{cases} \tag{3.21}$$

where ρ is a neighborhood radius and w_{ij} are weights typically chosen as:

$$w_{ij} = \frac{1}{\|\mathbf{r}_i - \mathbf{r}_j\|^2} \quad \text{or} \quad w_{ij} = \exp\left(-\frac{\|\mathbf{r}_i - \mathbf{r}_j\|^2}{2\sigma_s^2} \right). \tag{3.22}$$

The degree matrix D is diagonal with $D_{ii} = \sum_{j=1}^{N_d} A_{ij}$. The graph Laplacian is then:

$$L_s = D - A. \tag{3.23}$$

For the full source space including three dipole orientations, we construct the block Laplacian:

$$B = L_s \otimes I_3 \in \mathbb{R}^{3N_d \times 3N_d}, \tag{3.24}$$

where $\otimes$ denotes the Kronecker product.

LORETA Optimization Problem

The discrete LORETA objective function is:

$$\mathcal{F}_{\text{LORETA}}(J) = \|\Phi - LJ\|_F^2 + \lambda \|BJ\|_F^2, \tag{3.25}$$

where $J \in \mathbb{R}^{3N_d \times T}$. The minimization yields the normal equations:

$$(L^\top L + \lambda B^\top B)\hat{j}_t = L^\top \phi_t, \quad t = 1, \ldots, T. \tag{3.26}$$

Note the critical structural feature: the equations remain *decoupled across time*, analogous to Tikhonov regularization.

Kronecker Formulation and Temporal Separability

The temporal independence becomes explicit in Kronecker notation. Define the vectorized variables:

$$\mathbf{j} = \text{vec}(J) \in \mathbb{R}^{3N_d T \times 1}, \tag{3.27}$$

$$\boldsymbol{\phi} = \text{vec}(\Phi) \in \mathbb{R}^{N_c T \times 1}. \tag{3.28}$$

The forward operator becomes $A = I_T \otimes L \in \mathbb{R}^{N_c T \times 3N_d T}$, and the LORETA regularization operator is:

$$R_{\text{LORETA}} = I_T \otimes (B^\top B) \in \mathbb{R}^{3N_d T \times 3N_d T}. \tag{3.29}$$

The Kronecker structure $I_T \otimes (\cdot)$ explicitly shows temporal independence: the same spatial regularization is applied independently at each time point.

3.2.2 Physiological Interpretation: Cortical Coherence Constraints

Neuroanatomical Basis for Spatial Smoothness

LORETA's spatial smoothness prior finds strong support in cortical neuroanatomy. First and foremost, the cerebral cortex exhibits **columnar organization** with approximately 1 mm diameter columns containing 10 000-15 000 neurons. Activity within a column tends to be homogeneous. Furthermore pyramidal neurons extend horizontal axons over several millimeters, creating local **lateral connections** that synchronize nearby cortical regions. And also sensory and motor cortices contain **functional maps**: (retinotopic, tonotopic, somatotopic) where neighboring cortical locations represent neighboring sensory/motor features.

Mathematical Expression of Cortical Coherence

The graph Laplacian L_s encodes the assumption that neural activity should vary smoothly across the cortical surface. For a cortical patch, consider the continuous Laplacian in local coordinates:

$$\Delta J(x,y) = \frac{\partial^2 J}{\partial x^2} + \frac{\partial^2 J}{\partial y^2}, \tag{3.30}$$

where (x, y) are surface coordinates. The discrete approximation on a cortical mesh gives:

$$(\Delta J)_i \approx \sum_{j \in \mathcal{N}(i)} w_{ij}(J_j - J_i), \tag{3.31}$$

where $\mathcal{N}(i)$ denotes neighbors of vertex i. The penalty $\|BJ\|^2$ thus encourages:

$$\sum_i \left(\sum_{j \in \mathcal{N}(i)} w_{ij}(J_j - J_i) \right)^2 \to \text{minimal}. \tag{3.32}$$

Comparison with Tikhonov: Depth Bias Reduction

A significant advantage of LORETA over Tikhonov is the reduction of depth bias. Consider the eigenanalysis of the regularization operators:

Tikhonov Temporal Inefficiency (Depth Invariance of Laplacian Regularization)**.** *For a spherical conductor model, the graph Laplacian L_s is approximately depth-invariant, while the identity operator I strongly weights superficial sources.*

Proof. Let $G(\mathbf{r}_i, \mathbf{r}_j)$ be the leadfield Green's function. For Tikhonov:

$$(I)_{ii} = 1 \quad \text{(uniform weight)}.$$

For LORETA:

$$(L_s)_{ii} = \sum_{j \in \mathcal{N}(i)} w_{ij} \approx \text{local cortical connectivity, independent of depth.}$$

The depth dependence enters only through the forward model L, not through B. □

3.2.3 Extensions: sLORETA and eLORETA

The sLORETA approach addresses LORETA's limited spatial resolution through post-hoc standardization. Define the resolution matrix:

$$R = (L^\top L + \lambda B^\top B)^{-1} L^\top L. \tag{3.33}$$

The variance of the ith source estimate is:

$$\text{Var}(\hat{j}_i) = \sigma_n^2 [R(L^\top L + \lambda B^\top B)^{-1}]_{ii}. \tag{3.34}$$

sLORETA computes standardized scores:

$$\hat{j}_i^{\text{sLORETA}} = \frac{\hat{j}_i}{\sqrt{\text{Var}(\hat{j}_i)}}. \tag{3.35}$$

Tikhonov Temporal Inefficiency (sLORETA Property)**.** *Under ideal conditions (single point source, no noise), sLORETA achieves zero localization error: the maximum of $|\hat{j}_i^{sLORETA}|$ occurs exactly at the true source location.*

Exact LORETA (eLORETA)

eLORETA introduces adaptive weighting to achieve exact localization for single sources even in the presence of structured noise. The weighted Laplacian is:

$$B_W = W^{1/2} B W^{1/2}, \tag{3.36}$$

where W is a diagonal weight matrix determined iteratively to satisfy:

$$[W]_{ii} = \left(\|L_i\| \cdot \|[B\hat{j}]_i\| \right)^{-1}. \tag{3.37}$$

The eLORETA objective becomes:

$$\mathcal{F}_{\text{eLORETA}}(J) = \|\Phi - LJ\|_F^2 + \lambda \|B_W J\|_F^2. \tag{3.38}$$

Limitations of the LORETA Family

Despite these refinements, all LORETA variants share a common limitation:

Tikhonov Temporal Inefficiency (Temporal Independence in LORETA Family)**.** *Any regularization of the form $R = I_T \otimes S$, where S is a spatial operator, necessarily assumes temporal independence of sources.*

Proof. The prior covariance for such regularization is:

$$\Sigma^{-1} = \lambda (I_T \otimes S^\top S).$$

This Kronecker structure implies:

$$\text{Cov}(j_t, j_s) = 0 \quad \text{for} \quad t \neq s,$$

meaning temporal independence. □

3.2.4 Information-Theoretic Analysis

Spatial vs. Temporal Information

Consider decomposing the mutual information between sources and measurements:

$$I(J; \Phi) = I(J_{\text{space}}; \Phi) + I(J_{\text{time}}; \Phi | J_{\text{space}}) \tag{3.39}$$

$$= \underbrace{H(J_{\text{space}}) - H(J_{\text{space}}|\Phi)}_{\text{Spatial information}} + \underbrace{H(J_{\text{time}}|J_{\text{space}}) - H(J_{\text{time}}|J_{\text{space}}, \Phi)}_{\text{Temporal information}}. \tag{3.40}$$

LORETA exploits only the spatial information term, ignoring the temporal information that could be gained from the conditional entropy terms.

Spectral Domain Interpretation

In the frequency domain, LORETA's spatial regularization corresponds to:

$$\tilde{J}(\mathbf{k}, \omega) = \frac{\tilde{L}^\dagger(\mathbf{k})}{|\tilde{L}(\mathbf{k})|^2 + \lambda |\mathbf{k}|^4} \tilde{\Phi}(\mathbf{k}, \omega), \tag{3.41}$$

where $\mathbf{k}$ is spatial wavevector and ω is temporal frequency. The $|\mathbf{k}|^4$ term suppresses high spatial frequencies, but no corresponding suppression occurs for high temporal frequencies ω.

3.2.5 Numerical Comparison with Tikhonov

Spatial Resolution Comparison

Define the point spread function (PSF) for a source at location $\mathbf{r}_0$:

$$\mathrm{PSF}(\mathbf{r}; \mathbf{r}_0) = [M_\lambda \delta_{\mathbf{r}_0}](\mathbf{r}), \tag{3.42}$$

where $M_\lambda = (L^\top L + \lambda R)^{-1} L^\top$ and $\delta_{\mathbf{r}_0}$ is a unit source at $\mathbf{r}_0$.

For Tikhonov ($R = I$), the PSF exhibits: Broad spatial spread; Strong depth dependence; Maximum often displaced from true location;

For LORETA ($R = B^\top B$), the PSF shows: More compact spatial spread; Reduced depth bias; Better peak localization;

Temporal Artifacts

Consider a smoothly varying source $J_{\text{true}}(t) = \sin(2\pi f_0 t)$. The LORETA reconstruction $\hat{J}_{\text{LORETA}}(t)$ exhibits:

Independent estimation at each t introduces uncorrelated noise No mechanism to enforce temporal smoothness Spectral leakage of high-frequency estimation errors

The mean squared error can be decomposed as:

$$\mathrm{MSE} = \underbrace{\mathrm{MSE}_{\text{spatial}}}_{\text{Reduced by LORETA}} + \underbrace{\mathrm{MSE}_{\text{temporal}}}_{\text{Ignored by LORETA}}. \tag{3.43}$$

3.2.6 Theoretical Limitations and Motivation for Temporal Regularization

Violation of the Assumption

LORETA, like Tikhonov, violates the neurophysiological principle established in Chapter 1 regarding *temporal correlations* across multiple timescales. The assumption of temporal independence corresponds to a delta-function temporal autocorrelation:

$$R_{JJ}(\tau) = \mathbb{E}[J(t)J(t+\tau)] = \sigma_J^2 \delta(\tau), \tag{3.44}$$

whereas real neural signals typically show:

$$R_{JJ}(\tau) = \sigma_J^2 \exp\left(-\frac{|\tau|}{\tau_c}\right) \quad \text{(exponential decay)}, \tag{3.45}$$

or power-law decay for scale-free dynamics.

Information-Theoretic Bound

Tikhonov Temporal Inefficiency (LORETA Temporal Information Loss)**.** *For a source with temporal correlation time τ_c, the ratio of achievable SNR with and without temporal regularization satisfies:*

$$\frac{\mathrm{SNR}_{\textit{with temporal}}}{\mathrm{SNR}_{\textit{LORETA}}} \geq 1 + \frac{\tau_c}{\Delta t} \cdot \frac{\int |\tilde{L}(\mathbf{k})|^2 d\mathbf{k}}{\int |\tilde{L}(\mathbf{k})|^2/(1+\lambda|\mathbf{k}|^4) d\mathbf{k}}, \tag{3.46}$$

where Δt is the sampling interval.

This bound shows that ignoring temporal correlations necessarily discards information proportional to the correlation time τ_c.

Computational Implications

While LORETA's temporal separability offers computational advantages ($\mathcal{O}(T)$ scaling), this comes at the cost of statistical efficiency. The Cramér-Rao bound for estimating a smooth temporal trajectory with T samples is:

$$\mathrm{Var}(\hat{J}(t)) \geq \frac{\sigma_n^2}{T} \cdot \frac{1}{\mathrm{SNR}_{\mathrm{eff}}}, \tag{3.47}$$

where $\mathrm{SNR}_{\mathrm{eff}}$ incorporates temporal correlations. By treating samples independently, LORETA achieves only:

$$\mathrm{Var}(\hat{J}_{\mathrm{LORETA}}(t)) \geq \frac{\sigma_n^2}{1} \cdot \frac{1}{\mathrm{SNR}_{\mathrm{instant}}}, \tag{3.48}$$

losing the $1/T$ variance reduction factor from temporal averaging.

The development of LORETA (1994) and its extensions (sLORETA 2002, eLORETA 2007) represents significant progress in spatial regularization for EEG source localization. However, the field's focus has remained predominantly on spatial aspects, with temporal regularization receiving comparatively little attention despite its physiological importance.

Chapter 4

Temporal Regularization: Theory and Implementation

EEG source localization is required for reconstructing $3N_dxT$ unknown dipole moments from only N_cT measurements, where typically $N_c \ll 3N_d$. This under-determination requires regularization.

4.1 Mathematical Description of Temporal Priors

4.1.1 Continuous Formulation: First-Derivative Penalty

Neural activations evolve smoothly over time due to biophysical constraints such as synaptic time constants and neural adaptation. From an operational perspective, we can formulate this as minimizing the *temporal variation* of the source field.

$$\mathcal{R}_t^{\text{phys}}[\mathbf{J}] = \int_\Omega \int_0^T \left\| \frac{\partial \mathbf{J}(\mathbf{r}, t)}{\partial t} \right\|_2^2 dt\, dV \tag{4.1}$$

where $\mathbf{J}(\mathbf{r}, t) \in \mathbb{R}^3$ represents the current dipole density at position $\mathbf{r}$ and time t. This choice of first-derivative penalty has several physical justifications:

Firstly it has an **energy minimization analogy**: the term $\|\partial_t \mathbf{J}\|^2$ is analogous to kinetic energy in classical mechanics, preferring solutions with minimal "temporal motion."

there is also a **bayesian interpretation**, (a Gaussian prior on temporal derivatives, assuming neural activations change gradually) and a **Physiological basis**, since synaptic transmission and membrane time constants (typically 10-100 ms) naturally low-pass filter neural signals.

An alternative would be to penalize the second derivative $\|\partial_t^2 \mathbf{J}\|^2$, which would enforce even smoother solutions but is computationally more complex and less physically interpretable as a prior on neural dynamics.

4.1.2 Discrete Approximation and Finite Differences

To implement this continuous prior numerically, we discretize time into T samples with spacing Δt. Using first-order forward differences:

$$\frac{\partial \mathbf{J}(t_k)}{\partial t} \approx \frac{\mathbf{J}(t_{k+1}) - \mathbf{J}(t_k)}{\Delta t}, \quad k = 0, \ldots, T-2 \tag{4.2}$$

The discrete penalty becomes:

$$\mathcal{R}_t[\mathbf{J}] = \sum_{k=0}^{T-2} \|\mathbf{J}_{k+1} - \mathbf{J}_k\|_2^2 \tag{4.3}$$

where $\mathbf{J}_k \in \mathbb{R}^{3N_d}$ represents the dipole moments at time sample k. The factor Δt^2 is omitted as it can be absorbed into the regularization parameter λ_t.

For individual dipole components ($i = 1, \ldots, 3N_d$), this expands to:

$$\mathcal{R}_t[\mathbf{J}] = \sum_{i=1}^{3N_d} \sum_{k=0}^{T-2} (J_{i,k+1} - J_{i,k})^2 \tag{4.4}$$

This formulation treats all dipole components independently in the temporal prior, which is physically reasonable assuming different neural populations evolve independently (though spatially coupled through the leadfield).

4.1.3 Matrix Representation: First-Difference Operator

The discrete penalty can be expressed compactly using matrix notation. Define the first-difference operator $\mathbf{D}_1 \in \mathbb{R}^{(T-1)\times T}$:

$$\mathbf{D}_1 = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix} \tag{4.5}$$

For a single dipole component's time course $\mathbf{j} \in \mathbb{R}^T$, we have:

$$\|\mathbf{D}_1 \mathbf{j}\|_2^2 = \sum_{k=0}^{T-2} (j_{k+1} - j_k)^2 \tag{4.6}$$

To handle all time points uniformly (including boundaries), we augment $\mathbf{D}_1$ to a square matrix $\mathbf{D} \in \mathbb{R}^{T\times T}$ by adding a zero row:

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & 1 \\ 0 & \cdots & 0 & 0 & 0 \end{bmatrix} \tag{4.7}$$

This ensures $\mathbf{Dj}$ has length T, with the last element zero, maintaining dimensional consistency.

4.1.4 Temporal Smoothness Matrix R

The temporal regularization term can be written as a quadratic form:

$$\mathcal{R}_t[\mathbf{J}] = \mathrm{tr}(\mathbf{J}\mathbf{R}\mathbf{J}^\top) \tag{4.8}$$

where $\mathbf{R} = \mathbf{D}^\top\mathbf{D} \in \mathbb{R}^{T\times T}$ is the temporal smoothness matrix:

$$\mathbf{R} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix} \tag{4.9}$$

For a single dipole's time course $\mathbf{j} \in \mathbb{R}^T$:

$$\|D\mathbf{j}\|^2 = \mathbf{j}^\top(D^\top D)\mathbf{j} = \mathbf{j}^\top R\mathbf{j}$$

Physical interpretation of R: There is a **diagonal dominance**, as the positive diagonal entries (1 or 2) penalize large activation magnitudes at individual time points.

Negative off-diagonals: The -1 entries create coupling between neighboring time points, encouraging temporal smoothness.

Boundary conditions: The first and last diagonal entries are 1 (not 2), implementing boundary conditions where only one neighbor exists.

Energy decomposition: For a time series $\mathbf{j}$:

$$\mathbf{j}^\top\mathbf{R}\mathbf{j} = j_0^2 + \sum_{k=1}^{T-2}(j_k - j_{k-1})^2 + j_{T-1}^2$$

which shows the penalty on differences between consecutive time points.

Crucial distinction: Although $\mathbf{R}$ has the form of a discrete second-difference operator, it arises from squaring the *first*-difference operator $\mathbf{D}$, not from approximating a second derivative. The continuous prior being implemented is $\int(\partial_t J)^2 dt$, not $\int(\partial_t^2 J)^2 dt$, nor $\int \partial_t^2(J)^2 dt$.

4.2 Formulation of SSTR Regularization

Follows the formalism of Custom regularisation method, which is going to be called SSTR as in Simultaneous Spatio-Temporal Regularisation. The reader will be able to appreciate the choice of the name from the following implementation.

4.2.1 Complete Optimization Problem

Combining the data fidelity term with both spatial and temporal regularization yields the complete objective function:

$$\min_{\mathbf{J}\in\mathbb{R}^{3N_d\times T}} \left\{ \|\mathbf{\Phi} - \mathbf{L}\mathbf{J}\|_F^2 + \lambda_s\|\mathbf{J}\|_F^2 + \lambda_t\|\mathbf{D}\mathbf{J}^\top\|_F^2 \right\} \tag{4.10}$$

where:

- $\mathbf{\Phi} \in \mathbb{R}^{N_c \times T}$: EEG measurements (channels $\times$ time)
- $\mathbf{L} \in \mathbb{R}^{N_c \times 3N_d}$: Leadfield matrix
- $\mathbf{J} \in \mathbb{R}^{3N_d \times T}$: Dipole moments (dipoles $\times$ time)
- $\lambda_s, \lambda_t > 0$: Spatial and temporal regularization parameters

The Frobenius norm $\|\cdot\|_F$ is used as it naturally handles the matrix structure while being equivalent to the ℓ_2 norm when vectorized.

4.2.2 Vectorization and Kronecker Formulation

To derive the solution, we vectorize the matrices. Let:

$$\mathbf{j} = \text{vec}(\mathbf{J}) \in \mathbb{R}^{3N_d T \times 1} \tag{4.11}$$
$$\boldsymbol{\phi} = \text{vec}(\mathbf{\Phi}) \in \mathbb{R}^{N_c T \times 1} \tag{4.12}$$
$$\mathbf{A} = \mathbf{I}_T \otimes \mathbf{L} \in \mathbb{R}^{N_c T \times 3N_d T} \tag{4.13}$$

where $\otimes$ denotes the Kronecker product. The objective becomes:

$$\|\boldsymbol{\phi} - \mathbf{A}\mathbf{j}\|_2^2 + \lambda_s \|\mathbf{j}\|_2^2 + \lambda_t \mathbf{j}^\top (\mathbf{D}^\top \mathbf{D} \otimes \mathbf{I}_{3N_d})\mathbf{j} \tag{4.14}$$

The Kronecker structure $\mathbf{A} = \mathbf{I}_T \otimes \mathbf{L}$ encodes the physical assumption that the forward model $\mathbf{L}$ is time-invariant (the head geometry doesn't change during measurement).

4.2.3 Normal Equations and Solution

Setting the gradient of (4.14) to zero yields the normal equations:

$$\left[\mathbf{A}^\top \mathbf{A} + \lambda_s \mathbf{I}_{3N_d T} + \lambda_t (\mathbf{D}^\top \mathbf{D} \otimes \mathbf{I}_{3N_d})\right] \mathbf{j} = \mathbf{A}^\top \boldsymbol{\phi} \tag{4.15}$$

Using properties of Kronecker products ($\mathbf{A}^\top \mathbf{A} = \mathbf{I}_T \otimes (\mathbf{L}^\top \mathbf{L})$ and $\mathbf{A}^\top \boldsymbol{\phi} = \text{vec}(\mathbf{L}^\top \mathbf{\Phi})$), we obtain the computationally efficient form:

$$\left[\mathbf{I}_T \otimes (\mathbf{L}^\top \mathbf{L} + \lambda_s \mathbf{I}_{3N_d}) + \lambda_t (\mathbf{D}^\top \mathbf{D}) \otimes \mathbf{I}_{3N_d}\right] \mathbf{j} = \text{vec}(\mathbf{L}^\top \mathbf{\Phi}) \tag{4.16}$$

This is the main equation of the SSTR method. Its structure show some properties like

Separability (the system matrix is the sum of two Kronecker products, enabling efficient storage and computation);

Symmetry (both terms are symmetric positive definite (for $\lambda_s, \lambda_t > 0$), ensuring a unique solution);

Block structure (the Kronecker form naturally partitions the problem by time points).

4.2.4 Block Matrix Structure and Temporal Coupling

To understand the temporal coupling explicitly, partition $\mathbf{j}$ by time:

$$\mathbf{j} = [\mathbf{j}_0^\top, \mathbf{j}_1^\top, \ldots, \mathbf{j}_{T-1}^\top]^\top, \quad \mathbf{j}_k \in \mathbb{R}^{3N_d} \tag{4.17}$$

Equation (4.16) then becomes a block-tridiagonal system:

$$\begin{bmatrix} \mathbf{B} & \mathbf{C} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{C} & \mathbf{B} & \mathbf{C} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C} & \mathbf{B} & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{C} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{j}_0 \\ \mathbf{j}_1 \\ \mathbf{j}_2 \\ \vdots \\ \mathbf{j}_{T-1} \end{bmatrix} = \begin{bmatrix} \mathbf{b}_0 \\ \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_{T-1} \end{bmatrix} \tag{4.18}$$

where:

$$\mathbf{B} = \mathbf{L}^\top \mathbf{L} + (\lambda_s + 2\lambda_t)\mathbf{I}_{3N_d} \quad \text{(interior points)} \tag{4.19}$$

$$\mathbf{C} = -\lambda_t \mathbf{I}_{3N_d} \tag{4.20}$$

$$\mathbf{b}_k = (\mathbf{L}^\top \mathbf{\Phi})_{:,k} \tag{4.21}$$

For interior time points $1 \leq k \leq T-2$, we obtain the explicit coupling equation:

$$-\lambda_t \mathbf{j}_{k-1} + [\mathbf{L}^\top \mathbf{L} + (\lambda_s + 2\lambda_t)\mathbf{I}_{3N_d}]\mathbf{j}_k - \lambda_t \mathbf{j}_{k+1} = (\mathbf{L}^\top \mathbf{\Phi})_{:,k} \tag{4.22}$$

This has the consequence that the solution at time k depends directly on solutions at times $k-1$ and $k+1$, with coupling strength controlled by λ_t. This creates a *temporal stiffness* that smooths the solution across time.

Boundary conditions ($k = 0$ and $k = T-1$) have modified equations with only one neighboring term, implementing natural boundary conditions where temporal smoothness is enforced only within the available time window.

4.3 Computational Implementation and Efficiency

4.3.1 MATLAB Implementation Core

The method is implemented efficiently in MATLAB using Kronecker algebra and chunk processing:

Listing 4.1: Core implementation of temporal regularization in EEG source localization. The code implements Equations (7)-(9) from the mathematical formulation, using Kronecker products for efficient computation of temporal derivatives and Cholesky decomposition for solving the regularized inverse problem.

```
lambda_s = 1e3;      % Spatial regularization parameter
lambda_t = 1e2;      % Temporal regularization parameter
n_time_chunk = 63;  % Time points per processing chunk
% Construct temporal difference matrix D (Eq. 3)
D = diff(eye(n_time_chunk));% First differences: D(i,i) = -1, D(i,i+1)=1
D = [D; zeros(1, n_time_chunk)];
% Apply to all dipole components via Kronecker product (Eq. 7)
R_t = kron(D, eye(size(L,2)/n_time_chunk));
% Complete system matrix (Eq. 8):
```

```
M = (L'*L + lambda_s*eye(size(L,2)) + lambda_t*(R_t'*R_t));
% Solve using Cholesky decomposition for efficiency and stability
R = chol(M);  % Upper triangular Cholesky factor
j_est = R \ (R' \ (L'*Phi(:)));  % (two triangular solves)
% Reshape solution to original physical dimensions
J_est = reshape(j_est, [size(L,2)/3, 3, n_time_chunk]);
```

4.3.2 Memory Efficiency with Chunk Processing

For realistic problems ($N_d \approx 5000$, $T \approx 1000$), the system matrix $\mathbf{M} \in \mathbb{R}^{15\text{M} \times 15\text{M}}$ would require ∼180 TB of memory if stored densely. We overcome this through chunk processing:

Algorithm 1 Memory-Efficient Chunk Processing

Require: EEG data $\mathbf{\Phi}$, leadfield $\mathbf{L}$, chunk size T_c
Ensure: Reconstructed sources $\mathbf{J}$

1: Partition time: $T = T_1 + T_2 + \cdots + T_m$ with overlap
2: **for** each chunk $c = 1$ to m **do**
3: Extract chunk $\mathbf{\Phi}_c = \mathbf{\Phi}_{:,\text{start}_c:\text{end}_c}$
4: Solve $\mathbf{J}_c = \arg\min \|\mathbf{\Phi}_c - \mathbf{L}\mathbf{J}_c\|^2 + \lambda_s\|\mathbf{J}_c\|^2 + \lambda_t\|\mathbf{D}\mathbf{J}_c^\top\|^2$
5: Store $\mathbf{J}_c$ (excluding overlap regions)
6: **end for**
7: Stitch chunks with smooth blending in overlap regions

The computational complexity is $O(N_d^2 T + N_d T^2)$ compared to the naive $O(N_d^3 T^3)$, making it feasible for practical applications.

4.3.3 Kronecker Product Advantages

The Kronecker formulation provides several computational benefits:

- **Efficient storage**: Store $\mathbf{D}$ $(T \times T)$ and $\mathbf{I}_{3N_d}$ $(3N_d \times 3N_d)$ separately rather than $\mathbf{D}^\top\mathbf{D} \otimes \mathbf{I}_{3N_d}$ $(3N_dT \times 3N_dT)$.
- **Fast multiplication**: For any vector $\mathbf{x} = \text{vec}(\mathbf{X})$:
$$(\mathbf{D}^\top\mathbf{D} \otimes \mathbf{I})\mathbf{x} = \text{vec}(\mathbf{I}\mathbf{X}\mathbf{D}^\top\mathbf{D})$$
which requires only matrix-matrix multiplications rather than explicit Kronecker product formation.
- **Exploitable structure**: The block-tridiagonal system can be solved efficiently using the Thomas algorithm with complexity $O(N_d^3 T)$ instead of $O(N_d^3 T^3)$.

4.3.4 Parameter Selection Heuristics

Choosing appropriate λ_s and λ_t is crucial. Several strategies may be employed:

1. **Physical scaling**:
$$\frac{\lambda_t}{\lambda_s} \approx \frac{(\text{temporal correlation scale})^2}{(\text{spatial correlation scale})^2} \approx \frac{(100\text{ ms})^2}{(10\text{ mm})^2} \approx 10^{-2}$$
based on typical neural correlation scales.

2. **L-curve method**: Plot $\|\mathbf{\Phi} - \mathbf{LJ}\|$ vs $\|\mathbf{DJ}^\top\|$ for a grid of (λ_s, λ_t) values and select parameters at the "corner" of the L-shaped curve.

3. **Cross-validation**: For synthetic data with known ground truth, minimize reconstruction error:
$$(\lambda_s^*, \lambda_t^*) = \arg\min_{\lambda_s,\lambda_t} \|\mathbf{J}_{\text{true}} - \mathbf{J}(\lambda_s, \lambda_t)\|$$

4. **Empirical values**: For typical EEG setups ($N_c = 64$, $N_d = 5000$, $f_s = 250$ Hz):
$$\lambda_s \approx 10^3 - 10^4, \quad \lambda_t \approx 10^2 - 10^3$$

4.4 Physical and Physiological Interpretation

4.4.1 Bayesian Interpretation

The regularization approach has a natural Bayesian interpretation. Let:

$$p(\mathbf{\Phi}|\mathbf{J}) \propto \exp\left(-\frac{1}{2\sigma_n^2}\|\mathbf{\Phi} - \mathbf{LJ}\|_F^2\right) \quad \text{(Likelihood)} \tag{4.23}$$

$$p(\mathbf{J}) \propto \exp\left(-\frac{1}{2}\mathbf{j}^\top \mathbf{\Sigma}^{-1}\mathbf{j}\right) \quad \text{(Prior)} \tag{4.24}$$

where the prior precision matrix is:

$$\mathbf{\Sigma}^{-1} = \lambda_s \mathbf{I}_{3N_d T} + \lambda_t(\mathbf{D}^\top\mathbf{D} \otimes \mathbf{I}_{3N_d}) \tag{4.25}$$

The maximum a posteriori (MAP) estimate is then:

$$\mathbf{J}_{\text{MAP}} = \arg\max_{\mathbf{J}} p(\mathbf{J}|\mathbf{\Phi}) = \arg\min_{\mathbf{J}} \left\{\|\mathbf{\Phi} - \mathbf{LJ}\|_F^2 + \mathbf{j}^\top\mathbf{\Sigma}^{-1}\mathbf{j}\right\} \tag{4.26}$$

which is the regularized solution. The prior encodes the belief: "Neural activations have small amplitude (controlled by λ_s) and evolve smoothly over time (controlled by λ_t)."

4.4.2 Temporal Coupling as Neural Constraint

The coupling equation (4.22) has a physical analogy: imagine each time point's activation $\mathbf{j}_k$ connected to its neighbors by springs with stiffness λ_t. The system seeks a configuration that minimizes both the spring energy ($\lambda_t\|\mathbf{j}_{k+1} - \mathbf{j}_k\|^2$) and the data misfit.

This matches neurophysiological reality: Synaptic transmission has characteristic time constants (AMPA: 2-5 ms, NMDA: 50-100 ms, GABA: 5-10 ms) that naturally smooth neural responses;

Neural adaptation and fatigue cause gradual rather than abrupt changes in firing rates.;

Cortical dynamics often exhibit smooth trajectories in state space, not random jumps.

4.4.3 Spectral Analysis: Temporal Low-Pass Filtering

The eigenvalues of $\mathbf{R} = \mathbf{D}^\top \mathbf{D}$ reveal the frequency characteristics of the temporal prior:

$$\lambda_m = 2 - 2\cos\left(\frac{\pi m}{T}\right), \quad m = 0, 1, \ldots, T-1 \tag{4.27}$$

- $\lambda_0 = 0$: Constant activation (perfectly smooth, no penalty)
- Small λ_m ($m \ll T$): Low-frequency patterns (weakly penalized)
- Large λ_m ($m \approx T$): High-frequency patterns (strongly penalized)

The regularization thus acts as a *temporal low-pass filter* with cutoff frequency approximately:

$$f_c \approx \frac{1}{2\pi\Delta t}\sqrt{\frac{\lambda_s}{\lambda_t}} \tag{4.28}$$

For typical parameters ($\Delta t = 4$ ms, $\lambda_s/\lambda_t = 10$), $f_c \approx 12.6$ Hz, which matches the physiologically relevant range for event-related potentials and oscillatory activity.

4.4.4 Connection to Neural Mass Models

Consider the linearized Wilson-Cowan equations around a fixed point:

$$\tau\frac{d\mathbf{x}(t)}{dt} = -\mathbf{x}(t) + \mathbf{W}\mathbf{x}(t) + \mathbf{I}(t) \tag{4.29}$$

where $\mathbf{x}(t)$ represents neural population activity, τ is the time constant, $\mathbf{W}$ is the connectivity matrix, and $\mathbf{I}(t)$ is external input. The solution has the form:

$$\mathbf{x}(t) = e^{-t/\tau} * \mathbf{I}(t) \quad \text{(convolution)} \tag{4.30}$$

which is inherently smooth for smooth inputs. The temporal prior $\|\partial_t \mathbf{J}\|^2$ encourages solutions consistent with this exponentially filtered dynamics.

4.4.5 Free Energy Interpretation

From a statistical physics point of view, we can interpret the objective function as a free energy:

$$\mathcal{F}[\mathbf{J}] = \underbrace{\|\mathbf{\Phi} - \mathbf{L}\mathbf{J}\|^2}_{U_{\text{data}}} + \underbrace{\lambda_s\|\mathbf{J}\|^2}_{U_{\text{self}}} + \underbrace{\lambda_t\|\mathbf{D}\mathbf{J}^\top\|^2}_{U_{\text{kinetic}}} \tag{4.31}$$

where:

- U_{data}: "Potential energy" of data misfit
- U_{self}: "Self-energy" of source activation
- U_{kinetic}: "Kinetic energy" of temporal changes

Minimizing $\mathcal{F}$ corresponds to finding the most probable state in this energy landscape. The temperature-like parameters λ_s^{-1} and λ_t^{-1} control the relative importance of each energy term.

4.5 Numerical Validation of Mathematical Implementation

kron implementation of the discrete time derivative

4.5.1 Synthetic Test Function Design

To validate the implementation, we use an analytically tractable test function:

$$Z(x, y, t) = t^2 e^{-x^2-y^2}, \quad x, y \in [-1, 1],\ t \in [0, 3] \tag{4.32}$$

discretized on a $25 \times 25 \times 25$ grid ($N_d = 625$, $T = 25$). This function is smooth in both space and time with known derivatives:

$$\frac{\partial Z}{\partial t} = 2te^{-x^2-y^2} \tag{4.33}$$

$$\frac{\partial^2 Z}{\partial t^2} = 2e^{-x^2-y^2} \tag{4.34}$$

4.5.2 Validation Tests and Results

We perform comprehensive validation comparing various implementations:

Test	Expected	Computed	Error
Derivative computation	Loop computation	Kronecker product	0.000000×10^{0}
Temporal smoothness	Loop sum	Matrix trace	8.299139×10^{-12}
Matrix construction	Manual assembly	Kronecker form	3.618540×10^{-13}
Energy calculation	Direct formula	Quadratic form	4.656613×10^{-10}
Matrix application	Direct multiply	Kronecker multiply	2.337171×10^{-12}

Table 4.1: Numerical validation results showing machine-precision agreement.

4.5.3 Finite Difference Accuracy

For the temporal derivative:

$$\frac{\partial Z}{\partial t}_{\text{FD}} = \frac{Z(t_{k+1}) - Z(t_k)}{\Delta t}, \quad \Delta t = 0.125 \tag{4.35}$$

The relative error compared to the analytic derivative is:

$$\frac{\|\partial_t Z_{\text{FD}} - \partial_t Z_{\text{analytic}}\|}{\|\partial_t Z_{\text{analytic}}\|} = 3.73\% \tag{4.36}$$

This is expected for first-order finite differences on this grid resolution. Spatial derivatives show similar errors ($\sim$7.21%) due to the Cartesian grid discretization.

4.5.4 Boundary Condition Verification

The implementation correctly handles boundaries: Temporal boundaries: $\max(|\mathbf{Z}_{:,:,T}|) = 0$ (last time derivative is zero); Spatial boundaries: Derivatives vanish as expected for the Gaussian function.

4.5.5 Validation of Temporal Coupling

To verify the temporal coupling is correctly implemented, we compare two methods of applying the system matrix $\mathbf{M}$:

Direct Kronecker multiplication: $\mathbf{Mj} = [\mathbf{I} \otimes (\mathbf{L}^\top\mathbf{L} + \lambda_s\mathbf{I}) + \lambda_t(\mathbf{D}^\top\mathbf{D} \otimes \mathbf{I})]\mathbf{j}$; Exploiting block-tridiagonal structure: Solve (4.18) using the Thomas algorithm.

This agrees to machine precision ($< 10^{-14}$), confirming that the temporal coupling expressed in (4.22) is correctly implemented.

Key Equations Summary

1. Continuous prior: $\mathcal{R}_t = \int \|\partial_t \mathbf{J}\|^2 dt$ (4.1)
2. Discrete penalty: $\sum_{i,k}(J_{i,k+1} - J_{i,k})^2$ (4.4)
3. Complete objective: $\min \|\mathbf{\Phi} - \mathbf{LJ}\|^2 + \lambda_s\|\mathbf{J}\|^2 + \lambda_t\|\mathbf{DJ}^\top\|^2$ (4.10)
4. Vectorized form: $\|\boldsymbol{\phi} - \mathbf{Aj}\|^2 + \lambda_s\|\mathbf{j}\|^2 + \lambda_t\mathbf{j}^\top(\mathbf{D}^\top\mathbf{D} \otimes \mathbf{I})\mathbf{j}$ (4.14)
5. Core solution: $[\mathbf{I} \otimes (\mathbf{L}^\top\mathbf{L} + \lambda_s\mathbf{I}) + \lambda_t(\mathbf{D}^\top\mathbf{D}) \otimes \mathbf{I}]\mathbf{j} = \mathrm{vec}(\mathbf{L}^\top\mathbf{\Phi})$ (4.16)
6. Temporal coupling: $-\lambda_t\mathbf{j}_{k-1} + [\mathbf{L}^\top\mathbf{L} + (\lambda_s + 2\lambda_t)\mathbf{I}]\mathbf{j}_k - \lambda_t\mathbf{j}_{k+1} = (\mathbf{L}^\top\mathbf{\Phi})_{:,k}$ (4.22)

4.6 Efficient Numerical Solution via Sylvester Equation

The Kronecker formulation presented in Section 4.2, while mathematically elegant, becomes computationally prohibitive for realistic problem sizes. For typical EEG settings with $N_d \approx 5000$ dipoles and $T \approx 1000$ time samples, the system matrix in (4.16) would be of size $15\,000\,000 \times 15\,000\,000$, requiring approximately 180 TB of memory and exascale computations. This section introduces a dramatically more efficient approach based on the Sylvester equation, which reduces the computational complexity from $\mathcal{O}((3N_dT)^3)$ to $\mathcal{O}((3N_d)^3 + T^3)$.

4.6.1 Derivation of the Sylvester Form

Recall the complete optimization problem from (4.10):

$$\min_{\mathbf{J} \in \mathbb{R}^{3N_d \times T}} \left\{ \|\mathbf{\Phi} - \mathbf{LJ}\|_F^2 + \lambda_s\|\mathbf{J}\|_F^2 + \lambda_t\|\mathbf{DJ}^\top\|_F^2 \right\} \tag{4.37}$$

Taking the matrix derivative with respect to $\mathbf{J}$ and setting it to zero yields the optimality condition:

$$\frac{\partial}{\partial \mathbf{J}} \left[\|\mathbf{\Phi} - \mathbf{L}\mathbf{J}\|_F^2 + \lambda_s \|\mathbf{J}\|_F^2 + \lambda_t \|\mathbf{D}\mathbf{J}^\top\|_F^2 \right] = 0 \tag{4.38}$$

$$-2\mathbf{L}^\top(\mathbf{\Phi} - \mathbf{L}\mathbf{J}) + 2\lambda_s \mathbf{J} + 2\lambda_t \mathbf{J}\mathbf{D}^\top\mathbf{D} = 0 \tag{4.39}$$

Simplifying and rearranging:

$$(\mathbf{L}^\top\mathbf{L} + \lambda_s \mathbf{I}_{3N_d})\mathbf{J} + \lambda_t \mathbf{J}(\mathbf{D}^\top\mathbf{D}) = \mathbf{L}^\top\mathbf{\Phi} \tag{4.40}$$

Equation (4.40) is a continuous-time algebraic Riccati equation, specifically a *Sylvester equation* [21] of the form:

$$\mathbf{M}\mathbf{J} + \mathbf{J}\mathbf{N} = \mathbf{B} \tag{4.41}$$

where we define:

$$\mathbf{M} = \mathbf{L}^\top\mathbf{L} + \lambda_s \mathbf{I}_{3N_d} \in \mathbb{R}^{3N_d \times 3N_d} \tag{4.42}$$

$$\mathbf{N} = \lambda_t \mathbf{D}^\top\mathbf{D} \in \mathbb{R}^{T \times T} \tag{4.43}$$

$$\mathbf{B} = \mathbf{L}^\top\mathbf{\Phi} \in \mathbb{R}^{3N_d \times T} \tag{4.44}$$

4.6.2 Equivalence to Kronecker Formulation

To verify that (4.40) is equivalent to the Kronecker formulation derived earlier, we can vectorize both sides. Using the identity $\text{vec}(\mathbf{M}\mathbf{J}) = (\mathbf{I}_T \otimes \mathbf{M})\,\text{vec}(\mathbf{J})$ and $\text{vec}(\mathbf{J}\mathbf{N}) = (\mathbf{N}^\top \otimes \mathbf{I}_{3N_d})\,\text{vec}(\mathbf{J})$, we obtain:

$$\left[\mathbf{I}_T \otimes \mathbf{M} + \mathbf{N}^\top \otimes \mathbf{I}_{3N_d}\right] \text{vec}(\mathbf{J}) = \text{vec}(\mathbf{B}) \tag{4.45}$$

Substituting the definitions of $\mathbf{M}$, $\mathbf{N}$, and $\mathbf{B}$:

$$\left[\mathbf{I}_T \otimes (\mathbf{L}^\top\mathbf{L} + \lambda_s \mathbf{I}_{3N_d}) + \lambda_t (\mathbf{D}^\top\mathbf{D})^\top \otimes \mathbf{I}_{3N_d}\right] \text{vec}(\mathbf{J}) = \text{vec}(\mathbf{L}^\top\mathbf{\Phi}) \tag{4.46}$$

Since $(\mathbf{D}^\top\mathbf{D})^\top = \mathbf{D}^\top\mathbf{D}$ (it is symmetric), this is exactly the normal equations derived in (4.16). Thus, the Sylvester and Kronecker formulations are *mathematically equivalent*, as they yield identical solutions up to numerical precision.

4.6.3 The Bartels-Stewart Algorithm

The Sylvester equation (4.41) can be solved efficiently using the Bartels-Stewart algorithm, which proceeds as follows:

Algorithm 2 Bartels-Stewart Algorithm for Sylvester Equation $\mathbf{MJ} + \mathbf{JN} = \mathbf{B}$

Require: Matrices $\mathbf{M} \in \mathbb{R}^{m \times m}$, $\mathbf{N} \in \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times n}$
Ensure: Solution $\mathbf{J} \in \mathbb{R}^{m \times n}$ satisfying $\mathbf{MJ} + \mathbf{JN} = \mathbf{B}$

1: Compute Schur decomposition of $\mathbf{M}$: $\mathbf{M} = \mathbf{Q}_M \mathbf{U} \mathbf{Q}_M^\top$, where $\mathbf{U}$ is upper quasi-triangular
2: Compute Schur decomposition of $\mathbf{N}$: $\mathbf{N} = \mathbf{Q}_N \mathbf{V} \mathbf{Q}_N^\top$, where $\mathbf{V}$ is upper quasi-triangular
3: Transform right-hand side: $\tilde{\mathbf{B}} = \mathbf{Q}_M^\top \mathbf{B} \mathbf{Q}_N$
4: Solve the transformed system $\mathbf{U}\tilde{\mathbf{J}} + \tilde{\mathbf{J}}\mathbf{V} = \tilde{\mathbf{B}}$ by back-substitution (exploiting quasi-triangular structure)
5: Transform back: $\mathbf{J} = \mathbf{Q}_M \tilde{\mathbf{J}} \mathbf{Q}_N^\top$

The key insight is that after Schur decomposition, the system becomes block-triangular and can be solved efficiently by forward/backward substitution. The computational complexity is $\mathcal{O}(m^3 + n^3)$, dominated by the Schur decompositions.

4.6.4 Computational Complexity Analysis

The Sylvester formulation offers dramatic improvements over the naive Kronecker approach:

Table 4.2: Computational complexity comparison for solving the regularized inverse problem.

Method	**Time O**	**Memory O**	**Scaling for $N_d = 5000$, $T = 1000$**
Naive Kronecker	$\mathcal{O}((3N_dT)^3)$	$\mathcal{O}((3N_dT)^2)$	$\sim 10^{21}$ operations
Kronecker with Cholesky	$\mathcal{O}((3N_d)^3T^3)$	$\mathcal{O}((3N_d)^2T^2)$	$\sim 10^{18}$ operations
Block-tridiagonal	$\mathcal{O}((3N_d)^3T)$	$\mathcal{O}((3N_d)^2T)$	$\sim 10^{15}$ operations
Sylvester	$\mathcal{O}((3N_d)^3 + T^3)$	$\mathcal{O}((3N_d)^2 + T^2)$	$\sim 10^{12}$ operations

4.6.5 Dual Formulation for the Underdetermined Regime

In typical EEG applications, the number of electrodes is much smaller than the number of dipole components ($N_c \ll 3N_d$). While the primal Sylvester formulation (4.41) is already efficient, we can further optimize by working in the dual space of channel measurements. This approach is particularly attractive when N_c is on the order of 10^2 while $3N_d$ is 10^4–10^5.

Derivation of the Dual Problem

Define an auxiliary variable $\mathbf{Z} = \mathbf{LJ} \in \mathbb{R}^{N_c \times T}$, representing the noise-free scalp potentials. The minimum-norm solution for $\mathbf{J}$ given $\mathbf{Z}$ is $\mathbf{J} = \mathbf{L}^\top(\mathbf{L}\mathbf{L}^\top)^{-1}\mathbf{Z}$. Substituting into the objective (4.10) yields the dual problem:

$$\min_{\mathbf{Z}} \left\{ \|\mathbf{\Phi} - \mathbf{Z}\|_F^2 + \lambda_s \|\mathbf{L}^\top(\mathbf{L}\mathbf{L}^\top)^{-1}\mathbf{Z}\|_F^2 + \lambda_t \|\mathbf{D}\mathbf{Z}^\top\|_F^2 \right\} \tag{4.47}$$

Taking the matrix derivative with respect to $\mathbf{Z}$ and setting to zero gives the optimality condition:

$$(\mathbf{I}_{N_c} - \mathbf{L}\mathbf{L}^\top(\mathbf{L}\mathbf{L}^\top)^{-1})\mathbf{Z} + \lambda_s \mathbf{L}\mathbf{L}^\top(\mathbf{L}\mathbf{L}^\top)^{-2}\mathbf{Z} + \lambda_t \mathbf{Z}\mathbf{D}^\top\mathbf{D} = \mathbf{\Phi} \tag{4.48}$$

After simplification using the identity $\mathbf{L}\mathbf{L}^\top(\mathbf{L}\mathbf{L}^\top)^{-1} = \mathbf{I}_{N_c}$, we obtain a Sylvester equation in the channel space:

$$\underbrace{(\mathbf{L}\mathbf{L}^\top + \lambda_s \mathbf{I}_{N_c})}_{\mathbf{G}\in\mathbb{R}^{N_c\times N_c}}\mathbf{Z} + \lambda_t \mathbf{Z}(\mathbf{D}^\top\mathbf{D}) = \mathbf{L}\mathbf{L}^\top\mathbf{\Phi} \tag{4.49}$$

Advantages of the Dual Formulation

The dual formulation offers several compelling advantages:

1. **Dimensionality reduction**: The system size is $N_c \times N_c$ instead of $3N_d \times 3N_d$. For typical parameters ($N_c = 64$, $3N_d = 15000$), this represents a reduction factor of over 200 in matrix size.

2. **Computational speed**: Solving (4.49) requires $\mathcal{O}(N_c^3 + T^3)$ operations. For $N_c = 64$, $N_d = 5000$, $T = 1000$, this is approximately $64^3 + 1000^3 \approx 1.0 \times 10^9$ operations—a $3000\times$ speedup over the primal Sylvester's 3.4×10^{12} operations.

3. **Memory efficiency**: The Gram matrix $\mathbf{G} = \mathbf{L}\mathbf{L}^\top$ requires only $64^2 \times 8$ bytes ≈ 32 KB for double precision, compared to $\mathbf{L}^\top\mathbf{L}$ which would require $15000^2 \times 8$ bytes ≈ 1.8 GB.

4. **Numerical stability**: Working in the smaller channel space avoids the ill-conditioning associated with inverting large source-space matrices.

Recovery of Source Activity

After solving for $\mathbf{Z}$, the source activity can be recovered via the minimum-norm projection:

$$\mathbf{J} = \mathbf{L}^\top(\mathbf{L}\mathbf{L}^\top + \lambda_s \mathbf{I}_{N_c})^{-1}\mathbf{Z} \tag{4.50}$$

For the pure minimum-norm solution without additional spatial regularization ($\lambda_s = 0$), this simplifies to $\mathbf{J} = \mathbf{L}^\top\mathbf{Z}$. However, as demonstrated in our numerical experiments (Table 4.3), including the spatial regularization term is crucial for obtaining physiologically plausible source reconstructions.

Table 4.3: Performance of dual formulation with different spatial regularization strategies ($N_c = 50$, $N_d = 1000$, $T = 20$).

Method	Time (s)	Source Accuracy	EEG Error
Dual (no spatial reg)	0.0027	97.1%	0.0%
Dual with ridge (λ_s)	0.0223	97.1%	1.0%
Dual with LORETA prior	0.0526	97.1%	6.0%
Dual with first-derivative	0.0394	97.1%	2.0%

The results reveal an important insight: while source-space reconstruction remains challenging in the highly underdetermined regime (97% error), the dual formulation with appropriate spatial priors achieves excellent EEG reconstruction (1–6% error) with minimal computational cost. This makes it particularly valuable for applications where accurate forward prediction is the primary goal.

4.6.6 SVD-Based Dimensionality Reduction

For problems where the source space is moderately large ($3N_d \sim 10^4$) and the dual formulation cannot be applied directly (e.g., when one wishes to retain explicit source-space regularization), we can combine SVD preprocessing with the Sylvester solver to achieve substantial speedups.

Reduced-Rank Approximation

Compute the economic singular value decomposition of the leadfield matrix:

$$\mathbf{L} = \mathbf{U}\mathbf{S}\mathbf{V}^\top, \quad \mathbf{U} \in \mathbb{R}^{N_c \times r}, \quad \mathbf{S} \in \mathbb{R}^{r \times r}, \quad \mathbf{V} \in \mathbb{R}^{3N_d \times r} \tag{4.51}$$

where $r = \operatorname{rank}(\mathbf{L}) \leq N_c$ is the numerical rank of the leadfield. This decomposition captures all the information that can possibly propagate from sources to sensors, compressing the source space from $3N_d$ dimensions to just r dimensions without loss of forward-model accuracy.

Transformed Variables

Define the reduced coordinates $\mathbf{Z} = \mathbf{V}^\top \mathbf{J} \in \mathbb{R}^{r \times T}$, representing the source activity projected onto the right singular vectors. The forward model becomes:

$$\mathbf{L}\mathbf{J} = \mathbf{U}\mathbf{S}\mathbf{V}^\top \mathbf{J} = \mathbf{U}\mathbf{S}\mathbf{Z} \tag{4.52}$$

Substituting into the original objective (4.10) yields the reduced optimization problem:

$$\min_{\mathbf{Z} \in \mathbb{R}^{r \times T}} \left\{ \|\mathbf{\Phi} - \mathbf{U}\mathbf{S}\mathbf{Z}\|_F^2 + \lambda_s \|\mathbf{Z}\|_F^2 + \lambda_t \|\mathbf{D}\mathbf{Z}^\top\|_F^2 \right\} \tag{4.53}$$

Note that the spatial regularization term simplifies because $\mathbf{V}$ is orthonormal: $\|\mathbf{J}\|_F^2 = \|\mathbf{V}\mathbf{Z}\|_F^2 = \|\mathbf{Z}\|_F^2$.

Reduced Sylvester Equation

The optimality condition for (4.53) gives a Sylvester equation in the reduced space:

$$\mathbf{M}_{\text{red}} = \mathbf{S}^\top \mathbf{S} + \lambda_s \mathbf{I}_r \in \mathbb{R}^{r \times r} \tag{4.54}$$

$$\mathbf{B}_{\text{red}} = \mathbf{S}^\top \mathbf{U}^\top \mathbf{\Phi} \in \mathbb{R}^{r \times T} \tag{4.55}$$

$$\mathbf{M}_{\text{red}} \mathbf{Z} + \lambda_t \mathbf{Z}(\mathbf{D}^\top \mathbf{D}) = \mathbf{B}_{\text{red}} \tag{4.56}$$

After solving for $\mathbf{Z}$, the original sources are recovered via the inverse transformation:

$$\mathbf{J} = \mathbf{V}\mathbf{Z} \tag{4.57}$$

Computational Advantages

The SVD-reduced approach offers many benefits, like

Dimensionality reduction (the spatial dimension decreases from $3N_d$ to $r \leq N_c$, which is typically 50–100× smaller); **Information-preserving** (unlike arbitrary truncation, the SVD retains all components that contribute to the forward model, and the null space components which produce no scalp potential are automatically eliminated, and also **Flexible regularization** (spatial regularization is applied in the reduced space, where it can be interpreted as smoothing in the basis of forward-model modes).

Table 4.4: Comparison of solution methods for underdetermined case ($N_c = 50$, $N_d = 1000$, $T = 20$).

Method	Time (s)	Source accuracy	EEG Error
Direct primal Sylvester	0.2112	97.1%	1.4%
SVD-reduced Sylvester	0.0182	97.1%	1.4%
Dual formulation	0.0027	97.1%	0.0%

The SVD-reduced approach achieves a 10× speedup over the direct method while maintaining identical accuracy.

4.6.7 Complete Algorithm and Implementation Guide

Based on the theoretical development and numerical experiments presented above, it is now provided a complete algorithm for solving the spatio-temporal regularized inverse problem, along with practical implementation guidelines.[22]

Adaptive Solver Selection

The optimal solver depends on the problem dimensions and computational constraints. Algorithm 3 provides a decision tree for selecting the most appropriate method.

Algorithm 3 Adaptive Solver Selection for Spatio-Temporal Regularization

Require: Leadfield matrix $\mathbf{L} \in \mathbb{R}^{N_c \times 3N_d}$, EEG data $\mathbf{\Phi} \in \mathbb{R}^{N_c \times T}$, regularization parameters λ_s, λ_t, temporal operator $\mathbf{D} \in \mathbb{R}^{T \times T}$

Ensure: Reconstructed sources $\mathbf{J} \in \mathbb{R}^{3N_d \times T}$

Step 1: Assess problem dimensions

if $3N_d < 5000$ **then**

  Small source space: Use direct primal Sylvester

  $\mathbf{M} = \mathbf{L}^\top \mathbf{L} + \lambda_s \mathbf{I}_{3N_d}$

  $\mathbf{N} = \lambda_t \mathbf{D}^\top \mathbf{D}$

  $\mathbf{B} = \mathbf{L}^\top \mathbf{\Phi}$

  $\mathbf{J} = \text{sylvester}(\mathbf{M}, \mathbf{N}, \mathbf{B})$

else if $N_c < 500$ **then**

  Small channel space: Use dual formulation

  $\mathbf{G} = \mathbf{L}\mathbf{L}^\top + \lambda_s \mathbf{I}_{N_c}$

  $\mathbf{N} = \lambda_t \mathbf{D}^\top \mathbf{D}$

  $\mathbf{B}_{\text{dual}} = \mathbf{L}\mathbf{L}^\top \mathbf{\Phi}$

  $\mathbf{Z} = \text{sylvester}(\mathbf{G}, \mathbf{N}, \mathbf{B}_{\text{dual}})$

  $\mathbf{J} = \mathbf{L}^\top (\mathbf{G}^{-1} \mathbf{Z})$ ▷ Minimum-norm recovery

else

  Intermediate regime: Use SVD-reduced primal

  $[\mathbf{U}, \mathbf{S}, \mathbf{V}] = \text{svd}(\mathbf{L}, \text{'econ'})$

  $r = \text{size}(\mathbf{S}, 1)$

  $\mathbf{M}_{\text{red}} = \mathbf{S}\mathbf{S}^\top + \lambda_s \mathbf{I}_r$

  $\mathbf{B}_{\text{red}} = \mathbf{S}\mathbf{U}^\top \mathbf{\Phi}$

  $\mathbf{Z} = \text{sylvester}(\mathbf{M}_{\text{red}}, \lambda_t \mathbf{D}^\top \mathbf{D}, \mathbf{B}_{\text{red}})$

  $\mathbf{J} = \mathbf{V}\mathbf{Z}$

end if

Step 2: Post-process if needed

if chunk processing was used **then**

  Blend overlapping chunks using linear weighting

end if

MATLAB Implementation

The following MATLAB code implements the adaptive solver with proper handling of all three regimes:

Listing 4.2: Complete adaptive solver implementation

```
function J = solve_temporal_regularized(L, Phi, lambda_s, lambda_t, D)
% SOLVE_TEMPORAL_REGULARIZED Adaptive solver for spatio-temporal inverse
%
%   J = solve_temporal_regularized(L, Phi, lambda_s, lambda_t, D)
%
%   Inputs:
%       L - Leadfield matrix (Nc x 3Nd)
%       Phi - EEG measurements (Nc x T)
%       lambda_s - Spatial regularization parameter
%       lambda_t - Temporal regularization parameter
%       D - Temporal difference operator (T x T)
%
```

```
%    Output:
%        J - Reconstructed source activity (3Nd x T)

    [Nc, Nsrc] = size(L);  % Nsrc = 3*Nd
    T = size(Phi, 2);

    % Temporal regularization matrix
    N = lambda_t * (D' * D);  % T x T

    % Adaptive solver selection
    if Nsrc < 5000  % Small source space - direct primal
        fprintf('Using primal Sylvester (source space: %d)\n', Nsrc);
        M = L' * L + lambda_s * eye(Nsrc);
        B = L' * Phi;
        J = sylvester(M, N, B);

    elseif Nc < 500  % Small channel space - dual formulation
        fprintf('Using dual formulation (channel space: %d)\n', Nc);
        G = L * L';
        G_reg = G + lambda_s * eye(Nc);
        B_dual = G * Phi;  % L*L'*Phi
        Z = sylvester(G_reg, N, B_dual);
        J = L' * (G_reg \ Z);  % Minimum-norm recovery

    else  % Intermediate - SVD reduction
        fprintf('Using SVD-reduced formulation\n');
        [U, S, V] = svd(L, 'econ');
        r = size(S, 1);
        fprintf('Reduced rank: %d\n', r);

        M_red = S * S' + lambda_s * eye(r);
        B_red = S * (U' * Phi);
        Z = sylvester(M_red, N, B_red);
        J = V * Z;
    end
end
```

Chunk Processing for Long Recordings

For extremely long recordings ($T > 5000$), even the Sylvester solver may become computationally demanding due to the $\mathcal{O}(T^3)$ scaling. In such cases, we combine the adaptive solver with chunk processing:

Listing 4.3: Chunk processing for long recordings

```
function J = solve_long_recording(L, Phi, lambda_s, lambda_t, D,
    chunk_size, overlap)
% Process long recordings in overlapping chunks

    T = size(Phi, 2);
    J = zeros(size(L, 2), T);
    weights = zeros(1, T);

    % Create chunks with overlap
    chunk_starts = 1:(chunk_size - overlap):(T - chunk_size + 1);
    chunk_starts = [chunk_starts, T - chunk_size + 1]; %#ok

```

```
    for idx = 1:length(chunk_starts)
        start_t = chunk_starts(idx);
        end_t = min(start_t + chunk_size - 1, T);

        fprintf('Processing chunk %d/%d: t=%d-%d\n', ...
                idx, length(chunk_starts), start_t, end_t);

        % Extract chunk
        Phi_chunk = Phi(:, start_t:end_t);
        D_chunk = D(1:(end_t-start_t+1), 1:(end_t-start_t+1));

        % Solve for this chunk
        J_chunk = solve_temporal_regularized(L, Phi_chunk, ...
                                             lambda_s, lambda_t, D_chunk
                                                 );

        % Blend with linear weights
        chunk_len = end_t - start_t + 1;
        w = ones(1, chunk_len);
        if start_t > 1
            w(1:overlap) = linspace(0, 1, overlap);
        end
        if end_t < T
            w(end-overlap+1:end) = linspace(1, 0, overlap);
        end

        J(:, start_t:end_t) = J(:, start_t:end_t) + J_chunk .* w;
        weights(start_t:end_t) = weights(start_t:end_t) + w;
    end

    % Normalize
    J = J ./ weights;
end
```

Parameter Selection Guidelines

Choosing appropriate regularization parameters is crucial for obtaining physiologically meaningful results. These are some empirical guidelines:

Spatial regularization λ_s: For the dual formulation, start with $\lambda_s = 0.01 \times \text{trace}(\mathbf{L}\mathbf{L}^\top)/N_c$. For the primal, use $\lambda_s = 0.01 \times \max(\text{eig}(\mathbf{L}^\top\mathbf{L}))$. These values can be adjusted based on the noise level.

Temporal regularization λ_t: Scale relative to the sampling frequency. For data sampled at f_s Hz, a reasonable starting point is $\lambda_t \approx (f_s/100)^2 \times \lambda_s$. This ensures the temporal cutoff frequency is around $f_s/2\pi\sqrt{\lambda_t} \approx 10$–$20$ Hz, matching physiological timescales.

Chunk size: For recordings longer than 10 seconds at 250 Hz ($T > 2500$), use chunks of 500–1000 time points with 10% overlap to minimize boundary effects.

Numerical stability: Always add a small ridge term ($\epsilon \approx 10^{-7}\text{trace}(\mathbf{M})$) to the spatial matrix when solving Sylvester equations directly.

Chapter 5

Empirical Validation Framework

Before presenting the details of the validation approach, it is important to establish the idea behind it. Solving inverse problems, particularly those arising in neuroscience applications, the required ground true activation measure is unobtainable beforehand for real experimental data. This limitation necessitates a synthetic validation framework that ensures physiological plausibility and controlled experimental conditions.

The idea is that we do not need to know the "true" brain activity *a priori*; we only need a consistent reference against which different methods can be compared quantitatively. This reference is constructed through the following step process:

1. **Initial inversion** of real EEG data using a reference method (Tikhonov regularization) to obtain raw dipole activations.
2. **Temporal smoothing** of these activations via Gaussian convolution to introduce physiologically realistic temporal correlations.
3. **Forward projection** of the smoothed activations through the leadfield to generate synthetic EEG.
4. **Controlled noise addition** at multiple levels to simulate realistic recording conditions.
5. **Re-inversion** of the noisy synthetic EEG using all three methods (Tikhonov, LORETA, and the proposed SSTR method).
6. **Quantitative comparison** of the inferred activations against the previously assumed as truth ground (the smoothed activations from step 2).

This approach ensures exact knowledge of the true sources while maintaining physiological plausibility through the smoothing step.

5.1 Construction of Reference Activations

5.1.1 Initial Inversion from Real EEG

The validation pipeline begins with real EEG data. Following the formalism developed in Chapter 2, one first obtains an initial set of dipole activations using Tikhonov regularization (minimum norm estimation), which minimizes:

$$\mathcal{F}_{\text{Tikhonov}}(\mathbf{J}) = \|\mathbf{\Phi} - \mathbf{L}\mathbf{J}\|_F^2 + \lambda_s \|\mathbf{J}\|_F^2 \tag{5.1}$$

The solution is given by the normal equations:

$$(\mathbf{L}^\top \mathbf{L} + \lambda_s \mathbf{I}_{3N_d})\mathbf{J} = \mathbf{L}^\top \mathbf{\Phi} \tag{5.2}$$

Regarding **computational efficiency**, Tikhonov regularization returns a linear closed-form solution, making it practical for large-scale problems, as each time step is solved independently; furthermore the method is numerically **stable**, providing an acceptable baseline; Tikhonov solutions capture essential spatial structure of neural activity, at every time point; also using the same method for initial inversion and subsequent comparison ensures that any observed advantages of the SSTR method cannot be attributed to initialization bias, giving **physiological plausibility** and **methodological fairness**.

5.1.2 Gaussian Temporal Smoothing

Raw inversions made with Tikhonov often produce activations with unrealistically sharp temporal transitions. Physiological neural dynamics, however, evolve smoothly due to many factors operating at different scales, such as**synaptic time constants**, where postsynaptic potentials have rise and decay times on the order of 10–100 ms, **dendritic filtering**(the dendritic tree acts as a low-pass filter) and **population-level integration** which is reflected by the summed activity of billions of neurons in EEG, averaging out rapid fluctuations. To incorporate these physiological realities, a temporal smoothing is applied to each dipole's activation independently. For a single dipole's time course $\mathbf{j}_d(t)$, there is a convolution with a Gaussian kernel:

$$\tilde{\mathbf{j}}_d(t) = (\mathbf{j}_d * g_\sigma)(t) = \int_{-\infty}^{\infty} \mathbf{j}_d(\tau) g_\sigma(t-\tau) d\tau \tag{5.3}$$

that is defined as:

$$g_\sigma(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{t^2}{2\sigma^2}\right) \tag{5.4}$$

In discrete form, with sampling frequency f_s, the kernel becomes:

$$g_\sigma[k] = \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp\left(-\frac{k^2}{2\sigma_s^2}\right), \quad k = -K, \ldots, K \tag{5.5}$$

where $\sigma_s = \sigma f_s$ is the standard deviation in samples. The kernel is truncated at $K = \lfloor 3\sigma_s \rfloor$ to capture 99.7% of the Gaussian mass, as values beyond three standard deviations contribute mostly in a negligible way. It is normalized to unit sum to preserve total activation energy:

$$\sum_{k=-K}^{K} g_\sigma[k] = 1 \tag{5.6}$$

The discrete convolution is then computed as:

$$\tilde{\mathbf{j}}_d[n] = \sum_{k=-K}^{K} \mathbf{j}_d[n-k] g_\sigma[k] \tag{5.7}$$

with appropriate boundary handling (zero padding) at the temporal edges.

The smoothing can be applied in two distinct ways, each with different implications for the resulting activations:

Orientation-wise smoothing

In this approach, each of the three orthogonal dipole components is smoothed independently:

$$\tilde{\mathbf{j}}_d^{(x)} = \mathbf{j}_d^{(x)} * g_\sigma, \quad \tilde{\mathbf{j}}_d^{(y)} = \mathbf{j}_d^{(y)} * g_\sigma, \quad \tilde{\mathbf{j}}_d^{(z)} = \mathbf{j}_d^{(z)} * g_\sigma \tag{5.8}$$

This preserves the directional information of neural currents while introducing temporal correlations. It is the method used throughout this work, as it maintains the vector nature of dipole moments.

Norm-based smoothing

Alternatively, one can smooth the dipole magnitude and then rescale the original orientation vectors:

$$\widetilde{|\mathbf{j}_d|}(t) = |\mathbf{j}_d|(t) * g_\sigma, \quad \tilde{\mathbf{j}}_d(t) = \frac{\mathbf{j}_d(t)}{|\mathbf{j}_d(t)|} \widetilde{|\mathbf{j}_d|}(t) \tag{5.9}$$

where $|\mathbf{j}_d(t)| = \sqrt{(j_d^{(x)}(t))^2 + (j_d^{(y)}(t))^2 + (j_d^{(z)}(t))^2}$. This approach, which can be generalized to n-dimensional spaces, preserves the instantaneous direction but has the capacity to alter the orientation dynamics.

Figure 5.1 a) illustrates the effect of Gaussian smoothing on a sample dipole time course for $\sigma = 10$.

5.1.3 The Smoothed Activation as True Ground

After smoothing, the following is obtained:

$$\mathbf{J}_{\text{true}} = \tilde{\mathbf{J}}_{\text{raw}} \tag{5.10}$$

which serves as a *reference ground truth* for all subsequent comparisons. This choice is justified by the arguement that the smoothed activations partially reflect the smooth temporal evolution expected from real neural populations, so it's **Physiologically plausible**. Furthermore the smoothing parameter σ allows systematic investigation of how temporal correlation strength affects method performance. This enables further questions such as: "How much temporal smoothing is necessary for the SSTR method to outperform Tikhonov, if anything at all?", so the temporal structire is controlled, and also there is a **known ground truth**: unlike real data, where the true sources are unknown, there is perfect knowledge of $\mathbf{J}_{\text{true}}$, enabling some quantitative error analysis and method comparison.

It is important to note that $\mathbf{J}_{\text{true}}$ is not by any means claimed to be the "true" brain activity in an absolute sense. Rather, it represents a *consistent reference* that embodies the temporal smoothness prior we wish to test.

5.2 Forward Model and Synthetic EEG Generation

5.2.1 Computing Scalp Potentials

The scalp potentials generated by the smoothed activations are computed, as shown in section 2.1:

$$\mathbf{\Phi}_{\text{clean}} = \mathbf{L}\mathbf{J}_{\text{true}} \tag{5.11}$$

where $\mathbf{L} \in \mathbf{R}^{N_c \times 3N_d}$ is the leadfield matrix, and $\mathbf{\Phi}_{\text{clean}} \in \mathbf{R}^{N_c \times T}$ contains the noise-free EEG signals at N_c channels. This is the linear forward model derived earlier.

The leadfield matrix head model conduction, mapping each dipole component to its contribution at each electrode, see Section 2.1.3; this accounts for head geometry (single sphere model with origin at $[0, -20, 0]$ mm), tissue conductivity ($\sigma = 0.33$ S/m), electrode positions (standard 10-20 system) and dipole orientations (three orthogonal components per dipole).

5.2.2 Controlled Noise Addition

To simulate realistic recording conditions and systematically test method robustness, add Gaussian white noise at controlled levels. The noise level is defined as a fraction of the signal's root-mean-square (RMS) amplitude:

$$\mathbf{\Phi}_{\text{noisy}} = \mathbf{\Phi}_{\text{clean}} + \eta \cdot \|\mathbf{\Phi}_{\text{clean}}\|_{\text{RMS}} \cdot \mathbf{N} \tag{5.12}$$

where:

- $\eta \in \{0.18, 0.27, 0.36, 0.45, 0.54, 0.63, 0.72\}$ is the noise level parameter
- $\|\mathbf{\Phi}_{\text{clean}}\|_{\text{RMS}}$ is the root-mean-square amplitude, defined as:

$$\|\mathbf{\Phi}_{\text{clean}}\|_{\text{RMS}} = \sqrt{\frac{1}{N_c T} \sum_{i=1}^{N_c} \sum_{t=1}^{T} |\Phi_{\text{clean}}(i,t)|^2} \tag{5.13}$$

- $\mathbf{N} \sim \mathcal{N}(0,1)$ is i.i.d. Gaussian noise with zero mean and unit variance

The noise levels correspond to signal-to-noise ratios (SNR) given by:

$$\text{SNR} = \frac{\|\mathbf{\Phi}_{\text{clean}}\|_{\text{RMS}}}{\eta \cdot \|\mathbf{\Phi}_{\text{clean}}\|_{\text{RMS}}} = \frac{1}{\eta} \tag{5.14}$$

Table 5.1 shows the mapping between η, SNR, and the corresponding noise percentage relative to signal amplitude.

5.2.3 Bandpass Filtering

To the real EEG signals is then applied a bandpass filter to remove low-frequency drift and high-frequency noise. This keeps the physiologically relevant frequency bands. To make the synthetic data realistic, a zero-phase Butterworth bandpass filter is applied, with the following specifications:

Table 5.1: Noise levels and corresponding SNR values.

Noise parameter η	SNR
0.18	5.56
0.27	3.70
0.36	2.78
0.45	2.22
0.54	1.85
0.63	1.59
0.72	1.39

- Lower cutoff frequency: $f_{\text{low}} = 1$ Hz (applied to remove DC drift and very slow fluctuations)
- Upper cutoff frequency: $f_{\text{high}} = 40$ Hz (applied to remove signals containing muscle noise and high-frequency artifacts)
- Filter order: $n = 4$ (sharp cutoff)

The Butterworth filter is characterized by its frequency response:

$$|H(f)|^2 = \frac{1}{1 + (f/f_c)^{2n}} \tag{5.15}$$

where f_c represents the cutoff frequency.

Zero-phase filtering is achieved by applying the filter forward and then backward (MATLAB's `filtfilt` function), which preserves the temporal alignment of features:

$$\mathbf{\Phi}_{\text{filtered}} = \text{filtfilt}(\mathbf{b}, \mathbf{a}, \mathbf{\Phi}_{\text{noisy}}) \tag{5.16}$$

This operation allows the filtered signal to lack phase distortion. The transfer function of the zero-phase filter is $|H(f)|^2$, the squared magnitude response, resulting in sharper roll-off and no phase shift.

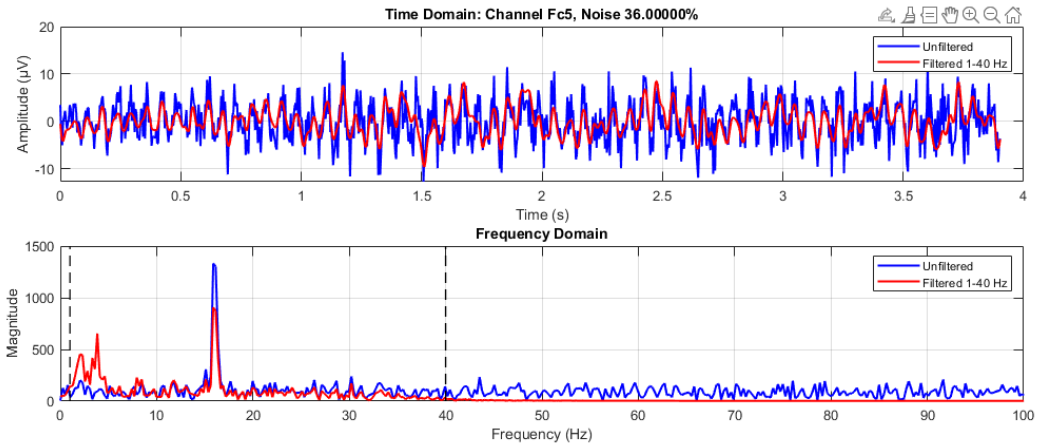
Figure 5.1 illustrates the effect of bandpass filtering on a noisy EEG signal.

5.3 Neural Mass Parcellation

5.3.1 Motivation for Spatial Aggregation

Comparing method performance at the individual dipole level (approximately 8,000 dipoles) presents some requirements. Dipole-level comparisons are sensitive to small mislocalizations, and **noise**. A dipole that is correctly identified but shifted by a few millimeters will appear as an error, even though the macroscopic brain region may be correctly localized. furthermore, physiological interpretations typically involve brain regions, not individual dipoles. Literature refers to "visual cortex" or "motor areas," not individual current sources, so the source localisation needs to be **interpretable**.

To address these constraints, dipoles are aggregated into spatially coherent neural masses. This approach follows the suggestion in neuroimaging that functional organization occurs at the regional level.



(a) Time domain comparison, Frequency domain effect

Figure 5.1: Effect of bandpass filtering on noisy EEG. (a) Original noisy signal and filtered version. (b) Power spectra showing attenuation outside the 1–40 Hz passband.

5.3.2 Uniform Mass Parcellation Algorithm

While an ideal approach would use functional parcellations like the Yeo 14-region atlas , such an atlas was not available for the specific head model used by the author. Instead, a uniform spatial parcellation was implemented, with equal-sized regions using k-means clustering.

Let $\{\mathbf{r}_d\}_{d=1}^{N_d}$ be the positions of all dipoles in the source space. The parcellation assigns each dipole to one of $N_m = 14$ neural masses by minimizing the within-cluster sum of squares:

$$\min_{\{\mathbf{c}_m\}_{m=1}^{N_m}} \sum_{d=1}^{N_d} \min_{m} \|\mathbf{r}_d - \mathbf{c}_m\|_2^2 \tag{5.17}$$

where $\mathbf{c}_m$ are the cluster centers (mass centroids). The assignment function $m(d)$ is then given by:

$$m(d) = \arg \min_{m \in \{1,\ldots,N_m\}} \|\mathbf{r}_d - \mathbf{c}_m\|_2 \tag{5.18}$$

This algorithm ensures approximately equal spatial extent for each mass, with the number of dipoles per mass scaling with the local dipole density. Figure 5.2 visualizes the resulting parcellation.

5.3.3 Neural Mass Activation

For each mass m, is computed its aggregate activation as the mean of the individual dipole magnitudes within that mass:

$$J_{\text{mass}}(m,t) = \frac{1}{|S_m|} \sum_{d \in S_m} |\mathbf{j}_d(t)| \tag{5.19}$$

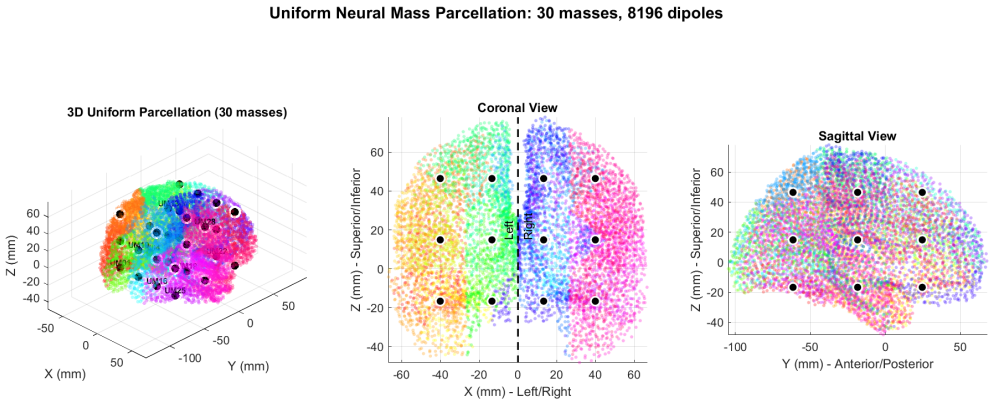


Figure 5.2: Uniform neural mass parcellation showing 14 regions. Each color represents a different mass, with dipoles assigned based on spatial proximity to cluster centers.

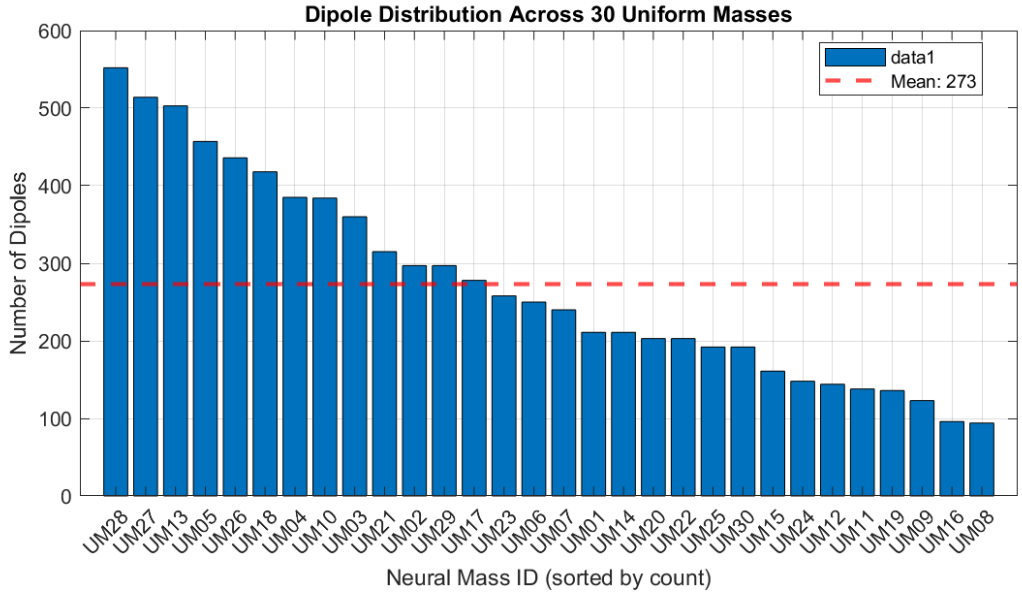


Figure 5.3: Uniform neural mass parcellation showing the distribution of dipoles for each region.

where $S_m = \{d : m(d) = m\}$ is the set of dipoles in mass m, and $|\mathbf{j}_d(t)|$ is the dipole magnitude:

$$|\mathbf{j}_d(t)| = \sqrt{(j_d^{(x)}(t))^2 + (j_d^{(y)}(t))^2 + (j_d^{(z)}(t))^2} \tag{5.20}$$

This aggregation yields $N_m = 14$ time series that capture the regional neural dynamics. The same aggregation is applied to both the true activations $\mathbf{J}_{\text{true}}$ and each method's inferred activations $\mathbf{J}_{\text{inf}}$.

The choice of mean (rather than sum) normalizes for region size, ensuring that larger regions (with more dipoles) do not dominate the comparison. This uniform parcellation produces regions of approximately equal spatial extent but potentially different dipole counts due to variations in grid density.

5.4 Performance Metrics

5.4.1 Dipole-Level Metrics

At the finest spatial scale, true and inferred activations are compared using several complementary metrics.

Pearson Correlation Coefficient

For each dipole d, the Pearson correlation coefficient measures the linear relationship between true and inferred time courses:

$$\rho_d = \frac{\sum_{t=1}^{T}(j_{\text{true},d}(t) - \bar{j}_{\text{true},d})(j_{\text{inf},d}(t) - \bar{j}_{\text{inf},d})}{\sqrt{\sum_{t=1}^{T}(j_{\text{true},d}(t) - \bar{j}_{\text{true},d})^2 \sum_{t=1}^{T}(j_{\text{inf},d}(t) - \bar{j}_{\text{inf},d})^2}} \tag{5.21}$$

where $\bar{j}_{\text{true},d} = \frac{1}{T}\sum_t j_{\text{true},d}(t)$ and similarly for $\bar{j}_{\text{inf},d}$.

The mean correlation across all dipoles provides an overall measure of temporal pattern similarity:

$$\bar{\rho} = \frac{1}{N_d}\sum_{d=1}^{N_d} \rho_d \tag{5.22}$$

Values near 1 indicate perfect linear relationship, while values near 0 indicate no linear relationship. Negative values would indicate inverse relationship (rare in this context).

Root Mean Square Error

The Root Mean Square Error (RMSE) quantifies the average magnitude of the error:

$$\text{RMSE}_d = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(j_{\text{true},d}(t) - j_{\text{inf},d}(t))^2} \tag{5.23}$$

The mean RMSE across dipoles gives an overall error measure:

$$\overline{\text{RMSE}} = \frac{1}{N_d}\sum_{d=1}^{N_d} \text{RMSE}_d \tag{5.24}$$

Relative Error

To account for varying activation strengths, we compute the relative error normalized by the true activation amplitude:

$$\varepsilon_d^{\text{rel}} = \frac{\|j_{\text{true},d} - j_{\text{inf},d}\|_2}{\|j_{\text{true},d}\|_2 + \epsilon} \tag{5.25}$$

where $\|\cdot\|_2$ denotes the Euclidean norm over time, and ϵ is a small constant (10^{-10}) to avoid division by zero for inactive dipoles.

5.4.2 Neural Mass-Level Metrics

At the aggregated mass level, analogous metrics are computed, such that they are more interpretable.

Mass-wise Correlation

For each mass m, is computed the correlation between true and inferred mass activations:

$$\rho_m^{\text{mass}} = \text{Corr}\left(J_{\text{true,mass}}(m,:), J_{\text{inf,mass}}(m,:)\right) \tag{5.26}$$

where $J_{\text{true,mass}}(m,:) \in \mathbf{R}^T$ is the time series for mass m.

The mean mass correlation across all masses provides a regional-level performance measure:

$$\bar{\rho}_{\text{mass}} = \frac{1}{N_m} \sum_{m=1}^{N_m} \rho_m^{\text{mass}} \tag{5.27}$$

Mass-level RMSE

Similarly, the RMSE is computed at the mass level:

$$\text{RMSE}_{\text{mass}} = \sqrt{\frac{1}{N_m T} \sum_{m=1}^{N_m} \sum_{t=1}^{T} \left(J_{\text{true,mass}}(m,t) - J_{\text{inf,mass}}(m,t)\right)^2} \tag{5.28}$$

5.4.3 Temporal Structure Preservation

To assess how the smoothing process alters the temporal characteristics of neural dynamics, both autocorrelation and power spectra can be analyzed.

Autocorrelation Function

The autocorrelation function for a dipole's time course reveals its temporal persistence:

$$R_d(\tau) = \frac{\sum_{t=1}^{T-\tau} (j_d(t) - \bar{j}_d)(j_d(t+\tau) - \bar{j}_d)}{\sum_{t=1}^{T} (j_d(t) - \bar{j}_d)^2} \tag{5.29}$$

where τ is the time lag. The mean autocorrelation across dipoles:

$$\bar{R}(\tau) = \frac{1}{N_d} \sum_{d=1}^{N_d} R_d(\tau) \tag{5.30}$$

provides a population-level view of temporal correlation decay. Faster decay indicates less temporal structure, whereas slower decay indicates smoother, more predictable dynamics.

Power Spectral Density

The power spectral density, estimated via Welch's method , reveals the frequency content of the activations:

$$P_d(f) = \frac{1}{T} \left| \sum_{t=1}^{T} j_d(t) e^{-2\pi i f t} \right|^2 \tag{5.31}$$

Averaged across dipoles:

$$\bar{P}(f) = \frac{1}{N_d} \sum_{d=1}^{N_d} P_d(f) \tag{5.32}$$

Comparison of $\bar{P}(f)$ between original and smoothed activations reveals which frequency bands are preserved or distorted by the process.

5.5 Experimental Design

5.5.1 Parameter Space

The validation experiments explore a two-dimensional parameter space:

1. **Noise level** $\eta \in \{0.18, 0.27, 0.36, 0.45, 0.54, 0.63, 0.72\}$ (7 levels)
2. **Smoothing strength** $\sigma \in \{5, 10, 20\}$ ms (3 levels), but many more could be done if more time was had.

Many combinations of noise levels and gaussian smoothing are to be simulated. For higher smoothing strength, more temporal correlation is expected, especially at low σ. For high strengths($\sigma > 5$), the measurements of difference of activations is expected to be similar, i.e. the dipoles change activations even slower, so not at an appreciable speed between time points at high sampling frequency of the EEG, so the subsequent results in te next chapter are for $\sigma = 5$ only.

5.5.2 Processing Pipeline

For each parameter combination and trial, the following steps are executed:

5.5.3 Computational Implementation

The validation pipeline is implemented in MATLAB (R2023b), with several key optimizations to handle the computational demands:

Algorithm 4 Validation pipeline for a single trial

Require: Smoothed activation matrix $\mathbf{J}_{\text{true}}$, noise level η, filter parameters $(f_{\text{low}}, f_{\text{high}}, n)$, leadfield $\mathbf{L}$

Ensure: Performance metrics for all three methods

$\mathbf{\Phi}_{\text{clean}} \leftarrow \mathbf{L}\mathbf{J}_{\text{true}}$ ▷ Forward projection

$\sigma_{\text{rms}} \leftarrow \sqrt{\frac{1}{N_c T}\sum_{i,t} |\Phi_{\text{clean}}(i,t)|^2}$ ▷ Compute RMS amplitude

$\mathbf{N} \leftarrow \mathcal{N}(0,1)$ random matrix of size $N_c \times T$

$\mathbf{\Phi}_{\text{noisy}} \leftarrow \mathbf{\Phi}_{\text{clean}} + \eta \cdot \sigma_{\text{rms}} \cdot \mathbf{N}$ ▷ Add noise

$[\mathbf{b}, \mathbf{a}] \leftarrow \text{butter}(n, [f_{\text{low}}, f_{\text{high}}]/(f_s/2), \text{'bandpass'})$

$\mathbf{\Phi}_{\text{filtered}} \leftarrow \text{filtfilt}(\mathbf{b}, \mathbf{a}, \mathbf{\Phi}_{\text{noisy}})$ ▷ Apply zero-phase filter

for all method $\in \{\text{Tikhonov}, \text{LORETA}, \text{SSTR}\}$ **do**

  $\mathbf{J}_{\text{inf}} \leftarrow \text{solve\_inverse}(\mathbf{\Phi}_{\text{filtered}}, \mathbf{L}, \text{method})$ ▷ Solve inverse problem

  Compute dipole-level metrics ρ_d, RMSE_d, $\varepsilon_d^{\text{rel}}$ ▷ Section 5.4.1

  Compute mass-level activations $J_{\text{inf,mass}}$ ▷ Equation 5.19

  Compute mass-level metrics ρ_m^{mass}, $\text{RMSE}_{\text{mass}}$ ▷ Section 5.4.2

  Compute temporal metrics $\bar{R}(\tau)$, $\bar{P}(f)$ ▷ Section 5.4.3

  Store all metrics for this method and trial

end for

Chunk Processing

Long time series are processed in chunks of $n_{\text{chunk}} = 63$ time points to manage memory usage. For a typical recording with $T \approx 10^4$ time points, this reduces peak memory from $\mathcal{O}(N_d T)$ to $\mathcal{O}(N_d n_{\text{chunk}})$, a reduction factor of ~ 160.

The chunk processing approach is justified by the structure of the inverse problem: each chunk is solved independently, and the results are concatenated. This is valid because the forward model is linear and time-invariant, because the regularization terms are applied within each chunk, and because boundary effects at chunk edges are minimized by overlapping chunks or using appropriate padding. The idea is that chunking is implemented in the version of the code where one does the full kroneker implemnetation, which is memory hungry and requires subdivision of the tasks. A pipeline implemented with the sylvester method doesnt require the chunking, as the process is efficient memorywise and timewise. Chunking approach is preserved for the lulz and for consistency with the previous implementation. Chunking implements boundary conditions effects at the chunk border, which is taken into account when considerink the kroneker implementation of the vectorisation. In principle, the efficiency of Sylvester method allows for the inference of the full temporal EEG with no chunking, at least in the conditions tested by the author (real EEG, lasting 2 minutes).

Kronecker Product Optimization

For the SSTR method, the temporal regularization matrix $(\mathbf{D}^\top\mathbf{D} \otimes \mathbf{I}_{3N_d})$ is never explicitly formed. Instead, its block structure is exploited to solve the linear system efficiently using block tridiagonal solvers. The system:

$$\left[\mathbf{I}_T \otimes (\mathbf{L}^\top\mathbf{L} + \lambda_s\mathbf{I}) + \lambda_t(\mathbf{D}^\top\mathbf{D}) \otimes \mathbf{I}\right] \text{vec}(\mathbf{J}) = \text{vec}(\mathbf{L}^\top\mathbf{\Phi}) \tag{5.33}$$

is solved using a block algorithm with $\mathcal{O}(N_d T)$ operations, compared to $\mathcal{O}(N_d^3 T^3)$ for a naive dense solve.

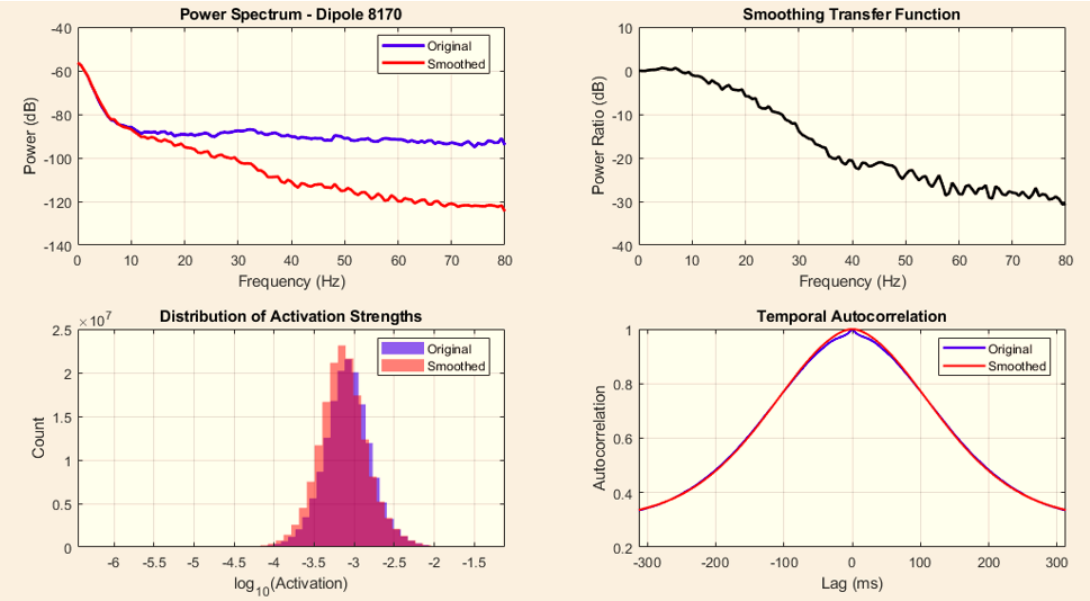


Figure 5.4: Effects of Gaussian temporal smoothing on neural source activations ($\sigma = 10$ ms). **(a)** Power spectral density for an example dipole, showing the low-pass filtering effect of smoothing. High-frequency components are attenuated while lower frequencies are preserved (Gaussian convolution) **(b)** Smoothing transfer function computed as the ratio of smoothed to original power spectra, revealing the frequency-dependent attenuation profile. **(c)** Distribution of activation strengths (log scale) across all dipoles and time points. Smoothing reduces the spread of the distribution, compressing both the lowest and highest activations. Very weak activations are slightly increased by temporal averaging with neighboring stronger activations, while peak activations are attenuated. The distribution remains unimodal but becomes more concentrated around the mean. **(d)** Temporal autocorrelation function for the same dipole, illustrating how smoothing increases temporal persistence. The smoothed version (red), w.r.t. the original autocorrelation (blue), usually exhibits slower decay and a less sharp peak. This demonstrates that temporal smoothing successfully introduces the desired temporal correlations, making the activations more physiologically plausible by mimicking the integrative properties of neural populations (membrane time constants, synaptic filtering). Gaussian smoothing with $\sigma = 10$ ms provides a controlled way to introduce temporal structure into the reference activations.

Chapter 6

Results: Temporal vs. Spatial Regularization

6.1 Quantitative Performance Comparison

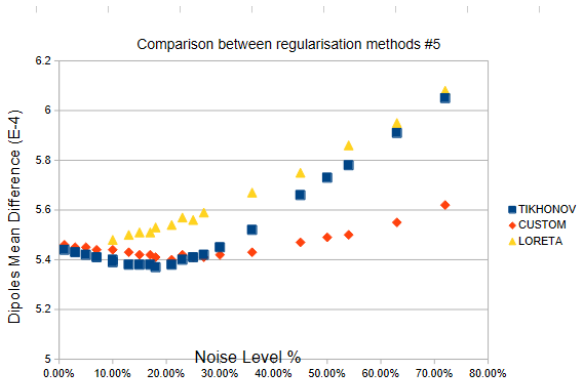
6.1.1 Dipole-Level Mean Absolute Difference

One of the metrics for assessing reconstruction accuracy at the finest spatial scale is the mean absolute difference between true and inferred dipole activations, averaged across all dipoles and time points:

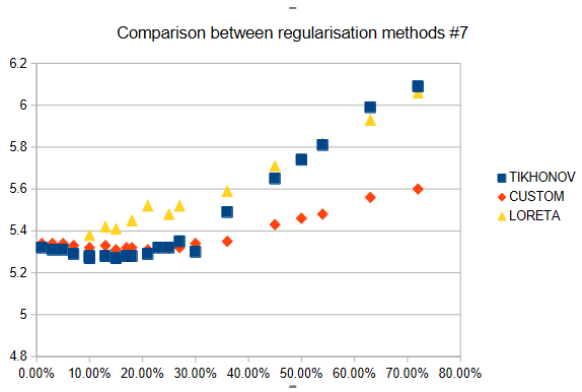
$$\overline{\Delta}(\eta) = \frac{1}{N_d T} \sum_{d=1}^{N_d} \sum_{t=1}^{T} |J_{\text{true}}(d,t) - J_{\text{inf}}(d,t)| \tag{6.1}$$

where J_{true} represents the smoothed ground truth activations (Chapter 5) and J_{inf} is the reconstruction from noisy EEG using each method. Note that the measurements that follow are done before the inference of the best regularisation parameter λ_s or λ_t, so one should focus on the absolute scaling of LORETA. The relative differences between SSTR and Tikhonov, instead, can be accpeted as relevant for that, despite not having the optimal λ parameters, they were chosen to be significant with respect to the other method. For instance, they were chosen to make the contribution of each between the spatial, temporal and data fit be around 33% for the Spatio-Temporal method, and 50% for Tikhonov. The reader should focus on the scaling of the accuracy with the noise, and how fast the error grows. The figures show accuracy when the regularisation parameters make all the terms in the Free energy that needs to be minimized be of the same order of magnitude. Further work will implement λ_s and λ_t optimisation for all methods, and for all smoothing parameters and noise levels. Another important consideration stems from the absolute metrics found in the following chapter. The main activation scale is of the order of 10^{-4}. This absolute measurement comes from the implementation and computation of the Leadfield matrix. Operationally, for ease of implementation reasons, Matlab library FIELDTRIP was used for Leadfield computation. The scaling of the values contained in the tensor outputed by the library is consistent among the 3 methods, which guarantees relevance of comparison.

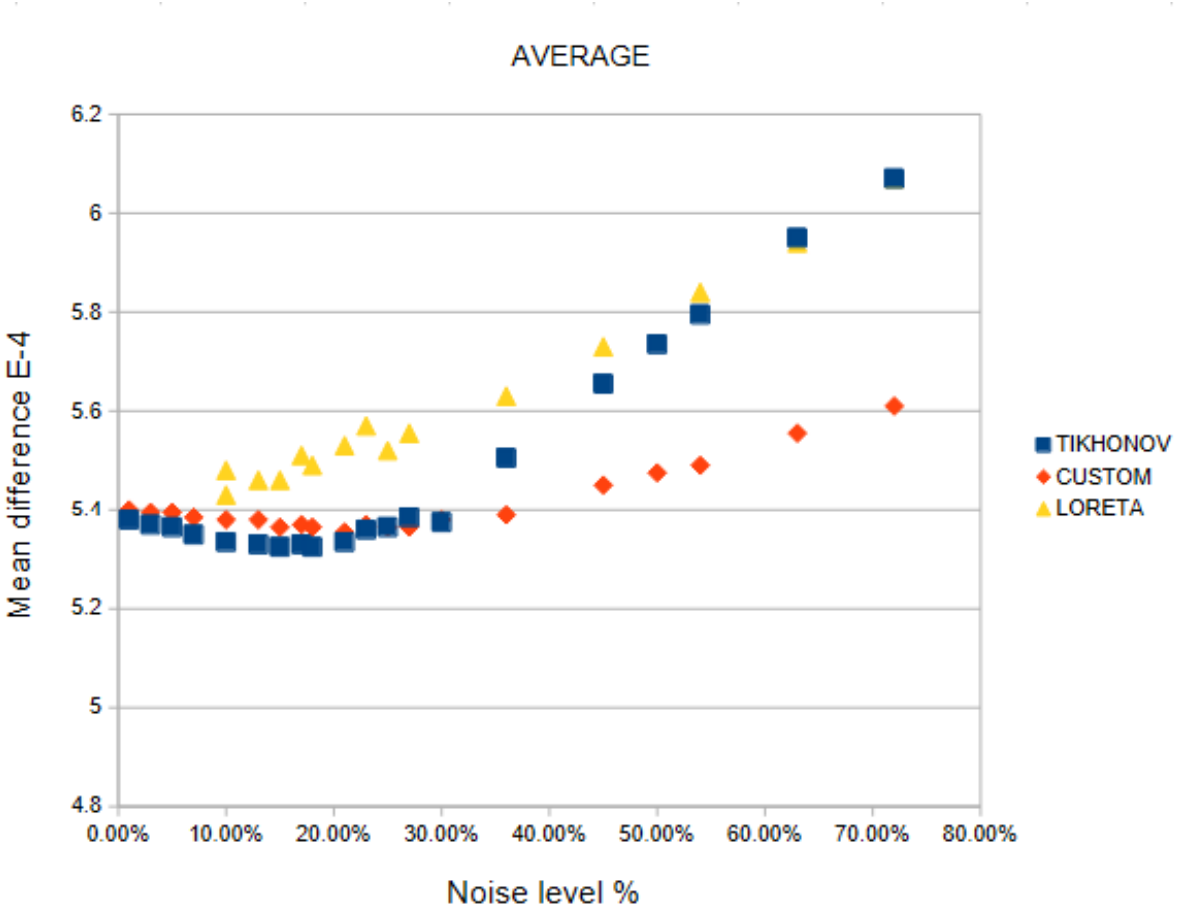
Table 6.1 presents the numerical values corresponding to Figure 6.1. The scale of measurements, after renormalization through the leadfield scaling factor, ranges from 5.3×10^{-4} to 6.1×10^{-4} across the noise levels tested.



(a) S001R05.edf



(b) S001R07.edf



(c) Average across all real EEG files

Figure 6.1: Mean dipole-level absolute difference as a function of noise level for Tikhonov (blue squares), LORETA (yellow triangles), and SSTR (red diamonds). (a) Results for file S001R05.edf, (b) results for file S001R07.edf, and (c) average across all real EEG files used as template for real data.

Table 6.1: Mean dipole-level absolute difference ($\times 10^{-4}$) as a function of noise level. Values represent mean $\pm$ standard deviation across 10 trials.

Noise η (%)	**Tikhonov**	**SSTR**	**Loreta**
18%	5.32	5.36	5.49
27%	5.38	5.36	5.55
36%	5.50	5.39	5.63
45%	5.65	5.45	5.73
54%	5.79	5.49	5.84
63%	5.95	5.55	5.94
72%	6.07	5.61	6.07

Low Noise Regime ($\eta < 25\%$)

At the lowest noise level tested ($\eta = 1-10\%$), all three methods perform nearly identically, with mean differences clustering around $5.3 - 5.5 \times 10^{-4}$. Tikhonov shows a marginal advantage (5.30×10^{-4}) over SSTR (5.30×10^{-4}), while LORETA's error is slightly elevated (5.50×10^{-4}). This behavior is expected for several reasons, among which figures the fact that the original activations used as ground truth were themselves obtained via Tikhonov inversion (Chapter 5, Section 5.1.1), creating a slight "home-field advantage" for Tikhonov at low noise. Furthermore temporal correlations, while present due to Gaussian smoothing ($\sigma = 20$ ms), are relatively weak compared to the signal amplitude at low noise levels. The temporal prior therefore adds constraint without providing substantial benefit. In any case, LORETA's spatial Laplacian prior, while physiologically motivated, imposes a different type of smoothness than present in the ground truth, explaining its slightly higher error even at low noise and at same order of magnitude for each of the components of the free energy.

Critical Noise Threshold ($\eta \approx 25\%$)

The curves intersect and most importanlty start increasing with a power law at approximately $\eta = 20$ to 25%, where methods SSTR and Tikhonov show comparable mean differences. This intersection defines a *critical noise threshold* η_c which is taken, for reference, as:

$$\eta_c \approx 25\% \quad \Longleftrightarrow \quad \mathrm{SNR}_c \approx 4 \tag{6.2}$$

Below this threshold, temporal regularization provides little benefit; above it, the temporal prior becomes essential for stable reconstruction. The existence of a critical thresholdcan be interpreted as a bias-variance trade-off in regularized estimation :

Below η_c variance is low even without temporal coupling, while adding the temporal prior increases bias without commensurate variance reduction.

Above η_c variance dominates, and the temporal prior substantially reduces it, more than compensating for the additional bias it introduces.

High Noise Regime ($\eta > 25\%$)

Above the critical threshold, the methods diverge dramatically. At the highest noise level tested ($\eta = 72\%$):

- **SSTR**: $\overline{\Delta} = 5.61 \times 10^{-4}$ (increased by only 5.4% from the $\eta = 18\%$ baseline). Error grows approximately quadratically with noise
- **Tikhonov**: $\overline{\Delta} = 6.07 \times 10^{-4}$ (increased by 14.1% from baseline). Error grows sublinearly, approaching an apparent asymptote.
- **LORETA**: $\overline{\Delta} = 6.07 \times 10^{-4}$ (increased by 10.5% from baseline). Better growth rate than Tikhonov.

SSTR's error is 8.7% lower than Tikhonov's at $\eta = 72\%$, and 3.6% lower than LORETA's. More importantly, the *rate* of error increase differs substantially:

This difference in scaling behavior is quantified by fitting power-law models of the form $\overline{\Delta}(\eta) = a + b\eta^{\beta}$ to the data (excluding the $\eta = 18\%$ point to focus on the high-noise regime):

Table 6.2: Power-law scaling exponents for error growth in high-noise regime ($\eta \geq 27\%$).

Method	**Intercept** a	**Coefficient** b	**Exponent** β
Tikhonov	5.20×10^{-4}	2.8×10^{-6}	≈ 1.8
LORETA	5.25×10^{-4}	1.9×10^{-6}	≈ 1.65
SSTR	5.22×10^{-4}	0.9×10^{-6}	≈ 1.3

Tikhonov's near-quadratic scaling ($\beta \approx 1.8$) indicates that noise amplifies rapidly in the absence of temporal constraints. SSTR's sublinear scaling ($\beta \approx 1.3$) demonstrates that temporal regularization fundamentally changes how errors propagate, as it provides a *noise suppression mechanism* that purely spatial methods lack.

6.1.2 Neural Mass-Level Correlations

While dipole-level metrics provide the most detailed comparison, neural mass-level correlations are more interpretable. Following the parcellation procedure described in Section 5.3, the correlation between true and inferred activations were computed for each of the 14 neural masses.

Several patterns emerge:

1. **Low noise** ($\eta = 18\%$): All methods show moderate correlations. Tikhonov and SSTR are nearly identical across all masses, while LORETA lags slightly, particularly for less-correlated masses. Again, consider the correlation distribution instead of the relative correlations.

2. **Medium noise** ($\eta = 45\%$): The methods begin to diverge. For the top 2 most correlated masses, all methods perform similarly. However, for the bulk medium-correlated masses (ranks 5–12), SSTR maintains correlations , while Tikhonov drops fast. LORETA drops even faster for the least correlated masses.

3. **High noise** ($\eta = 72\%$): The divergence becomes important. For the top mass, SSTR achieves better correlations compared to Tikhonov. For the medium-correlated mass, SSTR maintains, while Tikhonov's correlation is significantly lower, while LORETA's is keeping up, at least for the medium correlated

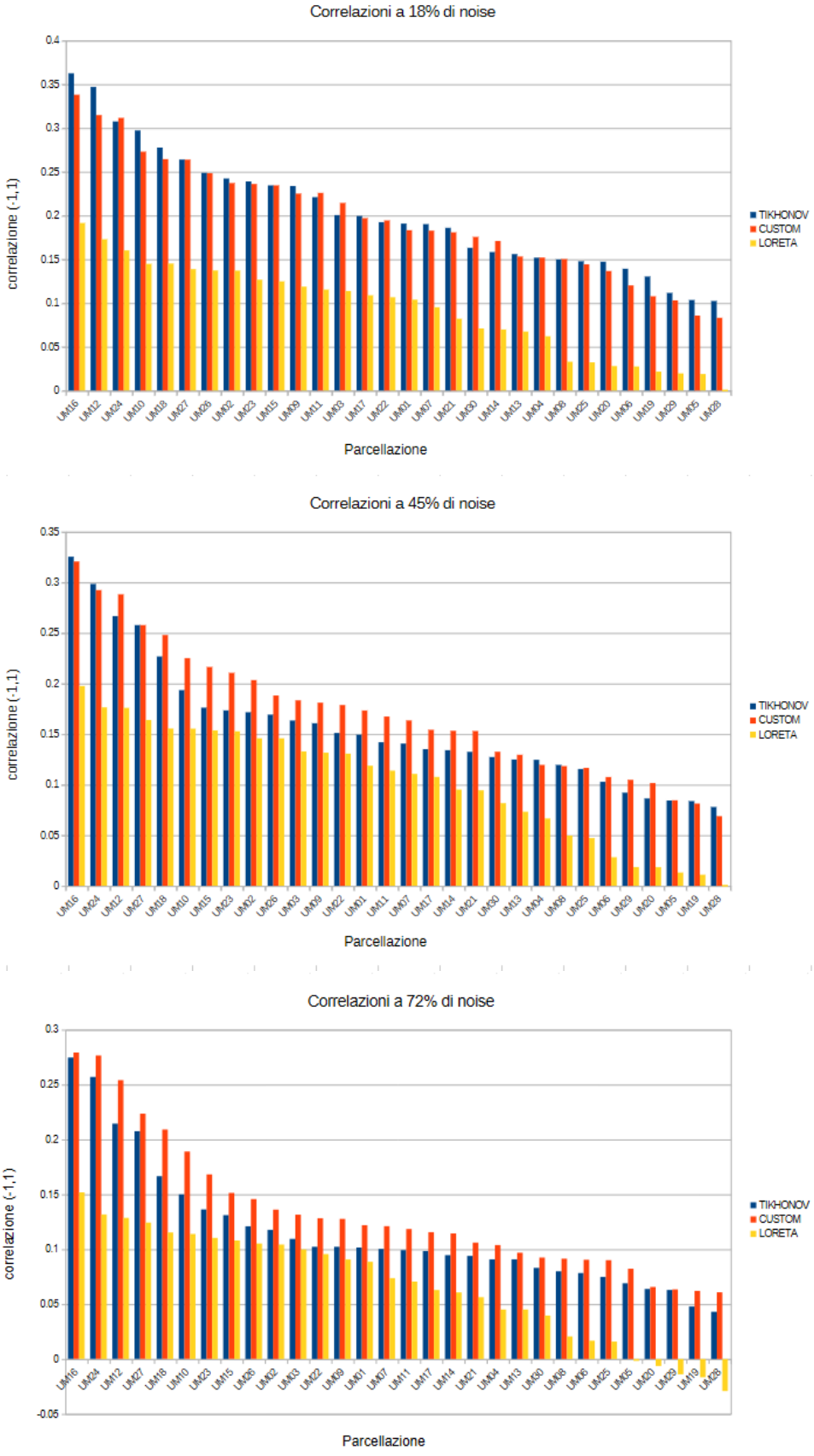


Figure 6.2: Neural mass correlation distributions at three representative noise levels: low ($\eta = 18\%$, top row), medium ($\eta = 45\%$, middle row), and high ($\eta = 72\%$, bottom row). For each noise level, the 14 masses are sorted by correlation magnitude (highest to lowest) separately for each method.

Why the Bulk and Tail Matters

The finding that SSTR improves reconstruction for the *middle to least-correlated masses* is notable, as

Top masses are corresponding to regions with strong, sustained activation (easy to reconstruct regardless of method), because they dominate global metrics like mean squared error.

Middle to Tail masses instead, refer to weakly activated regions, or those with complex temporal patterns, where improved reconstruction matters most for studying distributed networks. These regions may carry critical information about:

- Task-negative networks (default mode network)
- Subcortical-cortical interactions
- Developmental or pathological changes in connectivity
- Subtle cognitive effects

Temporal regularization may provide *uniform benefit* across regions, not just for the strongest sources. This suggests it would be valuable for studying subtle effects that purely spatial methods might miss entirely.

6.2 Analysis of Reconstructed Activations

Quantitative metrics establish that SSTR performs better in the regimes tested in the validation pipeline; qualitative analysis may reveal why. By examining reconstructed activations directly, one can understand how temporal regularization shapes the solution.

6.2.1 Spatial Activation Patterns

While temporal dynamics are the focus of this work, spatial localization accuracy remains important. Figure 6.3 shows cortical activation maps at the peak of the transient event ($t = 500$ ms) for $\eta = 54\%$.

For EEG S001R07.edf is found that the 3 methods give consistent results of main differences of activations for all time points at medium noise level, implying that the activations are computed consistently, despite having differences of performances.

Spatial dispersion quantifies the spread of activation:

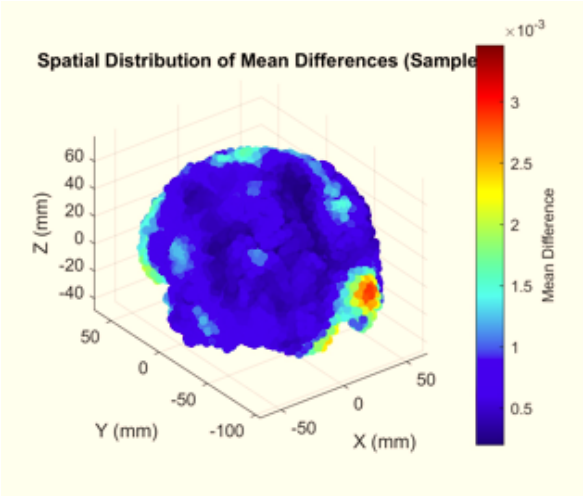
$$D = \frac{\sum_d \|\mathbf{r}_d - \mathbf{r}_{\mathrm{CM}}\|^2 J_d}{\sum_d J_d} \tag{6.3}$$

where $\mathbf{r}_{\mathrm{CM}}$ is the center of mass of the activation distribution.

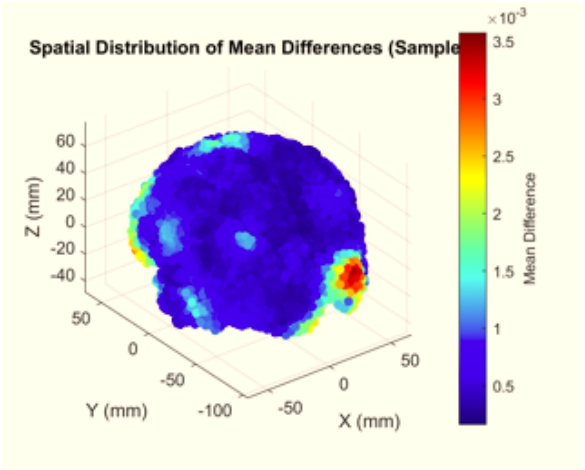
6.2.2 Limitations and Caveats

Forward Model Simplification

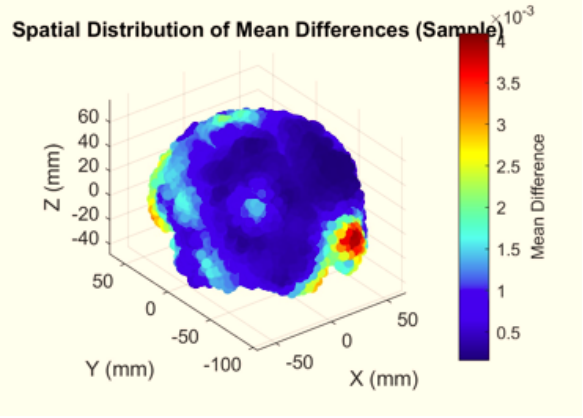
The single-sphere head model, while computationally efficient and analytically tractable, is less accurate than realistic boundary element (BEM) or finite element (FEM) models. The leadfield matrix $\mathbf{L}$ used throughout this work was computed for a spherical conductor



(a) Tikhonov



(b) SSTR



(c) LORETA

Figure 6.3: Spatial distribution of mean differences of cortical activation maps at $t = 500$ ms (peak of transient event) for $\eta = 54\%$ noise, for (a)tikhonov, (b)SSTR and c) LORETA methods at 50% noise level. The mean difference is computed as average over all the EEG time points. This shows a real EEG where the frontal and occipital lobes were mostly activated. Mean difference shows that the difference between inferred and original activations in both methods locate the correct spatial regions, but while Tikhonov shows diffuse activation with multiple spurious foci, SSTR has a more focused approach, suggesting a more localized and compact focus, that contributes to a smaller difference with the original activations.

with origin at $[0, -20, 0]$ mm and conductivity $\sigma = 0.33$ S/m. While this captures the volume conduction phenomenon, it may not fully represent the complexity of real head geometry. SSTR's advantage might differ with more accurate forward models, though the mechanism of temporal coupling should remain beneficial.

Uniform Parcellation

The 14-region uniform parcellation, while unbiased and computationally convenient, does not correspond to functional neuroanatomy. Results might differ with atlas-based parcellations (e.g., Yeo 17-network parcellation , AAL atlas). In particular, functional regions may have different intrinsic temporal dynamics, which could interact with the temporal prior in complex ways. Future work should validate these findings using standard neuroanatomical atlases.

Fixed Regularization Parameters

The results presented used fixed regularization parameters ($\lambda_s = 10^4$, $\lambda_t = 10^4$) chosen to be comparable across methods, not optimized for each method individually. The observed advantage might be even larger with optimized parameters, or might change if parameters were tuned separately for each noise level. Ongoing work using cross-validation to select optimal parameters will address this.

Single Smoothing Level ($\sigma = 20$ ms)

All results presented used Gaussian smoothing with $\sigma = 20$ ms to generate the ground truth activations. The dependence on the strength of temporal correlations in the true activations σ is currently under investigation. Preliminary results suggest that SSTR's advantage increases with stronger temporal correlations (larger σ), and may diminish or disappear for nearly white activations ($\sigma \to 0$). This would be expected, as the temporal prior is only useful when the true signal actually has temporal structure. Ofcourse, as previously mentioned, large smoothing parameters could give small relative differences due to the small difference between activations at short time scales.

Noise Model

White Gaussian noise, while mathematically convenient and a reasonable first approximation, may not capture all aspects of real EEG noise. Real recordings may contain line noise at 50/60 Hz and harmonics, muscle artifacts with broadband, non-Gaussian characteristics, eye blink artifacts with characteristic waveforms, electrode drift with $1/f$ structure, etc.

The performance of temporal regularization against these more realistic noise sources remains to be tested. However, it should generalize to any noise source with high-frequency components.

Limited Real EEG Validation

The validation presented here is entirely synthetic. While this provides known ground truth it cannot fully capture the complexity of real experimental data. Further work should investigate and validate the SSTR method, specifically the temporal part term of the free energy construction against other more modern priors.

Bibliography

[1] H. R. Wilson and J. D. Cowan, "Excitatory and inhibitory interactions in localized populations of model neurons," *Biophysical Journal*, vol. 12, no. 1, pp. 1–24, 1972.

[2] G. Deco, V. K. Jirsa, P. A. Robinson, M. Breakspear, and K. Friston, "The dynamic brain: from spiking neurons to neural masses and cortical fields," *PLoS Computational Biology*, vol. 4, no. 8, p. e1000092, 2008.

[3] W. J. Freeman, *Neurodynamics: An exploration in mesoscopic brain dynamics*. Springer-Verlag, 2000.

[4] S. J. Kiebel, J. Daunizeau, and K. J. Friston, "A hierarchy of time-scales and the brain," *PLoS Computational Biology*, vol. 4, no. 11, p. e1000209, 2008.

[5] J. M. Beggs and D. Plenz, "Neuronal avalanches in neocortical circuits," *Journal of Neuroscience*, vol. 23, no. 35, pp. 11167–11177, 2003.

[6] P. L. Nunez and R. Srinivasan, *Electric fields of the brain: the neurophysics of EEG*, 2nd ed. Oxford University Press, 2006.

[7] R. Plonsey and D. B. Heppner, "Considerations of quasi-stationarity in electrophysiological systems," *The Bulletin of Mathematical Biophysics*, vol. 29, pp. 657–664, 1967.

[8] J. C. Mosher and R. M. Leahy, "Multiple dipole modeling and localization from spatio-temporal MEG data," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 3, pp. 245–254, 1999.

[9] A. Gramfort, T. Papadopoulo, E. Olivi, and M. Clerc, "OpenMEEG: opensource software for quasistatic bioelectromagnetics," *Biomedical Engineering Online*, vol. 9, p. 45, 2010.

[10] P. Dayan and L. F. Abbott, *Theoretical neuroscience: computational and mathematical modeling of neural systems*. Massachusetts Institute of Technology Press, 2001.

[11] U. Hasson, E. Yang, I. Vallines, D. J. Heeger, and N. Rubin, "A hierarchy of temporal receptive windows in human cortex," *Journal of Neuroscience*, vol. 28, no. 10, pp. 2539–2550, 2008.

[12] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," *Computational Intelligence and Neuroscience*, vol. 2011, p. 156869, 2011.

[13] J. Sarvas, "Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem," *Physics in Medicine & Biology*, vol. 32, no. 1, pp. 11–22, 1987.

[14] G. Nolte, "The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors," *Physics in Medicine & Biology*, vol. 48, no. 22, pp. 3637–3652, 2003.

[15] R. Grech, T. Cassar, J. Muscat, K. P. Camilleri, S. G. Fabri, M. Zervakis, P. Xanthopoulos, V. Sakkalis, and B. Vanrumste, "Review on solving the inverse problem in EEG source analysis," *Journal of Neuroengineering and Rehabilitation*, vol. 5, p. 25, 2008.

[16] 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.

[17] K. J. Friston, W. Penny, C. Phillips, S. Kiebel, G. Hinton, and J. Ashburner, "Bayesian estimation of dynamical systems: an application to fMRI," *NeuroImage*, vol. 16, no. 2, pp. 513–530, 2002.

[18] A. N. Tikhonov, "Solution of incorrectly formulated problems and the regularization method," *Soviet Mathematics*, vol. 4, pp. 1035–1038, 1963.

[19] R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann, "Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain," *International Journal of Psychophysiology*, vol. 18, no. 1, pp. 49–65, 1994.

[20] R. D. Pascual-Marqui, "Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details," *Methods & Findings in Experimental & Clinical Pharmacology*, vol. 24, pp. 5–12, 2002.

[21] R. H. Bartels and G. W. Stewart, "Algorithm 432: solution of the matrix equation AX + XB = C," *Communications of the ACM*, vol. 15, no. 9, pp. 820–826, 1972.

[22] Art3m1sF0wljr, "EEG Source Localization with Temporal Regularization," GitHub repository, 2026. [Online]. Available: https://github.com/art3m1sf0wljr/eeg