Comparison of beamformer implementations for MEG source localization

Amit Jaiswal, Jukka Nenonen, Matti Stenroos, Alexandre Gramfort, Sarang
S. Dalal, Britta U Westner, Vladimir Litvak, John C. Mosher, Jan-Mathijs
Schoffelen, Caroline Witton, et al.

HAL Id: hal-02369296
https://inria.hal.science/hal-02369296

Submitted on 18 Nov 2019

1 Comparison of beamformer implementations for MEG source

2 localization

3 Amit Jaiswal[1,2], Jukka Nenonen[1], Matti Stenroos[2], Alexandre Gramfort[3], Sarang S. Dalal[4], Britta U. Westner[4],

4 Vladimir Litvak[5], John C. Mosher[6], Jan-Mathijs Schoffelen[7], Caroline Witton[9], Robert Oostenveld[7,8], Lauri

5 Parkkonen[1,2]

6 [1]Megin Oy[1], Helsinki, Finland.

7 [2]Department of Neuroscience and Biomedical Engineering, Aalto University School of Science, Espoo, Finland

8 [3]Inria, CEA/Neurospin, Universite Paris-Saclay, Paris, France

9 [4]Center of Functionally Integrative Neuroscience, Aarhus University, Denmark

10 [5]The Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology, London, UK

11 [6]Department of Neurology, University of Texas Health Science Center at Houston, Houston, Texas, USA

12 [7]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

13 [8]NatMEG, Karolinska Institutet, Stockholm, Sweden

14 [9]Aston Brain Centre, School of Life and Health Sciences, Aston University, Birmingham, UK

15 _____

16 Abstract

17 Beamformers are applied for estimating spatiotemporal characteristics of neuronal sources

18 underlying measured MEG/EEG signals. Several MEG analysis toolboxes include an

19 implementation of a linearly constrained minimum-variance (LCMV) beamformer. However,

20 differences in implementations and in their results complicate the selection and application of

21 beamformers and may hinder their wider adoption in research and clinical use. Additionally,

22 combinations of different MEG sensor types (such as magnetometers and planar gradiometers) and

23 application of preprocessing methods for interference suppression, such as signal space separation

24 (SSS), can affect the results in different ways for different implementations. So far, a systematic

25 evaluation of the different implementations has not been performed. Here, we compared the

26 localization performance of the LCMV beamformer pipelines in four widely used open-source

27 toolboxes (FieldTrip, SPM12, Brainstorm, and MNE-Python) using datasets both with and without

28 SSS interference suppression.

29 We analyzed MEG data that were i) simulated, ii) recorded from a static and moving phantom, and

30 iii) recorded from a healthy volunteer receiving auditory, visual, and somatosensory stimulation. We

31 also investigated the effects of SSS and the combination of the magnetometer and gradiometer

32 signals. We quantified how localization error and point-spread volume vary with SNR in all four

33 toolboxes.

---

[1] Former name: Elekta Oy

34 When applied carefully to MEG data with a typical SNR (3–15 dB), all four toolboxes localized the

35 sources reliably; however, they differed in their sensitivity to preprocessing parameters. As expected,

36 localizations were highly unreliable at very low SNR, but we found high localization error also at very

37 high SNRs. We also found that the SNR improvement offered by SSS led to more accurate

38 localization.

39 **Keywords**

40 MEG, EEG, source modeling, beamformers, LCMV, open-source analysis toolbox.

41

42

1. Introduction

43
44 MEG (magnetoencephalography) and EEG (electroencephalography) source imaging aims to
45 identify the spatiotemporal characteristics of neural source currents based on the recorded signals,
46 electromagnetic forward models and physiologically motivated assumptions about the source
47 distribution. One well-known method for estimating a small number of focal sources is to model each
48 of them as a current dipole with fixed location and fixed or changing orientation. The locations
49 (optionally orientations) and time courses of the dipoles are then collectively estimated (Mosher et
50 al., 1992; Hämäläinen et al., 1993). Such equivalent dipole models have been widely applied in basic
51 research (see e.g. Salmelin, 2010) as well as in clinical practice (Bagic et al., 2011a; 2011b; Burgess
52 et al., 2011). Distributed source imaging estimates source currents distribution across the whole
53 source space, typically the cortical surface. Examples of linear methods for distributed source
54 estimation are LORETA (low-resolution brain electromagnetic tomography; Pascual-Marqui et al.,
55 1994) and MNE (minimum-norm estimation; Hämäläinen and Ilmoniemi, 1994). From estimated
56 source distributions, one often computes noise-normalized estimates such as dSPM (dynamic
57 statistical parametric mapping; Dale et al., 2000). Also, various non-linear distributed inverse
58 methods have been proposed (Wipf et al., 2010; Gramfort et al., 2013b).

59 While dipole modeling and distributed source imaging estimate source distributions that reconstruct
60 (the relevant part of) the measurement, beamforming takes an adaptive spatial-filtering approach,
61 scanning independently each location in a predefined region of interest (ROI) within the source space
62 without attempting to reconstruct the data. Beamforming can be done in time-or frequency domain;
63 time-domain methods are typically based on the LCMV approach (Van Veen and Buckley, 1988;
64 1997; Spencer et al., 1992; Sekihara et al., 2006), and in frequency domain the DICS (Dynamic
65 Imaging of Coherent Sources) (Gross et al., 2001) approach is popular.

66 The LCMV beamformer estimates the activity for a source at a given location (typically a point
67 source) while simultaneously suppressing the contributions from all other sources and noise
68 captured in the data covariance matrix. For evaluation of the spatial distribution of the estimated
69 source activity, an image is formed by scanning a set of predefined possible source locations and
70 computing the beamformer output (often power) at each location in the scanning space. When the
71 scanning is done in a volume grid, the beamformer output is typically presented by superimposing it
72 onto an anatomical MRI.

73 Beamformers have been popular in basic MEG research studies (e.g. Hillebrand and Barnes, 2005;
74 Braca et al., 2011; Ishii et al., 2014; van Es and Schoffelen, 2019) as well as in clinical applications
75 such as in localization of epileptic events (e.g. Van Klink et al., 2017; Youssofzadeh et al., 2018; Hall
76 et al., 2018). Many variants of beamformers are implemented in several open-source toolboxes and
77 commercial software for MEG/EEG analysis. Presently, based on citation counts, the most used
78 open-source toolboxes for MEG data analysis are FieldTrip (Oostenveld et al., 2011), Brainstorm

79 (Tadel et al., 2011), MNE-Python (Gramfort et al., 2013a) and DAiSS in SPM12 (Litvak et al., 2011).

80 These four toolboxes have an implementation of an LCMV beamformer, based on the same

81 theoretical framework (van Veen et al., 1997; Sekihara et al., 2006). Yet, it has been anecdotally

82 reported that these toolboxes may yield different results for the same data. These differences may

83 arise not only from the core of the beamformer implementation but also from the previous steps in

84 the analysis pipeline, including data import, preprocessing, forward model computation, combination

85 of data from different sensor types, covariance estimation, and regularization method. Beamforming

86 results obtained from the same toolbox may also differ substantially depending on the applied

87 preprocessing methods; for example, Signal Space Separation (SSS; Taulu and Kajola 2005)

88 reduces the rank of the data, which could affect beamformer output unpredictably if not appropriately

89 considered in the implementation.

90 In this study, we evaluated the LCMV beamformer pipelines in the four open-source toolboxes and

91 investigated the reasons for possible inconsistencies, which hinder the wider adoption of

92 beamformers to research and clinical use where accurate localization of sources is required, e.g., in

93 pre-surgical evaluation. These issues motivated us to study the conditions in which these toolboxes

94 succeed and fail to provide systematic results for the same data and to investigate the underlying

95 reasons.

96

4

2. Materials and Methods

2.1. Datasets

To compare the beamformer implementations, we employed MEG data obtained from simulations, phantom measurements, and measurements of a healthy volunteer who received auditory, visual, and somatosensory stimuli. For all human data recordings, informed consent was obtained from all study subjects in agreement with the approval of the local ethics committee.

2.1.1. MEG systems

All MEG recordings were performed in a magnetically shielded room with a 306-channel MEG system (either Elekta Neuromag® VectorView or TRIUX™; Megin Oy, Helsinki, Finland), which samples the magnetic field distribution by 510 coils at distinct locations above the scalp. The coils are configured into 306 independent channels arranged on 102 triple-sensor elements, each housing a magnetometer and two perpendicular planar gradiometers. The location of the phantom or subject's head relative to the MEG sensor array was determined using four or five head position indicator (HPI) coils attached to the scalp. A Polhemus Fastrak® system (Colchester, VT, USA) was used for digitizing three anatomical landmarks (nasion, left and right preauricular points) to define the head coordinate system. Additionally, the centers of the HPI coils and a set of ~50 additional points defining the scalp were also digitized. The head position in the MEG helmet was determined at the beginning of each measurement using the 'single-shot' HPI procedure, where the coils are activated briefly, and the coil positions are estimated from the measured signals. The location and orientation of the head with respect to the helmet can then be calculated since the coil locations were known both in the head and in the device coordinate systems. After this initial head position measurement, continuous tracking of head movements (cHPI) was engaged by keeping the HPI coils activated to track the movement continuously.

2.1.2. Simulated MEG data

To obtain realistic MEG data with known sources, we superimposed simulated sensor signals based on forward modeling of dipolar sources onto measured spontaneous MEG data utilizing a special in-house simulation software. Structural MRI images, acquired from a healthy adult volunteer using a 3-tesla MRI scanner (Siemens Trio, Erlangen, Germany), were segmented using the MRI Segmentation Software of Megin Oy (Helsinki, Finland) and the surface enveloping the brain compartment was tessellated with triangles (5-mm side length). Using this mesh, a realistic single-shell volume conductor model was constructed using the Boundary Element Method (BEM; Hämäläinen and Sarvas, 1989) implemented in the Source modeling software of Megin Oy. We also segmented the cortical mantle with the FreeSurfer software (Dale et al., 1999; Fischl et al., 1999; Fischl, 2012) for deriving a realistic source space. By using the "ico4" subdivision in MNE-Python,

131   we obtained a source space comprising 2560 dipoles (average spacing 6.2 mm) in each hemisphere

132   (Fig. 1). Out of these, we selected 25 roughly uniformly distributed source locations in the left

133   hemisphere for the simulations (Fig. 1). All these points were at least 7.5 mm inwards from the

134   surface of the volume conductor model. We activated each of the 25 dipoles – one at a time – with

135   a 10-Hz sinusoid of 200-ms duration (2 cycles). The dipoles were simulated at eight source

136   amplitudes: 10, 30, 80, 200, 300, 450, 600 and 800 nAm.

137

138   Insert Fig.1 about here

139

140   A continuous resting-state MEG data with eyes open was recorded from the same volunteer who

141   provided the anatomical data, using an Elekta Neuromag® MEG system (at BioMag Laboratory,

142   Helsinki, Finland). The recording length was 2 minutes, the sampling rate was 1 kHz, and the

143   acquisition frequency band was 0.1–330 Hz. This recording provided the head position for the

144   simulations and defined their noise characteristics. MEG and MRI data were co-registered using the

145   digitized head shape points and the outer skin surface in the segmented MRI.

146   The simulated sensor-level evoked fields were superimposed on the unprocessed resting-state

147   recording with inter-trial-interval varying between 1000–1200 ms resulting in ~110 trials (epochs) in

148   each simulated dataset. The resting-state recording was used both as raw without preprocessing

149   and after SSS interference suppression. Altogether, we obtained 400 simulated MEG datasets (25

150   source locations at 8 dipole amplitudes, all both with the raw and SSS-preprocessed real data). Fig.

151   2 illustrates the generation of simulated MEG data.

152

153   Insert Fig. 2 about here

154   **2.1.3. Phantom data**

155   We used a commercial MEG phantom (Megin Oy, Helsinki, Finland) which contains 32 dipoles and

156   4 HPI coils at distinct fixed locations (see Fig 3a–c and Elekta Neuromag® TRIUX[TM] User's Manual).

157   The phantom is based on the triangle construction (Ilmoniemi et al., 1985): an isosceles triangular

158   line current generates on its relatively very short side a magnetic field distribution equivalent to that

159   of a tangential current dipole in a spherical conductor model, provided that the vertex of the triangle

160   and the origin of the model of a conducting sphere coincide. The phantom data were recorded from

161   8 dipoles, excited one by one (see Elekta Neuromag® TRIUX[TM] User's Manual), using a 306-channel

162   TRIUX[TM] system (at Aston University, Birmingham, UK). The distance from the phantom origin was

163   64 mm for dipoles 5 and 9 (the shallowest), 54 mm for dipoles 6 and 10, 44 mm for dipoles 7 and

164 11, and 34 mm for dipoles 8 and 12 (the deepest; see Fig 3c). The phantom was first kept stationary

165 inside the MEG helmet and continuous MEG data were recorded with 1-kHz sampling rate for three

166 dipole amplitudes (20, 200 and 1000 nAm); one dipole at a time was excited with a 20-Hz sinusoidal

167 current for 500 ms, followed by 500 ms of inactivity. The recordings were repeated with the 200-nAm

168 dipole strength while moving the phantom continuously to mimic head movements inside the MEG

169 helmet; see the movements in Fig. 3e and Suppl. Fig. 2 for all movement parameters.

170

171 <u>Insert Fig. 3 about here</u>

172

173 **2.1.4. Human MEG data**

174 We recorded MEG evoked responses from the same volunteer whose MRI and spontaneous MEG

175 data were utilized in the simulations. These human data were recorded using a 306-channel Elekta

176 Neuromag® system (at BioMag Laboratory, Helsinki, Finland). During the MEG acquisition, the

177 subject was receiving a random sequence of visual (a checkerboard pattern in one of the four

178 quadrants of the visual field), somatosensory (electric stimulation of the median nerve at the left/right

179 wrist at the motor threshold) and auditory (1-kHz 50-ms tone pips to the left/right ear) stimuli with an

180 interstimulus interval of ~500 ms. The Presentation software (Neurobehavioral Systems, Inc.,

181 Albany, CA, USA) was used to produce the stimuli.

182 **2.2. Preprocessing**

183 The datasets were analyzed in two ways: 1) omitting bad channels from the analysis, without

184 applying SSS preprocessing, and 2) applying SSS-based preprocessing methods (SSS/tSSS) to

185 reduce magnetic interference and perform movement compensation for moving phantom data. The

186 SSS-based preprocessing and movement compensation were performed in MaxFilter^{TM} software

187 (version 2.2; Megin Oy, Helsinki, Finland). After that, the continuous data were bandpass filtered

188 (passband indicated for each dataset later in the text) followed by the removing the dc. Then the

189 data were epoched to trials around each stimulus. We applied an automatic trial rejection technique

190 based on the maximum variance across all channels, rejecting trials that had variance higher than

191 the 98th percentile of the maximum or lower than the 2nd percentile (see Suppl. Fig. 4). This method

192 is available as an optional preprocessing step in FieldTrip, and the same implementation was applied

193 in the other toolboxes. Below we describe the detailed preprocessing steps for all datasets.

194 **2.2.1. Simulated data**

195 In each toolbox, the raw data with just bad channels removed or SSS-preprocessed continuous data

196 were filtered using a zero-phase filter with a passband of 2–40 Hz. The filtered data were epoched

197 into windows from –200 to +200 ms relative to the start of the source activity. The bad epochs were
198 removed using the variance-based automatic trial rejection technique, resulting in ~100 epochs.
199 Then the noise and data covariance matrices were estimated from these epochs for the time
200 windows of –200 to –20 ms and 20 to 200 ms, respectively.

201 **2.2.2. Phantom data**

202 All 32 datasets (static: 3 dipole strengths and 8 dipole locations; moving: 1 dipole strength and 8
203 dipole locations) were analyzed both without and with SSS-preprocessing. We applied SSS on static
204 phantom data to remove external interference. On moving-phantom data, combined temporal SSS
205 and movement compensation (tSSS_mc) were applied for suppressing external and movement-
206 related interference and for transforming the data from the continuously estimated positions into a
207 static reference position (Taulu and Kajola 2005; Nenonen et al., 2012). Then in each toolbox the
208 continuous data were filtered to 2–40 Hz using a zero-phase bandpass filter, and the filtered data
209 were epoched from –500 to +500 ms with respect to stimulus triggers. Bad epochs were removed
210 using the automated method based on maximum variance, yielding ~100 epochs for each dataset.
211 The noise and data covariance matrices were estimated in each toolbox for the time windows of –
212 500 to –50 ms and 50 to 500 ms, respectively.

213 **2.2.3. Human MEG data**

214 Both the unprocessed raw data and the data preprocessed with tSSS were filtered to 1–95 Hz using
215 a zero-phase bandpass filter in each toolbox. The trials with somatosensory stimuli (SEF) were
216 epoched between –100 to –10 and 10 to 100 ms for estimating the noise and data covariances,
217 respectively. The corresponding time windows for the auditory-stimulus trials (AEF) were –150 to –
218 20 and 20 to 150 ms, and for the visual stimulus trials (VEF) –200 to –50 and 50 to 200 ms,
219 respectively. Trials contaminated by excessive eye blinks (EOG > 250 µV) or by excessive magnetic
220 signals (MEG > 5000 fT or 3000 fT/cm) were removed with the variance-based automated trial
221 removal technique. Before covariance computation, baseline correction by the time window before
222 the stimulus was applied on each trial. The covariance matrices were estimated independently in
223 each toolbox.

224 Since the actual source locations associated with the evoked fields are not precisely known, we
225 defined reference locations using conventional dipole fitting in the Source Modeling Software of
226 Megin Oy (Helsinki, Finland). A single equivalent dipole was used to represent SEF and VEF
227 sources, and one dipole per hemisphere was used for AEF (see Suppl. Fig. 3). The dipole fitting was
228 performed at the time point of the maximum RMS value across all planar gradiometer channels
229 (global field power) of the average response amplitude.

230 **2.2.4. Forward model**

231 For the beamformer scan of simulated data, we used the default or the most commonly used forward
232 model of each toolbox: a single-compartment BEM model in MNE-Python, a single-shell corrected-
233 sphere model (Nolte, 2003) in FieldTrip, a single-shell corrected sphere model (Nolte, 2003) through
234 inverse normalization of template meshes (Mattout et al., 2007) in SPM12(DAiSS), and the
235 overlapping-spheres (Huang et al., 1999) model in Brainstorm. For constructing models for these
236 forward solutions, the segmentation of MRI images was performed in FreeSurfer for MNE-Python
237 and Brainstorm while FieldTrip and SPM12 used the SPM segmentation procedure. A volumetric
238 source space was represented by a rectangular grid with 5-mm resolution and 5-mm minimal
239 distance from the head model surface. Forward solutions were computed separately in each toolbox
240 using the head model, the volumetric grid sources, and sensor information from the MEG data. Since
241 each toolbox prepares a head model using a different method, the shape of the head models may
242 slightly differ from each other (see Fig. 4) which further may result in a shift between the positions of
243 the scanning grid in these toolboxes.

244

245 <u>Insert Fig. 4 about here</u>

246

247 For phantom data, a homogeneous spherical volume conductor model was defined in each toolbox
248 with the origin at the head coordinate system origin. An equidistant rectangular source-point grid with
249 5-mm resolution was placed inside the upper half of a sphere covering all 32 dipoles of the phantom;
250 see Fig. 3d. Forward solutions for these grids were computed independently in each toolbox. For
251 human MEG data, the head models and the source space were defined in the same way as for the
252 beamformer scanning of the simulated data.

253 **2.3. LCMV beamformer**

254 The linearly constrained minimum-variance (LCMV) beamformer is a spatial filter that relates the
255 magnetic field measured outside the head to the underlying neural activities using the covariance of
256 measured signals and models of source activity and signal transfer between the source and the
257 sensor (Spencer et al., 1992; van Veen et al. 1997; Robinson and Vrba, 1998). The spatial filter
258 weights are computed for each location in the region of interest (ROI).

259 Let **x** be an $M \times 1$ signal vector of MEG data measured with M sensors, and N is the number of grid
260 points in the ROI with grid locations r_j, $(j = 1, ..., N)$. Then the source $\mathbf{y}(r_j)$ at any location r_j can be
261 estimated as weighted combination of the measurement **x** as

262 $$\mathbf{y}(r_j) = \mathbf{W}^{\mathrm{T}}(r_j)\mathbf{x} \tag{1}$$

263 where the $M \times 3$ matrix $\mathbf{W}(r_j)$ is known as spatial filter for a source at location r_j. This type of spatial

264 filter provides a *vector type beamformer* by separately estimating the activity for three orthogonal

265 source orientations, corresponding to the three columns of the matrix. According to Eqs 16–23 in

266 van Veen et al. (1997), the spatial filter $\mathbf{W}(r_j)$ for vector beamformer is defined as

267
$$\mathbf{W}(r_j) = \left(\mathbf{L}^{\mathrm{T}}(r_j)\mathbf{C}^{-1}\mathbf{L}(r_j)\right)^{-1}\mathbf{L}^{\mathrm{T}}(r_j)\mathbf{C}^{-1} \tag{2}$$

268 Here $\mathbf{L}(r_j)$ is the $M \times 3$ local leadfield matrix that defines the contribution of a dipole source at location

269 r_j in the measured data $\mathbf{x}$, and $\mathbf{C}$ is the covariance matrix computed from the measured data samples.

270 To perform source localization using LCMV, the output variance (or output source power) $\mathrm{Var}(\mathbf{y}(r_j))$

271 is estimated at each point in the source space (see Eq (24) in van Veen et al., 1997), resulting in

272
$$\widehat{\mathrm{Var}}(\mathbf{y}(r_j)) = \mathrm{Trace}\{[\mathbf{L}^{\mathrm{T}}(r_j)\mathbf{C}^{-1}\mathbf{L}(r_j)]^{-1}\} \tag{3}$$

273 Usually, the measured signal is contaminated by non-uniformly distributed noise and therefore the

274 estimated signal variance is often normalized with projected noise variance $\mathbf{C}_\mathrm{n}$ calculated over some

275 baseline data (noise). Such normalized estimate is called Neural Activity Index (NAI; van Veen et

276 al., 1997) and can be expressed as

277
$$\mathrm{NAI}(r_j) = \mathrm{Trace}\{[\mathbf{L}^{\mathrm{T}}(r_j)\mathbf{C}^{-1}\mathbf{L}(r_j)]^{-1}\}/\mathrm{Trace}\{[\mathbf{L}^{\mathrm{T}}(r_j)\mathbf{C}_\mathrm{n}^{-1}\mathbf{L}(r_j)]^{-1}\} \tag{4}$$

278 Scanning over all the locations in the region of interest in source space transforms the MEG data

279 from a given measurement into an NAI map.

280 In contrast to a vector beamformer, a *scalar beamformer* (Sekihara and Scholz, 1996; Robinson and

281 Vrba, 1998) uses constant source orientation that is either pre-fixed or optimized from the input data

282 by finding the orientation that maximizes the output source power at each target location. Besides

283 simplifying the output, the optimal-orientation scalar beamformer enhances the output SNR

284 compared to the vector beamformer (Robinson and Vrba, 1998; Sekihara et al., 2004). The optimal

285 orientation $\eta_\mathrm{opt}(r_j)$, for location r_j can be determined by generalized eigenvalue decomposition

286 (Sekihara et al., 2004) using Rayleigh–Ritz formulation as

287
$$\eta_\mathrm{opt}(r_j) = \upsilon_\mathrm{min}\{\mathbf{L}^{\mathrm{T}}(r_j)\mathbf{C}^{-2}\mathbf{L}(r_j), \mathbf{L}^{\mathrm{T}}(r_j)\mathbf{C}^{-1}\mathbf{L}(r_j)\} \tag{5}$$

288 where υ_min indicates the eigenvector corresponding to the smallest generalized eigenvalue of the

289 matrices enclosed in Eq (5) curly braces. For further details, see Eq (4.44) and Section 13.3 in

290 Sekihara and Nagarajan (2008).

291 Denoting $\mathbf{l}_{\eta_\mathrm{opt}}(r_j) = \mathbf{L}(r_j)\eta_\mathrm{opt}(r_j)$ instead of $\mathbf{L}(r_j)$, the weight matrix in Eq (2) becomes $M \times 1$ weight

292 vector $\boldsymbol{w}(r_j)$,

293
$$\boldsymbol{w}(r_j) = \left(\mathbf{l}^{\mathrm{T}}_{\eta_\mathrm{opt}}(r_j)\mathbf{C}^{-1}\mathbf{l}_{\eta_\mathrm{opt}}(r_j)\right)^{-1}\mathbf{l}^{\mathrm{T}}_{\eta_\mathrm{opt}}(r_j)\mathbf{C}^{-1} \tag{6}$$

294 Using $\mathbf{l}_{\eta_{opt}}(r_j)$ in Eq (4), we find the estimate (NAI) of a scalar LCMV beamformer as

295
$$NAI(r_j) = \mathbf{l}_{\eta_{opt}}^T(r_j)\mathbf{C}_n^{-1}\mathbf{l}_{\eta_{opt}}(r_j)/\mathbf{l}_{\eta_{opt}}^T(r_j)\mathbf{C}^{-1}\mathbf{l}_{\eta_{opt}}(r_j) \tag{7}$$

296 When the data covariance matrix is estimated from a sufficiently large number of samples and it has

297 full rank, Eq (7) provides the maximum spatial resolution (Lin et al., 2008; Sekihara and Nagarajan,

298 2008). According to van Veen and colleagues (1997), the number of samples for covariance

299 estimation should be at least three times the number of sensors. Thus, sometimes, the amount of

300 available data may be insufficient to obtain a good estimate of the covariance matrices. In addition,

301 pre-processing methods such as signal-space projection (SSP) or signal-space separation (SSS)

302 reduce the rank of the data, which impacts the matrix inversions in Eq (7). These problems can be

303 mitigated using Tikhonov regularization (Tikhonov, 1963) by replacing matrix $\mathbf{C}^{-1}$ by its regularized

304 version $(\mathbf{C} + \lambda\mathbf{I})^{-1}$ in Eqs (2–7) where λ is called the regularization parameter.

305 All tested toolboxes set the λ with respect to the mean data variance, using ratio 0.05 as default:

306
$$\lambda = 0.05 \times \mathrm{Trace}(\mathbf{C})/M$$

307 If the data are not full rank, also the noise covariance matrix $\mathbf{C}_n$ needs to be regularized.

308 **2.4. Differences between the beamformer pipelines**

309 Though all the four toolboxes evaluated here use the same theoretical framework of the LCMV

310 beamformer, there are several implementation differences which might affect the exact outcome of

311 a beamformer analysis pipeline. Many of these differences pertain to specific handling of the data

312 prior to the estimation of the spatial filters, or to specific ways of (post)processing the beamformer

313 output. Some of the toolbox-specific features reflect the characteristics of the MEG system around

314 which the toolbox has evolved. Importantly, some of these differences are sensitive to input SNR,

315 and they can lead to differences in the results. Table 1 lists the main characteristics and settings of

316 the four toolboxes used in this study. We used the default settings of each toolbox (general practice)

317 for steps before beamforming but set the actual beamforming steps as similar as possible across

318 the toolboxes to be able to meaningfully compare the results.

319 <u>Insert Table 1 about here</u>

320 All toolboxes import data using either Matlab or Python import functions of the MNE software

321 (Gramfort et al., 2014) but represent the data internally either in T or fT (magnetometer) and T/m or

322 fT/mm (gradiometer); see Suppl. Fig. 5. Default filtering approaches across toolboxes change the

323 numeric values, so the linear correlation between the same channels across toolboxes deviates from

324 the identity line; see Suppl. Fig. 6. The default head model is also different across toolboxes; see

325   Section 2.2.4. The single-shell BEM and single-shell corrected sphere model (the "Nolte model") are

326   approximately as accurate but produce slightly different results (Stenroos et al., 2014).

327   For MEG–MRI co-registration, there are several approaches available across these toolboxes such

328   as an interactive method using fiducial or/and digitization points defining the head surface, using

329   automated point cloud registration methods e.g., the iterative closest point (ICP) algorithm. Despite

330   using the same source-space specifications (rectangular grid with 5-mm resolution), differences in

331   head models and/or co-registration methods change the forward model across toolboxes; see Fig. 4.

332   Though there are several approaches to compute data and noise covariances across the four

333   beamformer implementations, by default they all use the empirical/sample covariance. In contrast to

334   other toolboxes, Brainstorm eliminates the cross-modality terms from the data and noise covariance

335   matrices. Also, the regularization parameter λ is calculated and applied separately for gradiometers

336   and magnetometers channel sets in Brainstorm.

337   The combination of two MEG sensor types in the MEGIN triple-sensor array causes additional

338   processing differences in comparison to other MEG systems that employ only axial gradiometers or

339   only magnetometers. Magnetometers and planar gradiometers have different dynamic ranges and

340   measurement units, so their combination must be appropriately addressed in source analysis such

341   as beamforming. For handling the two sensor types in the analysis, different strategies are used for

342   bringing the channels into the same numerical range. MNE-Python and Brainstorm use pre-

343   whitening (Engemann et al., 2015; Ilmoniemi and Sarvas, 2019) based on noise covariance while

344   FieldTrip and SPM12 assume a single sensor type for all the MEG channels. This approach makes

345   SPM12 to favor magnetometer data (with higher numeric values of magnetometer channels) and

346   FieldTrip to favor gradiometer data (with higher numeric values of gradiometer channels). However,

347   users of FieldTrip and SPM12 usually employ only one channel type of the triple-sensor array for

348   beamforming (most commonly, the gradiometers). Due to the presence of two different sensor types

349   in the MEGIN systems and the potential use of SSS methods, the eigenspectra of data from these

350   systems can be idiosyncratic (see Suppl. Fig. 7) and differ from the single-sensor type MEG systems.

351   Rank deficiency and related phenomena are potential sources of beamforming failures with data that

352   have been cleaned with a method such as SSS.

353   Previous studies have shown that the scalar beamformer yields twofold higher output SNR compared

354   to the vector-type beamformer, if the source orientation for the scalar beamformer has been

355   optimized according to Eq 5 (Vrba J., 2000; Sekihara et al., 2004). Most of the beamformer analysis

356   toolboxes have an implementation of optimal-orientation scalar beamformer. In this study, we used

357   the scalar beamformer in MNE-Python, FieldTrip, and SPM12 but a vector-beamformer in Brainstorm

358   since the orientation optimization was not available. To keep the output dimensionality the same

359   across the toolboxes, we linearly summed the three-dimensional NAI values at each source location.

12

360 The general workflow for analysis pipelines across toolboxes used in this study is illustrated in Suppl.
361 Fig. 8.

2.5. Metrics used in comparison

363 In this study, a single focal source could be assumed to underlie the simulated/measured data. In
364 such studies, accurate localization of the source is typically desired. We calculated two metrics for
365 comparing the characteristics of the LCMV beamformer results from the four toolboxes: localization
366 error, and point spread volume. We also analyzed their dependence on input signal-to-noise ratio.

367 **Localization Error (LE):** True source locations were known for the simulated and phantom MEG
368 data and served as reference locations in the comparisons. Since the exact source locations for the
369 human subject MEG data were unknown, we applied the location of a single current dipole as a
370 reference location (see Section 2.1.4 "Human MEG data"). The Source Modelling Software (Megin
371 Oy, Helsinki, Finland) was used to fit a single dipole for each evoked-response category at the time
372 point around the peak of the average response providing the maximum goodness-of-fit value. The
373 beamformer localization error is computed as the Euclidean distance between the estimated and
374 reference source locations.

375 **Point-Spread Volume (PSV):** An ideal spatial filter should provide a unit response at the actual
376 source location and zero response elsewhere. Due to noise and limited spatial selectivity, there is
377 some filter leakage to the nearby locations, which spreads the estimated variance over a volume.
378 The focality of the estimated source, also called focal width, depends on several factors such as the
379 source strength, orientation, and distance from the sensors. PSV measures the focality of an
380 estimate and is defined as the total volume occupied by the source activity above a threshold value;
381 thus, a smaller PSV value indicates a more focal source estimate. We fixed the threshold to 50% of
382 the highest NAI in all comparisons. In this study, the volume represented by a single source in any
383 of the four source spaces (5-mm grid spacing) was 125 mm^3.

384 **Signal-to-Noise ratio (SNR):** Beamformer localization error depends on the input SNR, which varies
385 – among other factors – as a function of source strength and distance of the source from the sensor
386 array. Therefore, we evaluated beamformer localization errors and PSV as a function of the input
387 SNR of the evoked field data.

388 We estimated the SNR for each evoked field MEG dataset in MNE-Python using the estimated noise
389 covariance as follows: The data were whitened using the noise covariance and the effective number
390 of sensors was then calculated as

391 $$N = M - \Sigma(\sigma_n \leq 0) \tag{8}$$

392 where σ_n are the eigenvalues of noise covariance matrix $\mathbf{C}_n$.

393 Then, the input SNR was calculated as:

394
$$SNR_{dB} \ = \ 10 \log_{10} \left(\left[\tfrac{1}{N} \sum_{k=1}^{M} x_k^2(t) \right]_{t_{max}} \right) \tag{9}$$

395 where $x_k(t)$ is the signal on the kth sensor, M is the total number of sensors in the measurement,

396 t_{max} is the time point at maximum amplitude of whitened data across all channels and N is the

397 number of effective sensors defined in Eq (8). Since the same data were used in all toolboxes, we

398 used the same input SNR value for all of them.

399 **2.6. Data and code availability**

400 The codes we wrote to conduct these analyses are publicly available under a repository

401 https://zenodo.org/record/3471758 (DOI: 10.5281/zenodo.3471758). The datasets as well as the

402 specific versions of the four toolboxes used in the study are available at

403 https://zenodo.org/record/3233557 (DOI: 10.5281/zenodo.3233557).

404

14

3. Results

We computed the source localization error (LE) and the point spread volume (PSV) for each NAI estimate across all datasets from LCMV beamformer in all four toolboxes. We plotted the LE and PSV as a function of the input SNR computed according to Eq (9). To differentiate the localization among the implementations, we followed the following color convention: *MNE-Python: grey; FieldTrip: Lavender; SPM12 (DAiSS): Mint; and Brainstorm: coral*.

3.1. Simulated MEG data

Localization errors and PSV values were calculated for all simulated datasets and plotted against the corresponding input SNR. The SNR of all 200 simulated datasets ranged between 0.5 to 25 dB. Fig. 5a shows the plots between localization error and input SNR of each simulated dataset. The polynomial regressions of the maximum localization errors across LCMV implementations show the variation of localization errors over the range of input SNRs. The localization error goes high for all toolboxes for very low SNR (< 3 dB) signals (e.g. 20-nAm or deep sources). The localization error within the input SNR range 3–12 dB is stable and mostly within 15 mm, and SSS preprocessing widens this SNR range of stable performance to 3–15 dB. Unexpectedly, we also found high localization error at high SNR (> 15 dB) for the toolboxes other than SPM12 (DAiSS). Fig. 5b plots PSV values against input SNR for raw and SSS-preprocessed simulated data. The polynomial regression plots fit a nonlinear relationship between the input SNR and the corresponding maximal PSVs across the four LCMV implementations. The regression plots in Fig. 5b agree with the corresponding plots in Fig. 5a, i.e., lower PSV values (higher spatial resolution) for the SNR range with smaller localization errors and vice-versa, for all toolboxes. The low SNR signals (usually, weak or deep sources) shows high PSV values in Fig. 5b which also indicates improved spatial resolution after SSS preprocessing. Fig. 5a–b shows that none of the four toolboxes provides accurate localization for all SNR values and that the spatial resolution of LCMV is dependent on input SNR. SPM12 (DAiSS) shows lower localization errors and PSV values at very high SNR too.

Insert Fig. 5 about here

3.2. Static and moving phantom MEG data

In the case of phantom data, the background noise is very low and there is a single source underneath a measurement. Also, the phantom analysis uses a homogeneous sphere model that does not introduce any forward model inaccuracy, except the possible co-registration error. All four toolboxes show high localization accuracy and high resolution for phantom data, if the input SNR is not very low. Corresponding results for the static phantom data are presented in Fig. 6a–b. Fig. 6a indicates the localization error clear dependency on SNR. The nonlinear regression plots fitted over maximum localization errors indicate high localization errors at very low SNR raw data sets. The

15

440   high error is because of some unfiltered artifacts in raw data which was removed by SSS. After SSS,

441   the beamformer shows localization error under ~5 mm for all the datasets. Fig. 6b shows the

442   beamforming resolution in terms of PSV. The regression plots fitted over maximum PSV values show

443   a high spatial resolution for the data with SNR > 5 dB.

444

445   <u>Insert Fig. 6 about here</u>

446

447   In the cases of moving phantom, Fig. 7a shows high localization errors with unprocessed raw data

448   because of disturbances caused by the movement. The dipole excitation amplitude was 200 nAm,

449   which is enough to provide a good SNR. The most superficial dipoles (Dipoles 5 and 9 in Fig. 3c)

450   possess higher SNR but also higher localization error since they get more significant angular

451   displacement during movement. Because of differences in implementations and preprocessing

452   parameters listed in Section 2.4, apparent differences among the estimated localization error can be

453   seen. Overall, MNE-Python shows the lowest while SPM12 (DAiSS) shows the highest localization

454   error with the phantom data with movement artifact. After applying for spatiotemporal tSSS and

455   movement compensation, the improved SNR provided significantly better localization accuracies.

456   Fig. 7b shows the PSV for moving phantom data for raw and processed data. The regression plots

457   indicate improvement in SNR and spatial resolution after tSSS with movement compensation.

458

459   <u>Insert Fig. 7 about here</u>

460

461   **3.3. Human subject MEG data**

462   Since the correct source locations for the human evoked field datasets are unknown, we plotted the

463   localization difference across the four LCMV implementations for each data. These localization

464   differences were the Cartesian distance between an LCMV-estimated location and the

465   corresponding reference dipole location as explained in Section 2.1.4. Fig. 8a shows the plots for

466   the localization differences against the input SNRs computed using Eq (9) for four visual, two

467   auditory and two somatosensory evoked-field datasets. The localization differences for both

468   unprocessed raw and SSS preprocessed data are mostly under 20 mm in each toolbox. The higher

469   differences compared to the phantom and simulated dataset could be because of two reasons. First,

470   the recording might have been comprised by some head movement, which could not be corrected

471   because of the lack of continuous HPI. Second, the reference dipole location may not represent the

472   very same source as estimated by the LCMV beamformer. In contrast to dipole fitting, beamforming

473   utilizes data from the full covariance window, so some difference between the estimated localizations

16

474   is to be expected. For all SSS-preprocessed evoked field datasets, Fig. 8b shows the estimated

475   locations across the four LCMV implementation and the corresponding reference dipole locations.

476   For simplifying the visualization, all estimated locations in a stimulus category are projected onto a

477   single axial slice. All localizations seem to be in the correct anatomical regions, except the estimated

478   location from right-ear auditory responses by MNE-Python after SSS-preprocessing (Fig. 8b; red

479   circle). After de-selecting the channels close to the right auditory cortex, the MNE-Python-estimated

480   source location was correctly in the left cortex (Fig. 8b; green circle). The regression plots fitted over

481   the maxima of the localization differences across the LCMV implementations show the improvement

482   in input SNR and also localization improvement in some cases. Fig. 9 in Supplementary material

483   shows the PSV values as a function of the input SNR for the evoked-field datasets, demonstrating

484   the spatial resolution of beamforming.

485

486   <u>Insert Fig. 8 about here</u>

487

488

17

4. Discussion

The localization accuracy and beamformer resolution as a function of the input SNR were investigated and compared across the LCMV implementations in the four tested toolboxes. In the absence of background noise, the phantom data showed high localization accuracy and high spatial resolution if the input SNR >~5 dB. All implementations also showed high localization accuracy for data recording from a moving phantom after compensating the movement and applying tSSS. For the simulated datasets with realistic background noise, the regression curve fitted over the maximum localization error across the LCMV implementations indicates that the reliability of localization accuracy in these implementations depends on the SNR of input data and these implementations localize a single source reliably within the SNR range of ~3–15 dB. Small differences among the estimated source locations across the implementations even in this SNR range are caused by differing processing steps in defining the head model, spatial filter and performing the beamformer scan. For the human subject evoked-field MEG data, all implementations localize sources within about 20 mm.

Our results indicate that with the default parameter settings, none of the four implementations works universally reliable for all datasets and input SNR values. In the case of low SNR (typically less than 3 dB), the lower contrast between data and noise covariance may cause the beamformer scan to provide a flat peak in the output and so the localization error goes high. Unexpectedly, we found high localization error for high SNR signal and significant differences between the toolboxes. The regression curves fitted over averaged maximum PSV across all toolboxes showed higher values for low- and high-SNR simulated data. As expected, reliable localization provides higher spatial resolution across the implementations and vice-versa (Fig. 5 and 6). The lower spatial resolution (higher PSV) for the signal with low SNR also agrees with previous studies (Lin et al., 2008; Hillebrand and Barnes, 2003). We further discuss here the significant steps of the beamformer pipelines, which affect the localization accuracy and introduce discrepancies among the implementations.

4.1 Preprocessing with SSS

Due to the spatial-filter nature of the beamformer, it can reject external interference and therefore SSS-based pre-processing may have little effect on the results. Thus, although the SNR increases as a result of applying SSS, the localization accuracy does not necessarily improve, which is evident in the localization of the evoked responses (Fig. 8).

However, undetected artifacts, such as a large-amplitude signal jump in a single sensor, may in SSS processing spread to neighboring channels and subsequently reduce data quality. Therefore, channels with distinct artifacts should be noted and marked as bad prior to beamforming of unprocessed data or before applying SSS operations. In addition, trials with large artifacts should be removed based on an amplitude thresholding or by other means. Furthermore, SSS processing of

525 extremely weak signals (SNR < ~2 dB) may not improve the SNR for producing smaller localization
526 errors and PSV values. Hence the data quality should be carefully inspected before and after
527 applying preprocessing methods such as SSS, and channels or trials with low-quality data (or lower
528 contrast) should be omitted from the covariance estimation.

4.2. Effect of filtering and artifact-removal methods

530 All four toolboxes we tested employ either a MATLAB or Python implementation of the same MNE
531 routines (Gramfort et al. 2014) for reading FIFF data files and thus have internally the exact same
532 data at the very first stage (see Suppl. Fig. 6). The data import either keeps the data in SI-units (T
533 for magnetometers and T/m for gradiometers) or rescales the data (fT and fT/mm) before further
534 processing. The actual pre-processing steps in the pipeline may contribute to differences in the
535 results. The filtering step is performed to remove frequency components of no interest, such as slow
536 drifts, from the data. By default, FieldTrip and SPM use an IIR (Butterworth) filter, and MNE-Python
537 uses FIR filters. The power spectra of these filters' output signals show notable differences and the
538 output data from these two filters are not identical. Significant variations can be found between MNE-
539 Python-filtered and FieldTrip/SPM-filtered data. Although SPM and FieldTrip use the same filter
540 implementation, the filtering results are not identical because of numeric differences caused by
541 different channel scaling (Suppl. Fig 6). These differences affect the estimated covariance matrices,
542 which are a crucial ingredient for the spatial-filter computation and finally may contribute to
543 differences in beamforming results.

4.3. Effect of SNR on localization accuracy

545 We reduced the impact of the unknown source depth and strength to a well-defined metrics in terms
546 of the SNR. We observed that the localization accuracy is poor for very low SNR values, i.e. below
547 3 dB. The weaker, as well as the deeper sources, project less power on to the sensor array and thus
548 show lower SNR; see Eq (9). On the other hand, the LCMV beamformer may also fail to localize
549 accurately sources that produce very high SNR values, likely because the data covariance matrix is
550 over-fitted, or the scanning grid is too sparse with respect to the point spread of the beamformer
551 output. In this case the output is too focal and a small error in forward solution, introduced for
552 example by inaccurate coregistration, may lead to missing the true focal source and obtaining nearly
553 equal power estimates at many source locations, increasing the chance of mislocalization. Usually,
554 such high levels of SNR do not occur in typical human MEG experiments, however, the strength of
555 equivalent current dipoles (ECD) for interictal epileptiform discharges (IIEDs) typically ranges
556 between 50 and 500 nAm (Bagic et al., 2011a).

557 All four beamformer pipelines provided very similar results when the SNR is in the "suitable range"
558 of about ~3–15 dB. Unsatisfactory performance is typically due to the data; either the SNR is
559 extremely low, or there are some uncorrected artifacts in the data. The results of the phantom data

19

560   showed that all toolboxes provide equally good results if there are no uncorrected large artifacts in
561   the data and if the SNR is not extremely small or large.

4.4. Effect of the head model

563   Forward modelling requires MEG–MRI co-registration, segmentation of the head MRI and leadfield
564   computation for the source space. The four beamformer implementations use different approaches,
565   or similar approaches but with different parameters, which yields slightly different forward models.
566   From Eqs (2–7), it is evident that beamformers are quite sensitive to the forward model. Hillebrand
567   and Barnes (2003) showed that the spatial resolution and the localization accuracy of a beamformer
568   improve with accuracy of the forward model. Dalal and colleagues (2014) reported that co-
569   registration errors contribute greatly to EEG localization inaccuracy, likely due to their ultimate impact
570   on head-model quality. Chella and colleagues (2019) presented the dependency of beamformer-
571   based functional connectivity estimates on MEG-MRI co-registration accuracy.

572   The increasing inter-toolbox localization differences towards very low and very high input SNR is
573   also subject to the differences between the head models. Fig. 4 shows the three overlapped head
574   models prepared from the same MRI where a slight misalignment among head models can be easily
575   seen. This misalignment also affects source space. These differences in head models and thus in
576   forward solutions will contribute to differences in beamforming results across the toolboxes.

4.5. Covariance matrix

578   The data covariance matrix is a key component of the adaptive spatial filter in LCMV beamforming,
579   and any error in covariance estimation can cause an error in source estimation. We used 5% of the
580   mean variance of all sensors to regularize data covariance for making its inversion stable in FieldTrip,
581   SPM12 and MNE-Python. Brainstorm uses a slightly different approach and applies regularization
582   with 5% of mean variance of gradiometer and magnetometer channel sets separately and eliminate
583   cross-sensor-type entries from the covariance matrices. As SSS preprocessing reduces the rank of
584   the data, usually retaining at most 75 non-zero eigenvalues, the trace of the covariance matrix
585   decreases strongly. At very high SNRs (> 15 dB), overfitting of the covariance matrix becomes more
586   prominent; the condition number (ratio of the largest and the smallest eigenvalues) of the covariance
587   matrix becomes very high even after the default regularization, which can deteriorate the quality of
588   source estimates unless the covariance is appropriately regularized. Therefore, the seemingly same
589   5% regularization can have very different effects before and after SSS; see Suppl. Fig. 7. Thus, the
590   commonly used way of specifying the regularization level might not be appropriate to produce a good
591   and stable covariance model at high SNR, and this could be one of the explanations for the
592   anecdotally reported detrimental effects of SSS on beamforming results.

20

5. Conclusion

We conclude that with the current versions of LCMV beamformer implementations in the four open-source toolboxes — FieldTrip, SPM12(DAiSS), Brainstorm, and MNE-Python — the localization accuracy is acceptable (within ~10 mm for a true point source) for most purposes when the input SNR is 3–15 dB. Lower or higher SNR may compromise the localization accuracy and spatial resolution. To extend this useable range, a properly defined scaling strategy such as pre-whitening should be implemented across the toolboxes. The default regularization is often inadequate and may yield suboptimal results. Therefore, a data-driven approach for regularization should be adopted to alleviate problems with low- and high-SNR cases. Our further work will be focusing on optimizing regularization using a more data-driven approach.

References

Bagic, A. I., Knowlton, R. C., Rose, D. F., Ebersole, J. S., & ACMEGS Clinical Practice Guideline (CPG) Committee. (2011a). American clinical magnetoencephalography society clinical practice guideline 1: recording and analysis of spontaneous cerebral activity. *J Clin Neurophysiol.* **28**(4): 348–354. http://dx.doi.org/10.1097/WNP.0b013e3182272fed

Bagic, A. I., Knowlton, R. C., Rose, D. F., Ebersole, J. S., & ACMEGS Clinical Practice Guideline (CPG) Committee. (2011b). American Clinical Magnetoencephalography Society Clinical Practice Guideline 3: MEG–EEG Reporting. *J Clin Neurophysiol.* **28**(4): 362–365. http://dx.doi.org/10.1097/WNO.0b013e3181cde4ad

Barca, L., Cornelissen, P., Simpson, M., Urooj, U., Woods, W., & Ellis, A. W. (2011). The neural basis of the right visual field advantage in reading: an MEG analysis using virtual electrodes. *Brain Lang* **118**(3): 53–71. http://dx.doi.org/10.1016/j.bandl.2010.09.003

Burgess, R. C., Funke, M. E., Bowyer, S. M., Lewine, J. D., Kirsch, H. E., Bagić, A. I., & ACMEGS Clinical Practice Guideline (CPG) Committee (2011). American Clinical Magnetoencephalography Society Clinical Practice Guideline 2: Presurgical functional brain mapping using magnetic evoked fields. *J Clin Neurophysiol.* **28**(4): 355–361. http://dx.doi.org/10.1097/WNP.0b013e3182272ffe

625 Chella, F., Marzetti, L., Stenroos, M., Parkkonen, L., Ilmoniemi, R. J., Romani, G. L., & Pizzella, V.
626 (2019). The impact of improved MEG–MRI co-registration on MEG connectivity analysis.
627 *NeuroImage*, **197**: 354-367. https://doi.org/10.1016/j.neuroimage.2019.04.061

628 Dalal, S. S., Rampp, S., Willomitzer, F., & Ettl, S. (2014). Consequences of EEG electrode position
629 error on ultimate beamformer source reconstruction performance. *Front Neurosci.* **8**: 42.
630 http://dx.doi.org/10.3389/fnins.2014.00042

631 Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis I: Segmentation and
632 surface reconstruction. *Neuroimage* **9**(2): 179-194. https://doi.org/10.1006/nimg.1998.0395

633 Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., & Halgren, E.
634 (2000). Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution
635 imaging of cortical activity. *Neuron* **26**(1): 55–67. http://dx.doi.org/10.1016/S0896-6273(00)81138-1

636 Elekta Neuromag® TRIUX User's Manual, (Megin Oy, 2018)

637 Engemann, D. A., & Gramfort, A. (2015). Automated model selection in covariance estimation and
638 spatial whitening of MEG and EEG signals. *NeuroImage* **108**: 328–342.
639 http://dx.doi.org/10.1016/j.neuroimage.2014.12.040

640 Fischl, B. (2012). FreeSurfer. *Neuroimage* **62**(2): 774–781.
641 https://doi.org/10.1016/j.neuroimage.2012.01.021

642 Fischl, B., Sereno, M. I. & Dale, A. M. (1999). Cortical surface-based analysis II: Inflation, flattening,
643 and a surface-based coordinate system. *NeuroImage* **9**(2): 195–207.
644 https://doi.org/10.1006/nimg.1998.0396

645 Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., ... &
646 Hämäläinen, M. (2013a). MEG and EEG data analysis with MNE-Python. *Front Neurosci.* **7**: 267.
647 http://dx.doi.org/10.3389/fnins.2013.00267

648 Gramfort, A., Strohmeier, D., Haueisen, J., Hämäläinen, M. S., & Kowalski, M. (2013b). Time-
649 frequency mixed-norm estimates: Sparse M/EEG imaging with non-stationary source activations.
650 *NeuroImage* **70**: 410–422. http://dx.doi.org/10.1016/j.neuroimage.2012.12.051

651 Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., ... &
652 Hämäläinen, M. S. (2014). MNE software for processing MEG and EEG data. *Neuroimage* **86**: 446–
653 460. http://dx.doi.org/10.1016/j.neuroimage.2013.10.027

654 Gross, J., Kujala, J., Hämäläinen, M., Timmermann, L., Schnitzler, A., & Salmelin, R. (2001).
655 Dynamic imaging of coherent sources: studying neural interactions in the human brain. *Proc Natl
656 Acad Sci USA* **98**(2): 694-699. http://dx.doi.org/10.1073/pnas.98.2.694

657 Hall, M. B., Nissen, I. A., van Straaten, E. C., Furlong, P. L., Witton, C., Foley, E., ... & Hillebrand, A.
658 (2018). An evaluation of kurtosis beamforming in magnetoencephalography to localize the
659 epileptogenic zone in drug-resistant epilepsy patients. *Clin Neurophysiol.* **129**(6): 1221–1229.
660 http://dx.doi.org/10.1016/j.clinph.2017.12.040

661 Hämäläinen, M., Hari, R., Lounasmaa, O. V., Knuutila, J., & Ilmoniemi, R. J. (1993).
662 Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the
663 working human brain. *Rev Mod Phys* **65**: 413–497. https://doi.org/10.1103/RevModPhys.65.413

664 Hämäläinen, M. S., & Ilmoniemi, R. J. (1994). Interpreting magnetic fields of the brain: minimum
665 norm estimates. *Med Biol Eng Compt.* **32**(1): 35–42. https://doi.org/10.1007/BF02512476

666 Hämäläinen, M. S., & Sarvas, J. (1989). Realistic conductivity geometry model of the human head
667 for interpretation of neuromagnetic data. *IEEE Trans Biomed Eng.* **36**(2): 165–171.
668 https://doi.org/10.1109/10.16463

669 Hillebrand, A., & Barnes, G. R. (2003). The use of anatomical constraints with MEG beamformers.
670 *Neuroimage* **20**(4): 2302–2313. http://dx.doi.org/10.1016/j.neuroimage.2003.07.031

671 Hillebrand, A., & Barnes, G. R. (2005). Beamformer analysis of MEG data. *Int Rev Neurobiol* **68**:
672 149–171. http://dx.doi.org/10.1016/S0074-7742(05)68006-3

673 Huang, M. X., Mosher, J. C., & Leahy, R. M. (1999). A sensor-weighted overlapping-sphere head
674 model and exhaustive head model comparison for MEG. *Phys Med Biol.* **44**(2): 423-440.
675 http://dx.doi.org/10.1088/0031-9155/44/2/010

676 Ilmoniemi, R. J., Hämäläinen, M. S., & Knuutila, J. (1985). The forward and inverse problems in the
677 spherical model, in: Weinberg, H., Stroink, G. & Katila, T. (Eds.), Biomagnetism: Applications and
678 Theory. *Pergamon Press, New York*, 278–282

679 Ilmoniemi, R. J., & Sarvas, J. (2019). Brain Signals: Physics and Mathematics of MEG and EEG.
680 *MIT Press.*

681 Ishii, R., Canuet, L., Ishihara, T., Aoki, Y., Ikeda, S., Hata, M., ... & Iwase, M. (2014). Frontal midline
682 theta rhythm and gamma power changes during focused attention on mental calculation: an MEG
683 beamformer analysis. *Front Hum Neurosci.* **8**: 406. http://dx.doi.org/10.3389/fnhum.2014.00406

684 Lin, F. H., Witzel, T., Zeffiro, T. A., & Belliveau, J. W. (2008). Linear constraint minimum variance
685 beamformer functional magnetic resonance inverse imaging. *Neuroimage* **43**(2): 297–311.
686 http://dx.doi.org/10.1016/j.neuroimage.2008.06.038

687 Litvak, V., Mattout, J., Kiebel, S., Phillips, C., Henson, R., Kilner, J., ... & Penny, W. (2011). EEG and
688 MEG data analysis in SPM8. *Comput Intell Neurosci.* **2011**. http://dx.doi.org/10.1155/2011/852961

689 Mattout, J., Henson, R. N., & Friston, K. J. (2007). Canonical source reconstruction for MEG. *Comput*
690 *Intell Neurosci.* **2007**. http://dx.doi.org/10.1155/2007/67613

691 Mosher, J. C., Lewis, P. S., & Leahy, R. M. (1992). Multiple dipole modeling and localization from
692 spatio-temporal MEG data. *IEEE Trans Biomed Eng.* **39**(6): 541–557.
693 http://dx.doi.org/10.1109/10.141192

694 Nenonen J, Nurminen J, Kicic D, Bikmullina R, Lioumis P, Jousmäki V, Taulu S, Parkkonen L,
695 Putaala M, Kähkönen S (2012). Validation of head movement correction and spatiotemporal signal
696 space separation in magnetoencephalography. *Clin Neurophysiol.* **123**(11): 2180–2191.
697 http://dx.doi.org/10.1016/j.clinph.2012.03.080

698 Nolte, G. (2003). The magnetic lead field theorem in the quasi-static approximation and its use for
699 magnetoencephalography forward calculation in realistic volume conductors. *Phys Med Biol.* **48**(22):
700 3637. http://dx.doi.org/10.1088/0031-9155/48/22/002

701 Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for
702 advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci.*
703 **2011**. http://dx.doi.org/10.1155/2011/156869

704 Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1994). Low resolution electromagnetic
705 tomography: a new method for localizing electrical activity in the brain. *Int J Psychophysiol.* **18**(1):
706 49-65. https://doi.org/10.1016/0167-8760(84)90014-X

707 Robinson S.E., Vrba, J. (1998). Functional neuroimaging by synthetic aperture magnetometry
708 (SAM), in: Yoshimoto, T., Kotani, M., Kuriki, S., Karibe, H. & Nakasato, N. (Eds.), Recent Advances
709 in Biomagnetism. *Tohoku University Press, Japan,* 302–305.

710 Salmelin, R. (2010). Multi-dipole modeling in MEG, in: Hansen, P., Kringelbach, M. & Salmelin, R.
711 (Eds.), MEG: an introduction to methods. *Oxford university press*, 124–155.
712 http://dx.doi.org/10.1093/acprof:oso/9780195307238.003.0006

713 Sekihara, K., & Scholz, B. (1996). Generalized Wiener estimation of three-dimensional current
714 distribution from biomagnetic measurements. *IEEE Trans Biomed Eng.* **43**(3): 281-291.
715 http://dx.doi.org/10.1109/10.486285

716 Sekihara, K., Hild, K. E., & Nagarajan, S. S. (2006). A novel adaptive beamformer for MEG source
717 reconstruction effective when large background brain activities exist. *IEEE Trans Biomed Eng.* **53**(9),
718 1755–1764. http://dx.doi.org/10.1109/TBME.2006.878119

719 Sekihara, K., & Nagarajan, S., S. (2008). Adaptive spatial filters for electromagnetic brain imaging.
720 *Springer Science & Business Media.* https://doi.org/10.1007/978-3-540-79370-0

721 Sekihara, K., Nagarajan, S. S., Poeppel, D., & Marantz, A. (2004). Asymptotic SNR of scalar and
722 vector minimum-variance beamformers for neuromagnetic source reconstruction. *IEEE Trans*
723 *Biomed Eng.* **51**(10): 1726–1734. http://dx.doi.org/10.1109/TBME.2004.827926

724 Spencer, M. E., Leahy, R. M., Mosher, J. C., & Lewis, P. S. (1992). Adaptive filters for monitoring
725 localized brain activity from surface potential time series. in: *Conference Record of the Twenty-Sixth*
726 *Asilomar Conference on Signals, Systems & Computers.* **1992**: 156-161.
727 http://dx.doi.org/10.1109/ACSSC.1992.269278

728 Stenroos, M., Hunold, A., & Haueisen, J. (2014). Comparison of three-shell and simplified volume
729 conductor models in magnetoencephalography. *NeuroImage* **94**: 337-348.
730 https://doi.org/10.1016/j.neuroimage.2014.01.006

731 Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: a user-friendly
732 application for MEG/EEG analysis. *Comput Intell Neurosci.*, **2011**.
733 http://dx.doi.org/10.1155/2011/879716

734 Taulu, S., & Kajola, M. (2005). Presentation of electromagnetic multichannel data: the signal space
735 separation method. *J Appl Physics* **97**(12): 124905. http://dx.doi.org/10.1063/1.1935742

736 Tikhonov, A. N. (1963). Solution of incorrectly formulated problems and the regularization method.
737 *Soviet Math.*, **4**: 1035–1038.

738 van Es, M. W., & Schoffelen, J. M. (2019). Stimulus-induced gamma power predicts the amplitude
739 of the subsequent visual evoked response. *NeuroImage* **186**: 703-712.
740 http://dx.doi.org/10.1016/j.neuroimage.2018.11.029

741 van Klink, N., van Rosmalen, F., Nenonen, J., Burnos, S., Helle, L., Taulu, S., ... & Hillebrand, A.
742 (2017). Automatic detection and visualization of MEG ripple oscillations in epilepsy. *NeuroImage*
743 *Clin* **15**: 689–701. https://doi.org/10.1016/j.nicl.2017.06.024

744 Van Veen, B. D., & Buckley, K. M. (1988). Beamforming: A versatile approach to spatial filtering.
745 *IEEE ASSP Magazine* **5**(2): 4–24. http://dx.doi.org/10.1109/53.665

746 Van Veen, B. D., Van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain
747 electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans Biomed Eng.*
748 **44**(9): 867–880. http://dx.doi.org/10.1109/10.623056

749 Vrba, J. (2000). Differences between synthetic aperture magnetometry (SAM) and linear
750 beamformers. In: *Proceedings of the 12th International Conference on Biomagnetism, Espoo,*
751 *Finland.* **2000**: 681-684

752   Wipf, D. P., Owen, J. P., Attias, H. T., Sekihara, K., & Nagarajan, S. S. (2010). Robust Bayesian

753   estimation of the location, orientation, and time course of multiple correlated neural sources using

754   MEG. *NeuroImage* **49**(1): 641–655. http://dx.doi.org/10.1016/j.neuroimage.2009.06.083

755   Youssofzadeh, V., Agler, W., Tenney, J. R., & Kadis, D. S. (2018). Whole-brain MEG connectivity-

756   based analyses reveals critical hubs in childhood absence epilepsy. *Epilepsy Res.* **145**: 102-109.

757   http://dx.doi.org/10.1016/j.eplepsyres.2018.06.001

|	MNE-Python	FieldTrip	SPM12 (DAiSS)	Brainstorm
Version	0.18	20190922	20190924	20190926
Data import functions	MNE (Python)	MNE (Matlab)	MNE (Matlab)	MNE (Matlab)
Internal units of MEG data	T, T/m	T, T/m	fT, fT/mm	T, T/m
Band-pass filter type	FIR	IIR	IIR	FIR
MRI segmentation	FreeSurfer	SPM8/SPM12	SPM8/SPM12	FreeSurfer/SPM8
Head model	Single-shell BEM	Single-shell corrected sphere	Single-shell corrected sphere	Overlapping spheres
Source space	Rectangular grid (5 mm)	Rectangular grid (5 mm)	Rectangular grid (5 mm)	Rectangular grid (5 mm)
MEG–MRI coregistration	Point-cloud co-registration and manual correction	3-point manual co-registration followed by ICP co-registration	Point-cloud co-registration using ICP	Point-cloud co-registration using ICP
Data covariance matrix	Sample data covariance	Sample data covariance	Sample data covariance	Sample data covariance
Noise normalization for NAI computation	Sample noise covariance	Sample noise covariance	Sample noise covariance	Sample noise covariance
Combining data from multiple sensor types	Prewhitening (full noise covariance)	No scaling or prewhitening	No scaling or prewhitening	Prewhitening (full noise covariance but cross-sensor-type terms zeroed)
Beamformer type	**Scalar**	**Scalar**	**Scalar**	Vector
Beamformer output	**Neural activity index (NAI)**	**Neural activity index (NAI)**	**Neural activity index (NAI)**	**Neural activity index (NAI)**

Table 1. Characteristics of the four beamforming toolboxes. The non-default settings of each toolbox are shown in bold. The toolbox version is indicated either by the version number or by the download date (*yyyymmdd)* from GitHub.
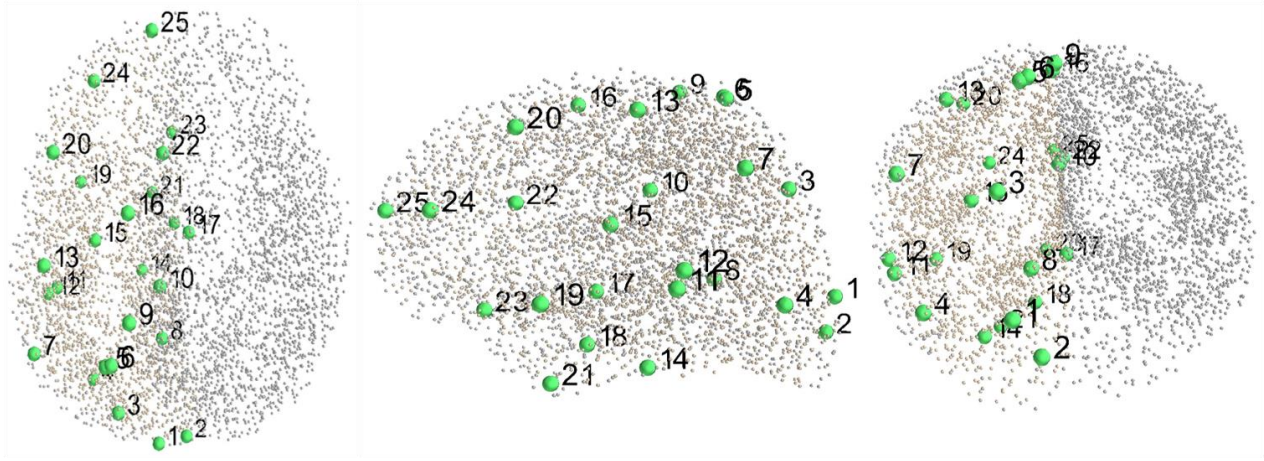
**Fig. 1.** The 25 simulated dipolar sources (green dots) in the source space (grey dots).
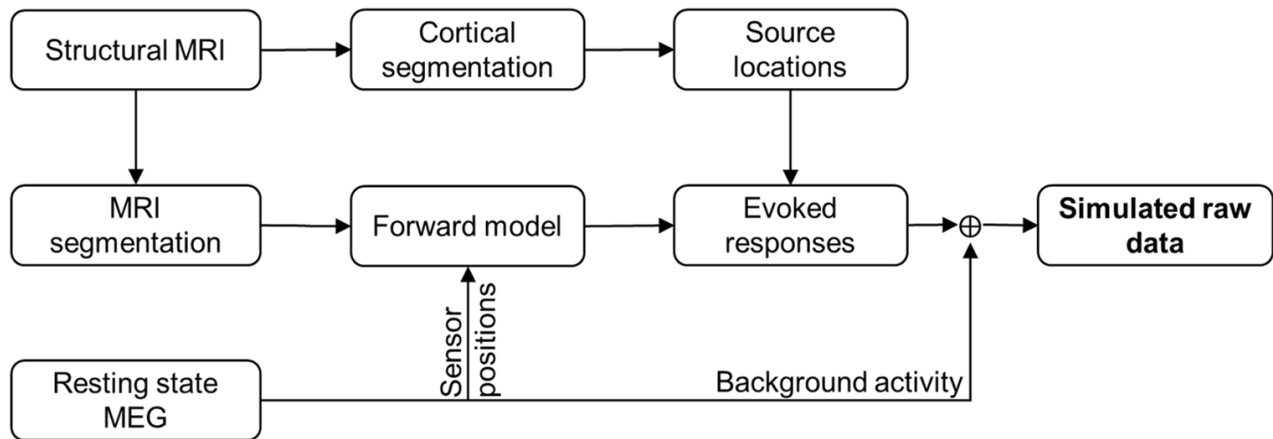
**Fig. 2**. Simulating MEG data (detailed workflow in Suppl. Fig. 1).
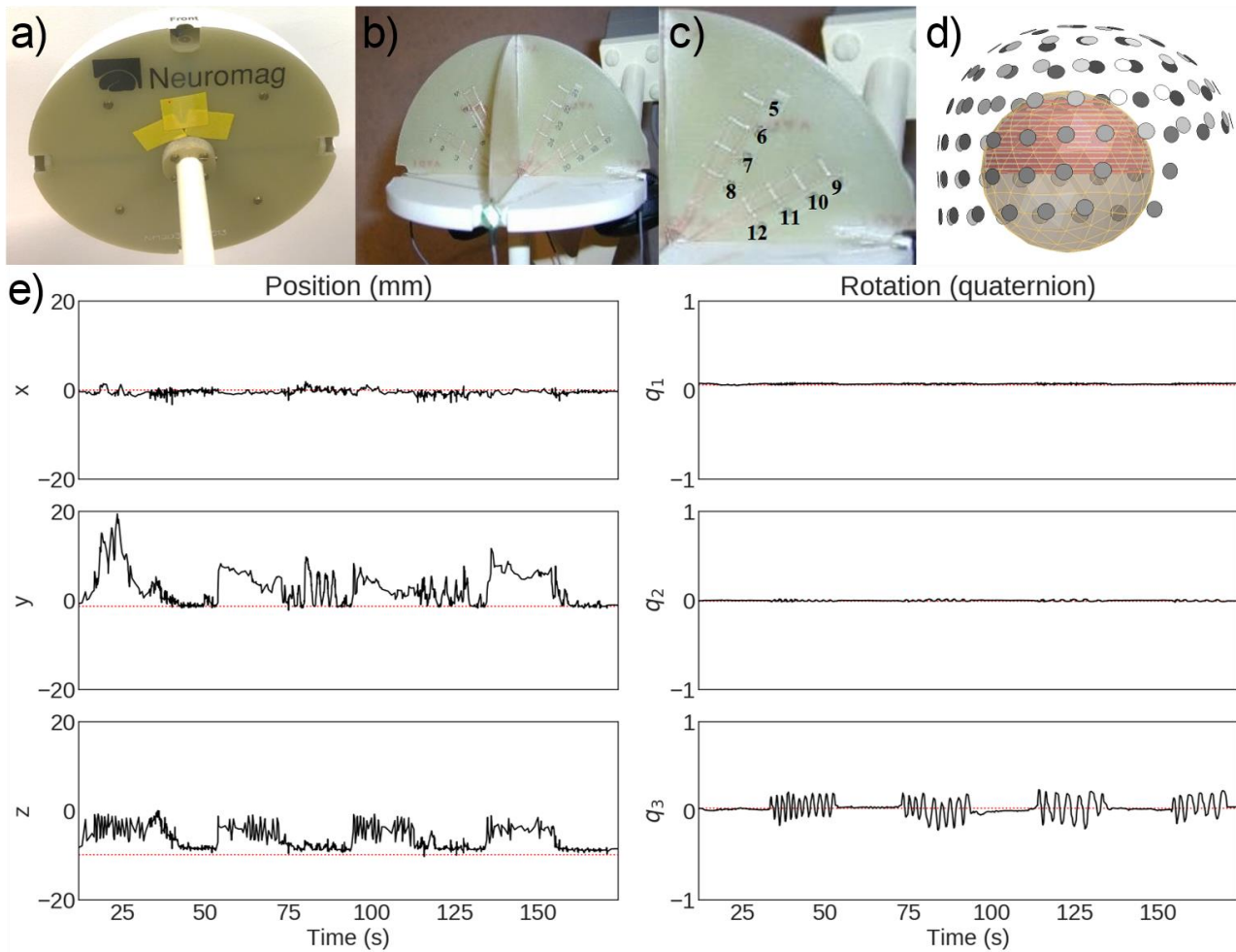
Fig. 3. The dry phantom measured in this study. (a) Outer view, (b) cross-section, (c) positions of the employed dipole sources, (d) phantom position with respect to the MEG sensor helmet, and (e) position and rotation of the phantom during one of the moving-phantom measurements (Dipole 9 activated).
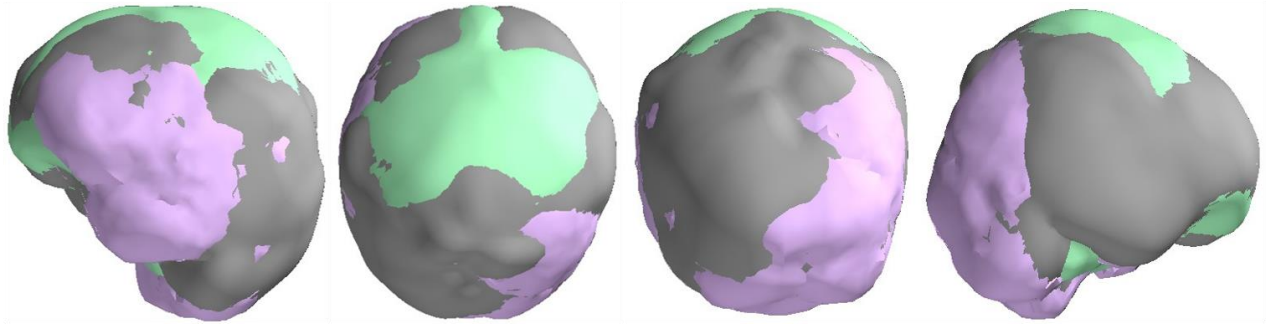
Fig. **4**: Correspondence between the head models used by MNE-Python (grey), FieldTrip (lavender) and SPM12 (mint). The Brainstorm head model is not included here as it uses overlapping spheres. The outermost surface (inner skull) across the toolboxes is rendered visible.
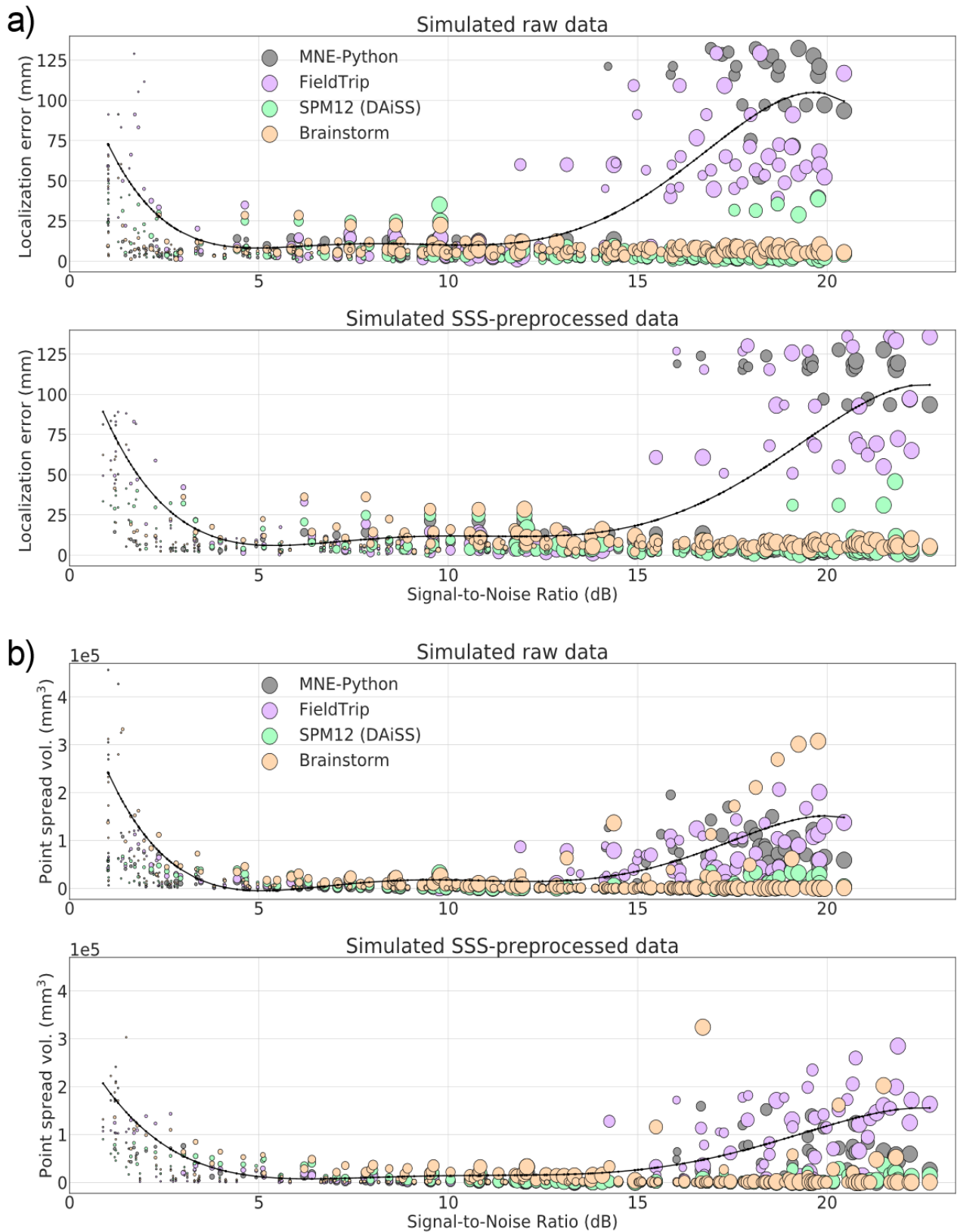
Fig. 5. Localization error (a) and point-spread volume (b) as a function of input SNR for raw and SSS-preprocessed simulated datasets. The markers size indicates the true dipole amplitude. The curves (black) indicate the polynomial regression of the maximal value across the four LCMV implementations.
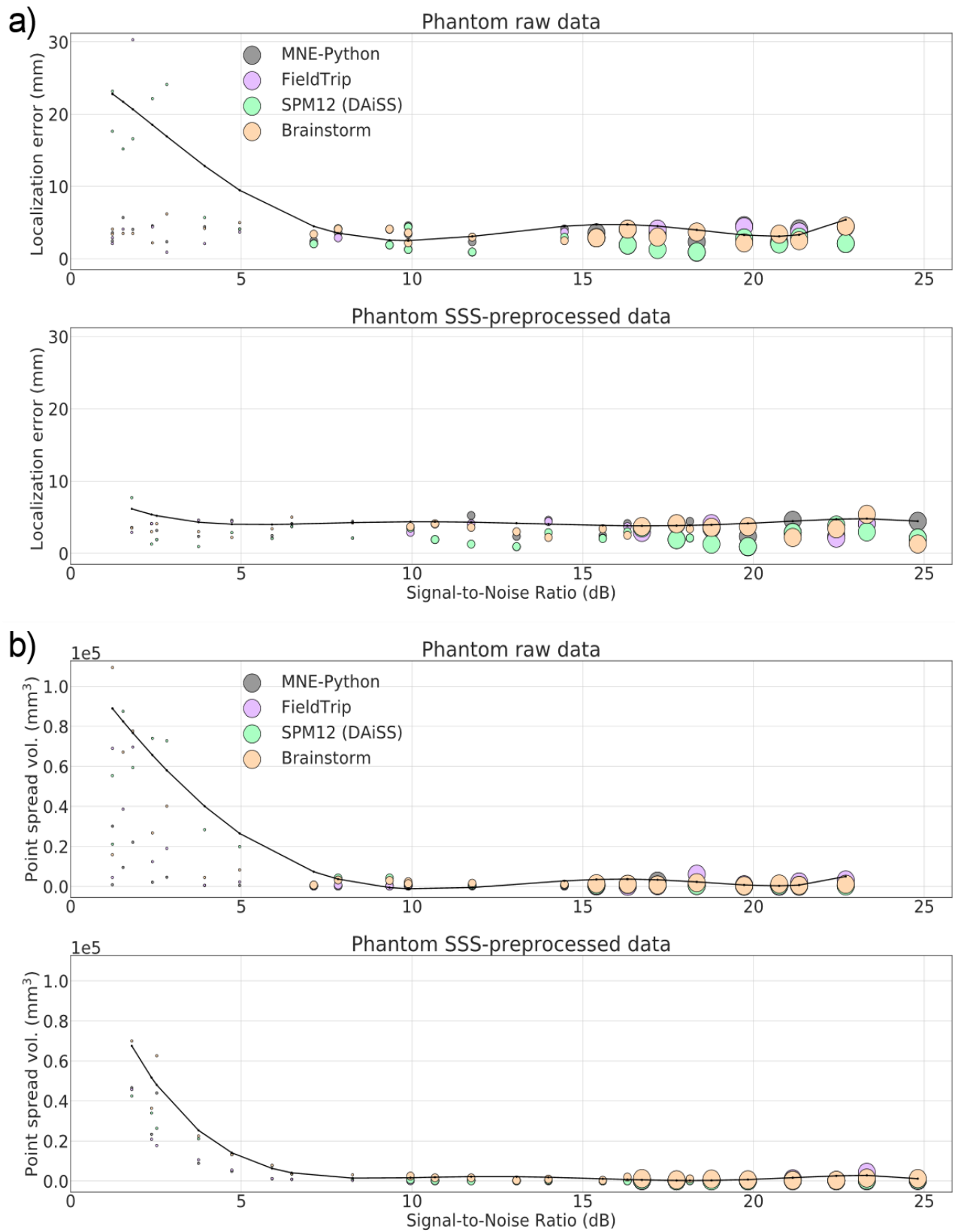
Fig. 6. Localization error (a) and point-spread volume (b) as a function of input SNR for phantom data recording in a stable position. The markers size indicates the true dipole amplitude. The curves (black) indicate the polynomial regression of the across the four LCMV implementations.
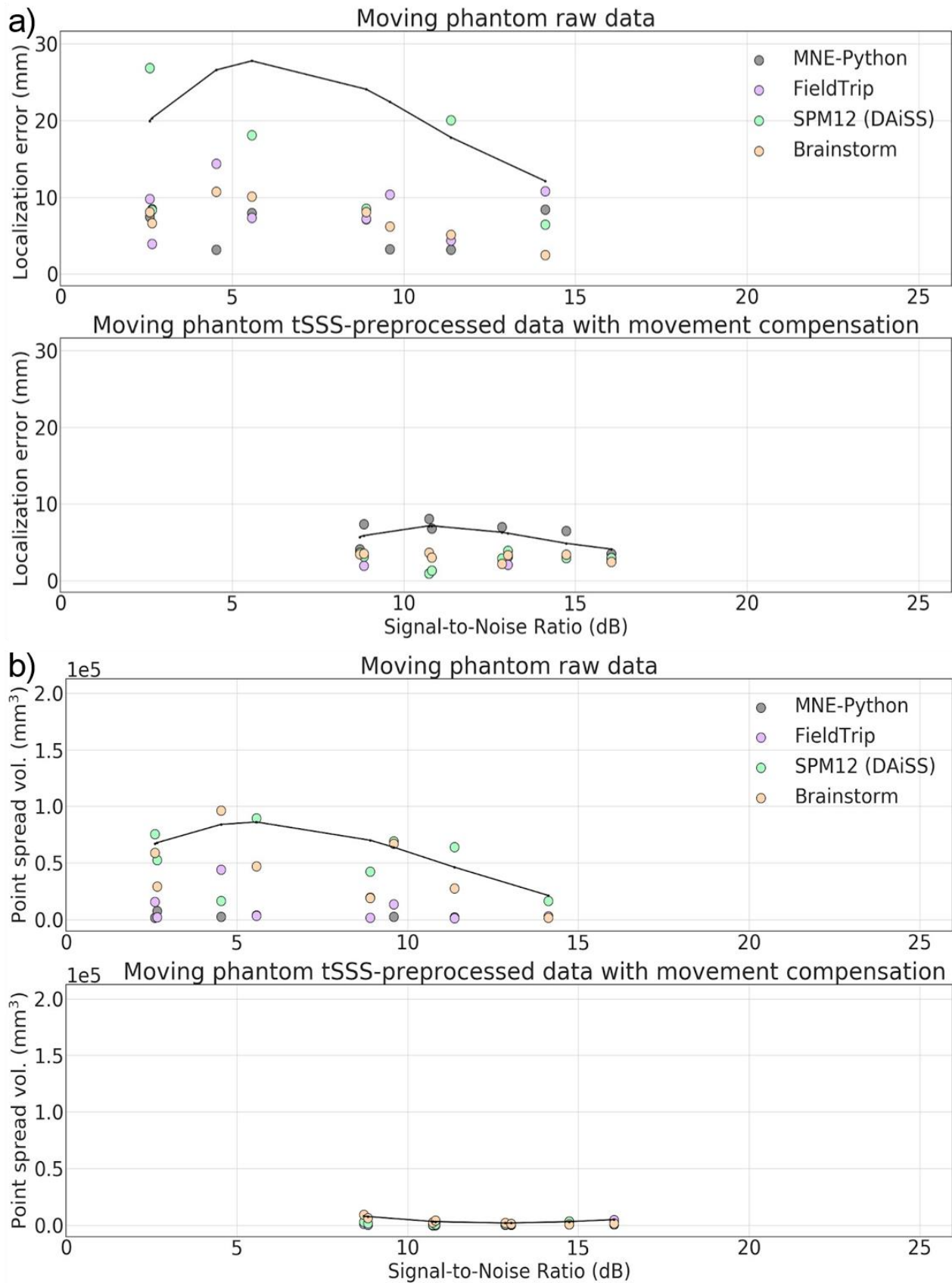
Fig. 7. Localization error (a) and point-spread volume (b) as a function of input SNR for data from the moving phantom. The curves (black) indicate the polynomial regression of the maximum value across the four LCMV implementations.
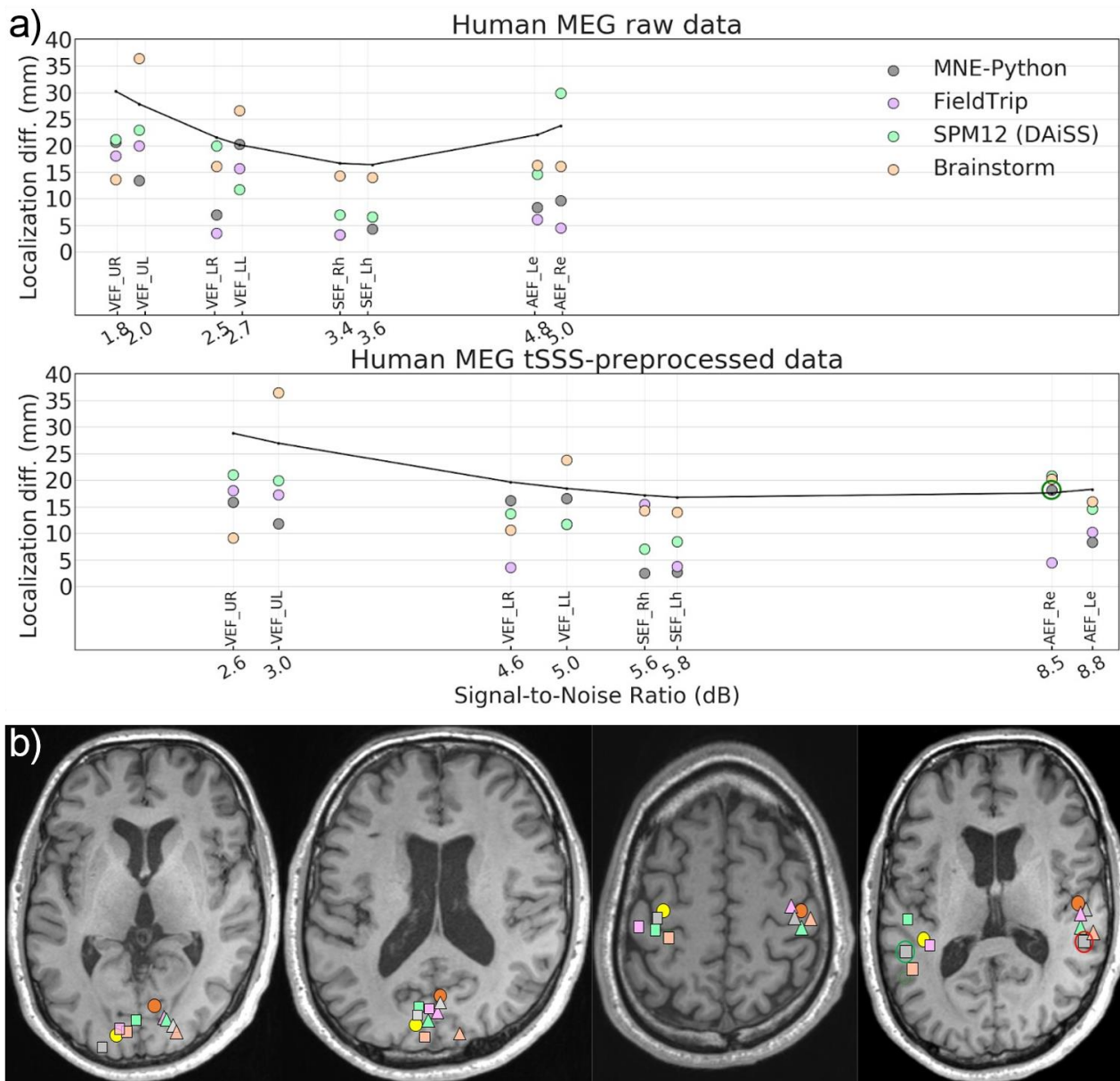
Fig. 8. Source estimates of human MEG data. (a) Localization difference from the reference dipole location for raw and tSSS-preprocessed data. (b) Peaks of the beamformer source estimate of tSSS-processed data. From left to right: visual stimuli presented to left (triangle) and right (square) upper and lower quadrant of the visual field (the two axial slices showing all sources); somatosensory stimuli to left (triangle) and right (square) wrist; auditory stimuli to the left (triangle) and right (square) ear. Reference dipole locations (yellow and orange circles).