Université de Rennes I
IMT Atlantique

EEG inverse problem

TP1 - Gibbs sampler

28 february 2024

Students: Franco Martin DI MARIA

franco.di-maria@imt-atlantique.net

1 Prerequisites

We aim to reconstruct the unknown source vector $\mathbf{s}$ given its observation :

$$x = As + n \tag{1}$$

where $\mathbf{A}$ is the known $N \times D$ mixing matrix ($N \ll D$), $\mathbf{x}$ the observation column vector ($N \times 1$), and $\mathbf{n}$ the white Gaussian noise of unknown variances σ_n^2 .

To impose sparsity in the solution, it is further supposed that s follows a Bernoulli-Gaussian (BG) model, an independent, identically distributed (iid) process defined in two stages. Firstly, the sparse nature is governed by the Bernoulli law:

$$P(\mathbf{q}) = \lambda^L (1 - \lambda)^{D-L} \tag{2}$$

with $\mathbf{q} = [q_1, ..., q_D]^t$ a binary sequence and $L = \sum_{d=1}^{D} q_d$, the number of non-zero entries of $\mathbf{q}$. Secondly, amplitudes $\mathbf{s} = [s_1, ..., s_D]^t$ are assumed iid zero-mean Gaussian conditionally to $\mathbf{q}$:

$$s|\mathbf{q} \sim N(0, \sigma_s^2 diag(\mathbf{q})) \tag{3}$$

where $diag(\mathbf{q})$ denotes a diagonal matrix whose diagonal is $\mathbf{q}$.

The problem becomes the estimation of $\Theta = \mathbf{s}, \mathbf{q}, \lambda, \sigma_n^2, \sigma_s^2$ given $\mathbf{x}$, for which the joint posterior distribution writes :

$$P(\Theta|\mathbf{x}) \propto g(\mathbf{x} - A\mathbf{s}; \sigma_s^2 I_N) P(\mathbf{q}; \lambda) g(\mathbf{s}; \sigma_s^2 diag(\mathbf{q})) P(\sigma_n^2) P(\sigma_s^2)(\lambda) \tag{4}$$

where $g(.; R)$ denotes the centered Gaussian density of covariance $\mathbf{R}$. A conjugate prior law is adopted for λ:

$$\lambda \sim Be(?, ?)$$

where Be represents the Beta distribution. For simplicity, fixed values are assumed for $\sigma_n^2 =?$ and $\sigma_s^2 =?$

The hyper-parameters (noted by ?) are to be adjusted according to the problem context.

2 Proposition

According to the Monte Carlo principle, a posterior mean estimator of the unknown random variables can be approximated by:

$$\widehat{\Theta} = \frac{1}{I - J} \sum_{k=J+1}^{I} \Theta^{(k)} \tag{5}$$

where the sum extends over the last I-J samples. In the MCMC framework, the samples are generated iteratively, so that asymptotically converges in distribution to the joint posterior probability in Equation 4.

In order to use the Gibbs algorithm the the following conditional probabilities are needed. They can be deduced from the joint probability in Equation 4.

3 Conditional probabilities

Equation 4 can be developed as follows:

3.1 $q_d|\Theta \setminus \{q_d, s_d\}, x$

$$
\begin{aligned}
P(\Theta|\mathbf{x}) &\propto g(\mathbf{x} - A\mathbf{s}; \sigma_s^2 I_N) P(\mathbf{q}; \lambda) g(\mathbf{s}; \sigma_s^2 diag(\mathbf{q})) P(\sigma_n^2) P(\sigma_s^2) P(\lambda) \\
P(\{\mathbf{s}, \mathbf{q}, \lambda, \sigma_n^2, \sigma_s^2\}|\mathbf{x}) &\propto g(\mathbf{x} - A\mathbf{s}; \sigma_s^2 I_N) P(\mathbf{q}; \lambda) g(\mathbf{s}; \sigma_s^2 diag(\mathbf{q})) P(\sigma_n^2) P(\sigma_s^2) P(\lambda) \\
P(\{\mathbf{s}_{-d}, s_d, \mathbf{q}_{-d}, q_d, \lambda, \sigma_n^2, \sigma_s^2\}|\mathbf{x}) &\propto g(\mathbf{x} - A\mathbf{s}; \sigma_s^2 I_N) P(\mathbf{q}; \lambda) g(\mathbf{s}; \sigma_s^2 diag(\mathbf{q})) P(\sigma_n^2) P(\sigma_s^2) P(\lambda) \\
P(\{s_d, q_d\}|\mathbf{s}_{-d}, \mathbf{q}_{-d}, \lambda, \sigma_n^2, \sigma_s^2, \mathbf{x}) &\propto g(\mathbf{x} - A\mathbf{s}; \sigma_s^2 I_N) P(s_d; \lambda) g(q_d; \sigma_s^2 q_d).1.1.1 \\
P(\{q_d, s_d\}|\Theta \setminus \{q_d, s_d\}, \mathbf{x}) &\propto g(\mathbf{x} - A\mathbf{s}; \sigma_s^2 I_N) P(q_d; \lambda) g(s_d; \sigma_s^2 q_d) \\
&\propto \frac{e^{-\frac{1}{2\sigma_n^2}||x-As||^2}}{(2\pi\sigma_n^2)^{N/2}} \lambda^{q_d} (1-\lambda)^{1-q_d} \frac{e^{-\frac{s_d^2}{2\sigma_s^2 q_d}}}{(2\pi\sigma_s^2 q_d)^{1/2}}
\end{aligned}
$$

3.1.1 Case $q_d = 0$

By Equation 3, we know that if $q_d \to 0$, then s_d tends to Dirac function as follows:

$$
\begin{aligned}
s_d|(q_d \to 0) &\sim N(0,0) \\
s_d|(q_d \to 0) &\sim Dirac(s_d) \\
P(s_d|q_d \to 0) &= \delta(s_d)
\end{aligned}
$$

Therefore, the joint probability distribution of $q_d = 0, s_d$ can be written as follows

$$
P(\{q_d = 0, s_d\}|\Theta \setminus \{q_d, s_d\}, \mathbf{x}) \propto \frac{e^{-\frac{1}{2\sigma_n^2}||x-As||^2}}{(2\pi\sigma_n^2)^{N/2}} \lambda^0 (1-\lambda)^{1-0} \delta(s_d)
$$

$$
P(\{q_d = 0, s_d\}|\Theta \setminus \{q_d, s_d\}, \mathbf{x}) \propto \frac{e^{-\frac{1}{2\sigma_n^2}||x-As||^2}}{(2\pi\sigma_n^2)^{N/2}} (1-\lambda) \delta(s_d) \tag{6}
$$

$$
P(q_d = 0|\Theta \setminus \{q_d, s_d\}, \mathbf{x}) \propto \int \frac{e^{-\frac{1}{2\sigma_n^2}||x-As||^2}}{(2\pi\sigma_n^2)^{N/2}} (1-\lambda) \delta(s_d) ds_d
$$

$$
P(q_d = 0|\Theta \setminus \{q_d, s_d\}, \mathbf{x}) \propto \frac{e^{-\frac{1}{2\sigma_n^2}||x-A_{-d}xs_{-d}||^2}}{(2\pi\sigma_n^2)^{N/2}} (1-\lambda) \tag{7}
$$

where A_{-d} is the A matrix without the d^{th} column and where $\mathbf{s}_{-d}$ is the vector $\mathbf{s}$ without the d^{th} value.

3.1.2 Case $q_d = 1$

$$
\begin{aligned}
P(\{q_d = 1, s_d\}|\Theta \setminus \{q_d, s_d\}, \mathbf{x}) &\propto \frac{e^{-\frac{1}{2\sigma_n^2}||x-As||^2}}{(2\pi\sigma_n^2)^{N/2}} \lambda^1 (1-\lambda)^{1-1} \frac{e^{-\frac{s_d^2}{2\sigma_s^2 1}}}{(2\pi\sigma_s^2 1)^{1/2}} \\
&\propto \frac{e^{-\frac{1}{2\sigma_n^2}||x-As||^2}}{(2\pi\sigma_n^2)^{N/2}} \lambda \frac{e^{-\frac{s_d^2}{2\sigma_s^2}}}{(2\pi\sigma_s^2)^{1/2}} \\
&\propto \frac{e^{-\left(\frac{||x-As||^2}{2\sigma_n^2} + \frac{s_d^2}{2\sigma_s^2}\right)}}{(2\pi\sigma_n^2)^{N/2}(2\pi\sigma_s^2)^{1/2}} \lambda
\end{aligned}
$$

We develop the exponent by completing squares:

$$
\begin{aligned}
\frac{||\mathbf{x}-A\mathbf{s}||^2}{2\sigma_n^2}+\frac{s_d^2}{2\sigma_s^2} &= \frac{||\mathbf{x}-A_{-d}\mathbf{s}_{-d}-A_d s_d||^2}{2\sigma_n^2}+\frac{s_d^2}{2\sigma_s^2} \\
&= \frac{||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2-2(\mathbf{x}-A_{-d}\mathbf{s}_{-d})^t(A_d s_d)+||A_d||^2 s_d^2}{2\sigma_n^2}+\frac{s_d^2}{2\sigma_s^2} \\
&= \frac{||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2}{2\sigma_n^2}-\frac{2(\mathbf{x}-A_{-d}\mathbf{s}_{-d})^t(A_d s_d)}{2\sigma_n^2}+\frac{||A_d||^2 s_d^2}{2\sigma_n^2}+\frac{s_d^2}{2\sigma_s^2} \\
&= \frac{||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2}{2\sigma_n^2}-\frac{(\mathbf{x}-A_{-d}\mathbf{s}_{-d})^t A_d}{\sigma_n^2}s_d+\left(\frac{||A_d||^2}{2\sigma_n^2}+\frac{1}{2\sigma_s^2}\right)s_d^2 \\
&= \frac{||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2}{2\sigma_n^2}-\frac{(\mathbf{x}-A_{-d}\mathbf{s}_{-d})^t A_d}{\sigma_n^2}s_d+\frac{\sigma_s^2||A_d||^2+\sigma_n^2}{\sigma_n^2\sigma_s^2}\frac{s_d^2}{2} \\
&= \frac{\sigma_s^2||A_d||^2+\sigma_n^2}{\sigma_n^2\sigma_s^2}\frac{s_d^2}{2}-\frac{(\mathbf{x}-A_{-d}\mathbf{s}_{-d})^t A_d}{\sigma_n^2}s_d+\frac{||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2}{2\sigma_n^2} \\
&= \frac{1}{2}\frac{\sigma_s^2||A_d||^2+\sigma_n^2}{\sigma_n^2\sigma_s^2}\left(s_d^2-2\frac{\sigma_n^2\sigma_s^2}{\sigma_s^2||A_d||^2+\sigma_n^2}\frac{(\mathbf{x}-A_{-d}\mathbf{s}_{-d})^t A_d}{\sigma_n^2}s_d\right)+\frac{||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2}{2\sigma_n^2} \\
&= \frac{1}{2\left(\frac{\sigma_n^2\sigma_s^2}{\sigma_s^2||A_d||^2+\sigma_n^2}\right)}\left(s_d^2-2\left(\frac{\sigma_n^2\sigma_s^2}{\sigma_s^2||A_d||^2+\sigma_n^2}\right)\left(\frac{(\mathbf{x}-A_{-d}\mathbf{s}_{-d})^t A_d}{\sigma_n^2}\right)s_d\right)+\frac{||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2}{2\sigma_n^2} \\
&= \frac{1}{2\sigma_d^2}\left(s_d^2-2\sigma_d^2\left(\frac{(\mathbf{x}-A_{-d}\mathbf{s}_{-d})^t A_d}{\sigma_n^2}\right)s_d+\left(\sigma_d^2\frac{(\mathbf{x}-A_{-d}\mathbf{s}_{-d})^t A_d}{\sigma_n^2}\right)^2-\left(\sigma_d^2\frac{(\mathbf{x}-A_{-d}\mathbf{s}_{-d})^t A_d}{\sigma_n^2}\right)^2\right) \\
&+\frac{||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2}{2\sigma_n^2} \\
&= \frac{1}{2\sigma_d^2}\left(s_d-\sigma_d^2\frac{(\mathbf{x}-A_{-d}\mathbf{s}_{-d})^t A_d}{\sigma_n^2}\right)^2-\frac{1}{2\sigma_d^2}\left(\sigma_d^2\frac{(\mathbf{x}-A_{-d}\mathbf{s}_{-d})^t A_d}{\sigma_n^2}\right)^2+\frac{||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2}{2\sigma_n^2} \\
&= \frac{1}{2\sigma_d^2}\left(s_d-\frac{\sigma_d^2}{\sigma_n^2}A_d^t(\mathbf{x}-A_{-d}\mathbf{s}_{-d})\right)^2-\frac{1}{2\sigma_d^2}\frac{\sigma_d^4}{\sigma_n^4}A_d^t(\mathbf{x}-A_{-d}\mathbf{s}_{-d})(\mathbf{x}-A_{-d}\mathbf{s}_{-d})^t A_d \\
&+\frac{1}{2\sigma_n^2}||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2 \\
&= \frac{1}{2\sigma_d^2}\left(s_d-\frac{\sigma_d^2}{\sigma_n^2}A_d^t(\mathbf{x}-A_{-d}\mathbf{s}_{-d})\right)^2-\frac{1}{2\sigma_d^2}\frac{\sigma_d^2}{\sigma_n^2}A_d^t(\mathbf{x}-A_{-d}\mathbf{s}_{-d})\frac{\sigma_d^2}{\sigma_n^2}A_d^t(\mathbf{x}-A_{-d}\mathbf{s}_{-d}) \\
&+\frac{1}{2\sigma_n^2}||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2 \\
&= \frac{1}{2\sigma_d^2}\left(s_d-\frac{\sigma_d^2}{\sigma_n^2}A_d^t(\mathbf{x}-A_{-d}\mathbf{s}_{-d})\right)^2-\frac{\mu_d^2}{2\sigma_d^2}+\frac{1}{2\sigma_n^2}||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2 \\
&= \frac{1}{2\sigma_n^2}||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2+\frac{1}{2\sigma_d^2}\left(s_d-\frac{\sigma_d^2}{\sigma_n^2}A_d^t(\mathbf{x}-A_{-d}\mathbf{s}_{-d})\right)^2-\frac{\mu_d^2}{2\sigma_d^2}
\end{aligned}
$$

Then,

$$
P(\{q_d=1,s_d\}|\Theta\setminus\{q_d,s_d\},\mathbf{x})\propto\frac{e^{-\frac{1}{2\sigma_n^2}||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2}}{(2\pi\sigma_n^2)^{N/2}}\frac{e^{-\frac{(s_d-\mu_d)^2}{2\sigma_d^2}}}{(2\pi\sigma_s^2)^{1/2}}e^{\frac{\mu_d^2}{2\sigma_d^2}}\lambda
$$

$$
P(\{q_d=1,s_d\}|\Theta\setminus\{q_d,s_d\},\mathbf{x})\propto\frac{e^{-\frac{1}{2\sigma_n^2}||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2}}{(2\pi\sigma_n^2)^{N/2}}\frac{\sigma_d}{\sigma_s}\frac{e^{-\frac{(s_d-\mu_d)^2}{2\sigma_d^2}}}{(2\pi\sigma_d^2)^{1/2}}e^{\frac{\mu_d^2}{2\sigma_d^2}}\lambda \tag{8}
$$

$$
P(q_d=1|\Theta\setminus\{q_d,s_d\},\mathbf{x})\propto\int\frac{e^{-\frac{1}{2\sigma_n^2}||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2}}{(2\pi\sigma_n^2)^{N/2}}\frac{\sigma_d}{\sigma_s}\frac{e^{-\frac{(s_d-\mu_d)^2}{2\sigma_d^2}}}{(2\pi\sigma_d^2)^{1/2}}e^{\frac{\mu_d^2}{2\sigma_d^2}}\lambda ds_d
$$

$$P(q_d = 1|\Theta \setminus \{q_d, s_d\}, \mathbf{x}) \propto \frac{e^{-\frac{1}{2\sigma_n^2}||\mathbf{x}-A_{-d}\mathbf{s}_{-d}||^2}}{(2\pi\sigma_n^2)^{N/2}} \frac{\sigma_d}{\sigma_s} \lambda e^{\frac{\mu_d^2}{2\sigma_d^2}} \tag{9}$$

with

$$\mu_d = \frac{\sigma_d^2}{\sigma_n^2} A_d^t(\mathbf{x} - A_{-d}\mathbf{s}_{-d}) \qquad \sigma_d^2 = \frac{\sigma_n^2\sigma_s^2}{\sigma_s^2||A_d||^2 + \sigma_n^2} \tag{10}$$

Finally, by Equations 7, 9 and 10, we have that $q_d|\Theta \setminus \{q_d, s_d\}, \mathbf{x}$ follows as Bernoulli distributions as follows:

$$\boxed{(q_d|\Theta \setminus \{q_d, s_d\}, \mathbf{x}) \sim Bi(\frac{\gamma_d}{\gamma_d + 1 - \lambda})} \tag{11}$$

with

$$\boxed{\gamma_d = \lambda \frac{\sigma_d}{\sigma_s} e^{-\frac{\mu_d^2}{2\sigma_d}}} \tag{12}$$

3.2 $s_d|\Theta \setminus \{s_d\}, \mathbf{x}$

By Equations 6 and 8, it is easy to see that s_d follows a normal distribution as follows:

$$\boxed{(s_d|\Theta \setminus \{s_d\}, \mathbf{x}) \sim N(\mu_d, \sigma_d^2) 1_{[q_d \neq 0]}} \tag{13}$$

3.3 $\lambda|\Theta \setminus \{\lambda\}, \mathbf{x}$

Let's suppose that the prior distribution of λ is a Beta distribution with parameters α and β as follows:

$$\lambda \sim Be(\alpha, \beta)$$

Then,

$$\begin{aligned} P(\lambda|\Theta \setminus \{\lambda\}, \mathbf{x}) &\propto g(\mathbf{x} - A\mathbf{s}; \sigma_s^2 I_N) P(\mathbf{q}; \lambda) g(\mathbf{s}; \sigma_s^2 diag(\mathbf{q})).1.1.P(\lambda) \\ &\propto P(\mathbf{q}; \lambda) P(\lambda) \\ &\propto \lambda^{\sum q_d} (1-\lambda)^{D - \sum q_d} \frac{1}{B(\alpha, \beta)} \lambda^{\alpha - 1} (1-\lambda)^{\beta - 1} \\ &\propto \frac{1}{B(\alpha, \beta)} \lambda^{\alpha - 1 + \sum q_d} (1-\lambda)^{\beta - 1 + D - \sum q_d} \end{aligned}$$

$$\boxed{(\lambda|\Theta \setminus \{\lambda\}, \mathbf{x}) \sim Be(\alpha + \sum q_d, \beta + D - \sum q_d)} \tag{14}$$

4 Pseudo-code deduction

Gibbs Sampling Scheme:

1. Current configuration $\Theta^{(k)}$
2. Choose (cyclically or randomly) i
3. Draw $\Theta^{(k+1)} \sim \Theta_i|\Theta^{(k)} \setminus \Theta_i^{(k),z}$
4. Repeat from the beginning

4.1 Current configuration $\Theta^{(k)}$

$e \leftarrow x - As$
$e_i \leftarrow e + A_i S_i$ % A_i is the i-the column of A
$\sigma_i^2 \leftarrow \sigma_n^2 \sigma_s^2 / (\sigma_n^2 + \sigma_s^2 ||A_i||^2)$
$\mu_i \leftarrow (\sigma_i^2/\sigma_s^2) A_i^t e_i$
$\nu_i \leftarrow \lambda(\sigma_i/\sigma_s) exp(\mu_i^2/(2\sigma_i^2))$
$\lambda_i \leftarrow \nu_i/(\nu_i + 1 - \lambda)$

4.2 Choose (cyclically or randomly) i

$e \leftarrow x - As$
for i = 1 to D **do**
  $e_i \leftarrow e + A_i S_i$ % A_i is the i-the column of A
  $\sigma_i^2 \leftarrow \sigma_n^2 \sigma_s^2 / (\sigma_n^2 + \sigma_s^2 ||A_i||^2)$
  $\mu_i \leftarrow (\sigma_i^2/\sigma_s^2) A_i^t e_i$
  $\nu_i \leftarrow \lambda(\sigma_i/\sigma_s) exp(\mu_i^2/(2\sigma_i^2))$
  $\lambda_i \leftarrow \nu_i/(\nu_i + 1 - \lambda)$
  $e \leftarrow e_i - A_i s_i$
end for

4.3 Draw $\Theta^{(k+1)} \sim \Theta_i | \Theta^{(k)} \setminus \Theta_i^{(k),z}$

$e \leftarrow x - As$
for i = 1 to D **do**
  $e_i \leftarrow e + A_i S_i$ % A_i is the i-the column of A
  $\sigma_i^2 \leftarrow \sigma_n^2 \sigma_s^2 / (\sigma_n^2 + \sigma_s^2 ||A_i||^2)$
  $\mu_i \leftarrow (\sigma_i^2/\sigma_s^2) A_i^t e_i$
  $\nu_i \leftarrow \lambda(\sigma_i/\sigma_s) exp(\mu_i^2/(2\sigma_i^2))$
  $\lambda_i \leftarrow \nu_i/(\nu_i + 1 - \lambda)$
  Sample $q_i \sim B(\lambda_i)$
  Sample $s_i \sim N(\mu_i, \sigma_i^2)$ if $q_i = 1$, $s_i = 0$ otherwise
  $e \leftarrow e_i - A_i s_i$
end for
Sample $\lambda \sim Be(\alpha + L, \beta + D - L)$, with $L = \sum q_d$

4.4 Repeat from the beginning

$\mathbf{q} \leftarrow \mathbf{0}$ $\mathbf{s} \leftarrow \mathbf{0}$
% Sample λ using the prior law and choose appropriate values for σ_n^2, σ_s^2
repeat
  % ——————Step 1: Sample (q_i, s_i)——————
  $e \leftarrow x - As$
  for i = 1 to D **do**
    $e_i \leftarrow e + A_i S_i$ % A_i is the i-the column of A
    $\sigma_i^2 \leftarrow \sigma_n^2 \sigma_s^2 / (\sigma_n^2 + \sigma_s^2 ||A_i||^2)$
    $\mu_i \leftarrow (\sigma_i^2/\sigma_s^2) A_i^t e_i$
    $\nu_i \leftarrow \lambda(\sigma_i/\sigma_s) exp(\mu_i^2/(2\sigma_i^2))$
    $\lambda_i \leftarrow \nu_i/(\nu_i + 1 - \lambda)$
    Sample $q_i \sim B(\lambda_i)$
    Sample $s_i \sim N(\mu_i, \sigma_i^2)$ if $q_i = 1$, $s_i = 0$ otherwise
    $e \leftarrow e_i - A_i s_i$
  end for
  % —————— Step 2: Sample λ——————
  Sample $\lambda \sim Be(\alpha + L, \beta + D - L)$, with $L = \sum q_d$
until Convergence

5 Gibbs sampler implementation

Listing 1: Gibbs_sampler.m

```
function [SOut,LambdaOut]=Gibbs_sampler(X, A, sigma_n2, sigma_s2, alpha, beta, Niter)

%get number of sensors and number of dipoles
[~,D]=size(A);

% constants
nA = sum(A.^2,1);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% implement the Gibbs sampler here according to the pseudocode
%
% store vectors q and s of each iteration in matrices Q (D x niter) and S
% (D x niter)
% use variables sigma_n2 and sigma_s2 for the variances of noise and
% signals
Q = zeros(D,Niter);
S = zeros(D,Niter);
lambda = betarnd(alpha, beta);
for n = 1:Niter
    fprintf('Gibbs Iter %d\n',n);
    e = X - A * S(:,n);
    for i = 1:D
        e_i = e + A(:,i)*S(i,n);
        sigma_i2 = (sigma_n2 * sigma_s2) / (sigma_n2 + sigma_s2 * nA(i));
        mu_i = (sigma_i2 / sigma_n2) * A(:,i)' * e_i;
        nu_i = lambda * (sqrt(sigma_i2) / sqrt(sigma_s2)) * exp((mu_i^2) / (2*sigma_i2));
        if nu_i > 1e10
            lambda_i = 1;
        else
            lambda_i = nu_i / (nu_i + 1 - lambda);
        end
        Q(i,n) = binornd(1,lambda_i, 1, 1);
        S(i,n) = Q(i,n) .* normrnd(mu_i, sqrt(sigma_i2));

        e = e_i - A(:,i)*S(i,n);
    end
    L = sum(Q(:,n));
    lambda = betarnd(1 + L, 1 + D - L);
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Make the estimation from Q and S using the MAP criterium
cout = zeros(Niter/2,1);
for j = 1:Niter/2
    fprintf('Est. Iter %d \n',j);
    q = Q(:,Niter/2+j);
    idx = find(q==1);
    R = (A(:,idx)'*A(:,idx))/sigma_n2 + eye(length(idx))/sigma_s2;
    S_opt = (R\(A(:,idx)'*X))/sigma_n2;
    cout(j) = - S_opt'*R*S_opt/sigma_n2;
end

[~, OptIdx] = min(cout);
q = Q(:,Niter/2+OptIdx);
idx = find(q==1);
R = (A(:,idx)'*A(:,idx))/sigma_n2 + eye(length(idx))/sigma_s2;
```

```
SOut = zeros(D,1);
SOut(idx) = (R\(A(:,idx)'*X))/sigma_n2;
LambdaOut = length(idx)/D;
```

Listing 2: TP_inverse_problem_tp1.m

```
%% TP1
clear;
close all;
clc;

load TP_data;

%% Data
%generate linear mixture of source signals
Xs=G*S;

%determine maximum of the signal of interest (here an epileptic spike) to
%apply source loclization algorithms to this time point in the following
[~,id]=max(mean(S,1));

%generate Gaussian random noise
Noise=randn(size(Xs));

%normalize noise
Noise=Noise/norm(Noise,'fro')*norm(Xs,'fro');

%% Original source
%visualize original source distribution
figure; trisurf(mesh.f,mesh.v(:,1),mesh.v(:,2),mesh.v(:,3),S(:,id));
title('Original source','FontSize',18); axis off;

%% Gibbs sampler
%signal to noise ratios
SNRs = [0.1, 1, 10];
for i = 1:length(SNRs)
    %signal to noise ratio
    SNR = SNRs(i);
    fprintf("SNR = %d\n", SNR);
    %generate noisy data according to given SNR
    X=Xs+1/sqrt(SNR)*Noise;
    sigma_n2 = 1/SNR;
    sigma_s2 = 1/SNR;
    alpha    = 1;
    beta     = 1;
    NIter    = 100;
    [Shat,Lambdahat] = Gibbs_sampler(X(:,id),G,sigma_n2, sigma_s2, alpha, beta, NIter);
    %visualize data (for a reduced number of sensors whose indices are
    %specified by idx_electrodes)
    plot_eeg(X(idx_electrodes,:),max(max(X(idx_electrodes,:))),256,channel_names);
    title(sprintf('EEG data (SNR=%d)', SNR),'FontSize',18);
    %visualize reconstructed source distribution
    figure; trisurf(mesh.f,mesh.v(:,1),mesh.v(:,2),mesh.v(:,3),Shat(:));
    title(sprintf('Gibbs sampling (SNR=%d)', SNR),'FontSize',18); axis off;
end
```

6 Gibbs sampler application

In this particular application scenario, our objective is to identify the peak amplitude of the signal of interest, such as an epileptic spike. This determination is crucial for performing source localization.

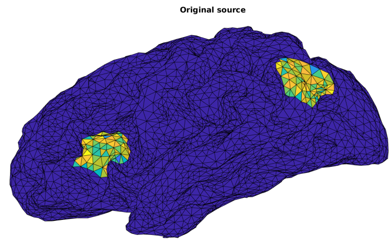

Figure 1: Original source (ground truth)

The diagram depicted in Figure 1 illustrates the simulated ground truth source that we aim to estimate.

Various levels of noise, characterized by different Signal-to-Noise Ratios (SNR), are introduced to the original source. This variation in noise levels serves to evaluate the efficacy of the Gibbs sampling algorithm's performance.

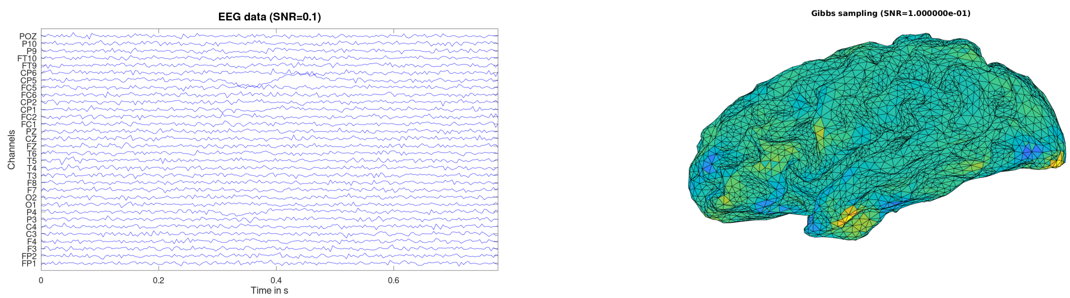

Figure 2: EGG data and Gibbs sampling reconstruction source with SNR = 0.1

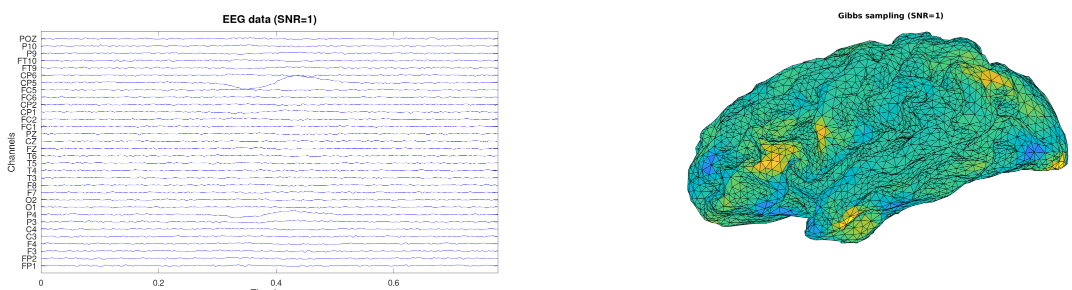

Figure 3: EGG data and Gibbs sampling reconstruction source with SNR = 1

Figure 4: EGG data and Gibbs sampling reconstruction source with SNR = 10

As depicted in Figures 2, 3, and 4, the Gibbs sampling algorithm demonstrates effective source localization capabilities when confronted with low levels of noise (SNR = 10). However, its performance diminishes notably when presented with highly noisy data (SNR = 0.1).

7 Conclusion

In summary, the Gibbs sampling algorithm proves to be proficient in estimating source localization particularly under conditions of low signal-to-noise ratio, where the optimal solution tends to be sparse. Nevertheless, it's important to highlight that this algorithm requires significant computational resources due to its extensive execution time.