NeuroImage 95 (2014) 330–335

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Technical Note

Multi-session statistics on beamformed MEG data

Henry T. Luckhoo [a,b], Matthew J. Brookes [c], Mark W. Woolrich [a,d,*]

[a] Oxford Centre for Human Brain Activity, University of Oxford, Warneford Hospital, Oxford, UK
[b] Centre for Doctoral Training in Healthcare Innovation, Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, UK
[c] Sir Peter Mansfield Magnetic Resonance Centre, School of Physics and Astronomy, University of Nottingham, University Park, Nottingham, UK
[d] FMRIB Centre, University of Oxford, Oxford, UK

ARTICLE INFO

Article history:
Accepted 14 December 2013
Available online 9 January 2014

Keywords:
MEG
Group statistics
Beamforming
Source reconstruction

ABSTRACT

Beamforming has been widely adopted as a source reconstruction technique in the analysis of magnetoencephalography data. Most beamforming implementations incorporate a spatially-varying rescaling (which we term *weights normalisation*) to correct for the inherent depth bias in raw beamformer estimates. Here, we demonstrate that such rescaling can cause critical problems whenever analyses are performed over multiple sessions of separately beamformed data, for example when comparing effect sizes between different populations. Importantly, we show that the weights-normalised beamformer estimates of neural activity can even lead to a reversal in the inferred sign of the effect being measured. We instead recommend that no weights normalisation be carried out; any depth bias is instead accounted for in the calculation of multi-session (e.g. group) statistics. We demonstrate the severity of the weights normalisation confound with a 2-D simulation, and in real MEG data by performing a group statistical analysis to detect differences in alpha power in eyes-closed rest compared with continuous visual stimulation.

Introduction

Magnetoencephalography (MEG) is a neuroimaging modality that provides an electrophysiological assessment of neural activity across the whole brain, sampled at millisecond temporal resolution. However, to fully leverage the rich spatio-temporal information contained within MEG data, it is necessary to infer the source-level neural activity from the sensor-level fields. A wide range of source reconstruction techniques have been developed to estimate the source space activity including, but not limited to, beamforming (Robinson and Vrba, 1998; Van Veen et al., 1997), minimum norm estimates (Hämäläinen and Ilmoniemi, 1984), and multiple sparse priors (Friston et al., 2008).

Beamformers are spatial filters designed to have a unit pass band at a specific spatial location of interest whilst minimising the overall variance passed by the spatial filter (Robinson and Vrba, 1998; Van Veen et al., 1997). However, it is well known that the raw estimates of the variance of beamformed data contain a spatial bias that leads to the variance increasing with depth (Hall et al., 2013). As we will show, this increase in variance is due to the contribution of projected sensor noise to the beamformed time series. This projection is greater for deep regions of the brain because they are further from the sensors. In order to correct for this depth bias, beamformer implementations typically contain a spatial correction to remove it. This equates to dividing the raw variance estimate by some estimate of the variance of the projected sensor noise. Examples include the neural activity index (Van Veen et al., 1997) and the pseudo-Z-statistic (Vrba and Robinson, 2001). In this paper, we consider one type of rescaling, which we refer to as *weights normalisation*. However, our findings are applicable to any rescaling correction to remove the beamformer depth bias.

It is important to note that weights normalisation is, by definition, specific to an individual session of beamformed data. In this paper, we show that this has major implications when performing any analyses over multiple, separately beamformed sessions. One approach is to estimate a common set of beamformer weights for all sessions being analysed. This can be achieved by concatenating all sessions prior to estimating a single covariance matrix. However, this has two key limitations. Firstly it cannot be applied to group analyses with multiple subjects as the forward models will differ. Secondly, even in multi-session analyses that use the same subject, movement of the participant's head across the concatenated sessions will limit the accuracy of the source reconstruction.

Instead of using common spatial filters, session-specific spatial filters can be used, but then a common normalisation is applied. This method has been previously alluded to in Vrba et al.'s pseudo-T-statistics for comparing active versus control blocks within a single session of data (Vrba and Robinson, 2001). In that analysis, active and control states from a single session of data are beamformed separately, but then a common correction for the projected sensor noise is applied to the

* Corresponding author at: Oxford Centre for Human Brain Activity, University of Oxford,Warneford Hospital, Oxford, UK.
E-mail address: mark.woolrich@ohba.ox.ac.uk (M.W. Woolrich).

http://dx.doi.org/10.1016/j.neuroimage.2013.12.026
1053-8119

two states. Whilst Vrba et al.'s approach is valid, it has the potential to scale poorly to large numbers of sessions. As more sessions of data are compared in a given group-level analysis, applying a single group average correction will become increasingly inaccurate.

In this technical note, we consider the confound that weights normalisation introduces into group analyses performed on separately beamformed multiple sessions data. This includes analyses that infer group-level differences, for example between different populations (e.g. between clinical and normal populations), or between different conditions in resting-state or task-positive paradigms. To avoid this confound, we instead recommend that no weights normalisation is carried out. Any depth bias is instead accounted for in the calculation of multi-session (e.g. group) statistics, as these naturally contain variance normalisations which include the increased variance due to reconstructing activity in deeper brain regions.

To demonstrate the weights normalisation confound, we start by reviewing its theoretical justification. We show mathematically that weights normalisation provides an unbiased estimate of the noise contribution for a voxel at the cost of providing a biased estimate of the true dipole variance. We go on to show that this leads to an inaccurate estimate of the difference between variances of two sessions. We use a simple 3 dipole/2 sensor simulation to show an example case of how, under certain conditions, using the weights-normalised variance estimates to compare two sessions can lead to a reversal in the inferred direction of effect. Finally, we show that this issue can cause severe problems in the analysis of real MEG data. We compare the oscillatory alpha power in eyes-closed rest with a visual active-state (watching a movie). In this paradigm we expect increased alpha power in the *eyes-closed* condition compared with the *active-state* (Berger, 1929; Buzsáki, 2006). We show that the weights normalisation confound actually leads to an inferred *decrease* in alpha power during the *eyes-closed* condition compared with the *active-state*. However, with no weights normalisation the alpha power is correctly inferred as increasing in occipital areas.

Theory

The MEG forward problem

MEG data at the sensor-level can be modelled by using the quasi-static approximation to Maxwell's equations (Hämäläinen et al., 1993). The sensor data, $\mathbf{Y}$ (which is an $N \times P$ matrix, where N is the number of sensors and P is the number of samples), can be modelled as follows:

$$\mathbf{Y} = \sum_{q=1}^{Q} \mathbf{h}_q \mathbf{x}_q + \varepsilon \tag{1}$$

where $\mathbf{h}_q$ is the N by 1 lead-field vector that describes the flux measured at the sensor array due to a unit dipole of specific orientation at location q whose time course of activity is described by $\mathbf{x}_q$. ε is the N sensors by P samples matrix of uncorrelated sensor noise and Q is the total number of dipoles that model the activity across the brain. For the purpose of this technical note, we shall assume that the dipole orientation is known a priori. We can then reduce our source reconstruction problem from vectors to scalars without loss of generality when considering the issue of weights normalisation.

Using beamforming to infer underlying neural activity

Under the assumption of stationarity, beamformers estimate the underlying neural activity that drives each MEG recording by designing a spatial filter specific to each location. These spatial filters estimate the underlying activity as a weighted sum of the sensor-level MEG data, as shown below:

$$\hat{\mathbf{x}}_q = \mathbf{w}_q \mathbf{Y} \tag{2}$$

where $\mathbf{w}_q$ is the vector of beamformer weights for the q^{th} voxel and $\hat{\mathbf{x}}_q$ is the raw beamformer estimate of the time course of the current dipole at q.[1] The beamformer weights, $\mathbf{w}_q$, at the q^{th} voxel are estimated by imposing a minimisation of the overall projected power, subject to unit gain at the location of interest. The set of weights that satisfies these two constraints can be found by using Lagrange multipliers (Van Veen et al., 1997).

Beamformer weights normalisation

In this section we revisit the motivation for applying a spatial correction (weights normalisation) to each voxel. This has been previously described elsewhere (Van Veen et al., 1997; Vrba and Robinson, 2001). Consider the time course of activity of a dipole at the q'^{th} voxel. The variance, $\hat{\sigma}^2_{q'}$, of our beamformer estimate of the neural activity (i.e. the average source power over time) can be evaluated as:

$$\hat{\sigma}^2_{q'} = \frac{1}{P} \hat{\mathbf{x}}_{q'} \hat{\mathbf{x}}^T_{q'} \tag{3}$$

We estimate the neural activity at q' as a weighted sum of the sensor data (Eq. (2)), which in turn is a function of the true neural activity and the lead-fields (Eq. (1)). Substituting Eqs. (1) and (2) into Eq. (3), and separating out the terms corresponding to the q'^{th} dipole from terms corresponding to the $q \neq q'$ dipoles, we get:

$$\hat{\sigma}^2_{q'} = \frac{1}{P} \mathbf{w}_{q'} \left(\mathbf{h}_{q'} \mathbf{x}_{q'} + \sum_{q \neq q'} \mathbf{h}_q \mathbf{x}_q + \varepsilon \right) \left(\mathbf{h}_{q'} \mathbf{x}_{q'} + \sum_{q \neq q'} \mathbf{h}_q \mathbf{x}_q + \varepsilon \right)^T \mathbf{w}^T_{q'} \tag{4}$$

If we substitute the unity pass band beamformer constraint ($\mathbf{w}_{q'} \mathbf{h}_{q'} = 1$) and make the following assumptions:

1. All the dipoles are uncorrelated with each other. This is an implicit requirement for beamforming (Brookes et al., 2007; Van Veen et al., 1997).
2. All the dipoles are uncorrelated with the sensor-level noise, ε.

we can then express the expected value of the variance of the beamformer estimate as:

$$\begin{aligned} \hat{\sigma}^2_{q'} &= \frac{1}{P} \left[\mathbf{x}_{q'} \mathbf{x}^T_{q'} + \mathbf{w}_{q'} \left(\sum_{q \neq q'} \mathbf{h}_q \mathbf{x}_q \mathbf{x}^T_q \mathbf{h}^T_q \right) \mathbf{w}^T_{q'} + \mathbf{w}_{q'} \varepsilon \varepsilon^T \mathbf{w}^T_{q'} \right] \\ &= \overbrace{\sigma^2_{q'}}^{\text{true variance}} + \overbrace{\mathbf{w}_{q'} \left(\sum_{q \neq q'} \mathbf{h}_q \sigma^2_q \mathbf{h}^T_q \right) \mathbf{w}^T_{q'}}^{\text{signal leakage variance}} + \overbrace{\sigma^2_\varepsilon \mathbf{w}_{q'} \mathbf{w}^T_{q'}}^{\text{noise variance}} \end{aligned} \tag{5}$$

where we assume that the sensor-level noise covariance matrix can be approximated as the estimated noise variance, ${\sigma_\varepsilon}^2$, averaged across sensors multiplied by an N by N identity matrix.

The sensor noise term in Eq. (5) has a spatially-varying bias ($\mathbf{w}_{q'} \mathbf{w}^T_{q'}$) which leads to an overestimation of the variance in the centre of the brain (Hall et al., 2013). Commonly, beamforming implementations rescale the estimate of the dipole's time course by dividing it by some quantity D (Huang et al., 2004) in order to down-weight deeper voxels. We term this *weights normalisation*.

Different types of weights normalisation have been proposed. The one we use here is $D = \sqrt{\mathbf{w}_{q'} \mathbf{w}^T_{q'}}$ (Borgiotti and Kaplan, 1979; Huang et al., 2004; Sekihara et al., 2001). This will give an unbiased estimate of the noise variance across all voxels. An alternative weights normalisation is to incorporate an estimate of the actual noise covariance matrix (i.e. $D = \sqrt{\mathbf{w}_{q'} \Sigma \mathbf{w}^T_{q'}}$, where $\Sigma = \frac{1}{P} \varepsilon \varepsilon^T$ is the noise covariance matrix) (Hall et al., 2013; Robinson and Vrba, 1998). However, it should be

[1] Notation: in the following section and subsequently through this technical note, the following notation will be adopted: The true underlying value of a variable has no accent: e.g. $\mathbf{x}$. The raw beamformer estimate of a variable is denoted with a circumflex: e.g. $\hat{\mathbf{x}}$. The weights-normalised beamformer estimate of a variable is denoted with a tilde: e.g. $\tilde{\mathbf{x}}$.

noted that any normalisation approach, where the normalisation depends on the weights, will result in a bias in the estimate of the true variance.

Without weights normalisation, we observe an increase in the variance of the estimated neural activity with depth even if the true neural activity has homogeneous variance (Van Veen et al., 1997). Whilst weights normalisation corrects for this, it also creates a biased estimate of the true variance. This means that the weights-normalised beamformer estimate of a dipole's variance, $\tilde{\sigma}_{q'}{}^2$, is no longer equal to the true variance, $\sigma^2_{q'}$, plus some error contributions due to signal leakage and sensor noise. Instead, the beamformer estimate is equal to the true variance scaled by the weights normalisation. This is shown below for the case of $D = \sqrt{\mathbf{w}_{q'}\mathbf{w}_{q'}^T}$:

$$\tilde{\sigma}^2_{q'} = \frac{\hat{\sigma}^2_{q'}}{\mathbf{w}_{q'}\mathbf{w}_{q'}^T} \tag{6}$$

$$= \frac{\sigma^2_{q'}}{\mathbf{w}_{q'}\mathbf{w}_{q'}^T} + \mathbf{w}_{q'}\frac{\left(\sum_{q\neq q'}\mathbf{h}_q\sigma^2_q\mathbf{h}_q^T\right)\mathbf{w}_{q'}^T}{\mathbf{w}_{q'}\mathbf{w}_{q'}^T} + \sigma^2_\varepsilon \tag{7}$$

Deriving the weights normalisation confound

In the previous section we derived an expression for the raw (non-weights-normalised) beamformer estimate of the variance, $\hat{\sigma}^2_{q'}$, at a voxel, q', using weights $\mathbf{w}_{q'}$, in terms of the true variance, $\sigma_{q'}{}^2$, plus contributions due to signal leakage and the projection of sensor noise (Eq. (5)). We derived an equivalent expression for the weights-normalised estimate of the variance, $\tilde{\sigma}_{q'}{}^2$ (Eq. (6)). If we consider any analysis that attempts to compare variances, or any quantities that scale with variance, across multiple beamformer sessions, we can demonstrate that we should use the non-weights-normalised estimates of the variance, $\hat{\sigma}_{q'}{}^2$, to get an unbiased estimate of the difference in variances. Consider two sessions of separately beamformed data, *1* and *2*, between which we intend to test for differences in variance (e.g. to see if we can detect a change in oscillatory power in a particular frequency band). At the q'^{th} voxel, we have *non-weights-normalised* estimates $\hat{\sigma}^2_{q',1}$ and $\hat{\sigma}^2_{q',2}$. The difference between these estimates is:

$$\hat{\sigma}^2_{q',1} - \hat{\sigma}^2_{q',2} = \sigma_{q',1}{}^2 - \sigma_{q',2}{}^2 + e \tag{8}$$

where e represents the additional error terms (i.e. the contributions to our variance estimate due to signal leakage and projected sensor noise, defined in Eq. (5)). We assume that these contributions are approximately equal across *sessions 1* and *2*. This assumption requires that the noise variance of the scanner is constant across the two sessions. If we perform our statistics on the weights-normalised variances, $\tilde{\sigma}^2_{q',1}$ and $\tilde{\sigma}^2_{q',2}$, Eq. (6) tells us that we get a biased estimate of the difference in variance due to the differences between the weights normalisation terms in the different sessions:

$$\tilde{\sigma}^2_{q',1} - \tilde{\sigma}^2_{q',2} = \frac{\sigma_{q',1}{}^2}{\mathbf{w}_{q',1}\mathbf{w}_{q',1}^T} - \frac{\sigma_{q',2}{}^2}{\mathbf{w}_{q',2}\mathbf{w}_{q',2}^T} + \tilde{e} \tag{9}$$

where $\tilde{e}$ represents the additional error terms (due to signal leakage and the projection of sensor noise) with their respective session-specific weights normalisations applied. We term this bias in the estimate the *weights normalisation confound*.

Simulating the weights normalisation confound

Critically, we can show that under certain conditions this confound can actually cause the apparent direction of an effect to be reversed. We demonstrated this with a very simple simulation involving three dipoles (*A*, *B* and *C*)in a 2-sensor MEG system.[2] Fig. 1A is a schematic showing the approximate geometry of this simulation. Fig. 1B shows the 2D lead-field vectors (constant across *sessions 1* and *2*) for the three dipoles. For simplicity, we made all the lead-field vectors unit magnitude (although this does not affect the generality of our argument). We generated three orthogonal, normally distributed time courses for the three dipoles for each session. Fig. 1C shows the standard deviations of the dipoles. In our simulation, *dipole A* shows a small reduction in standard deviation between *sessions 1* and *2*, *dipole B* shows a large reduction, and *dipole C* remains unchanged. We projected both sessions of data through our simulated lead-fields and then separately beamformed each session. Fig. 1D shows the difference in standard deviation between *session 1* and *2* for *dipole A* for the ground truth, the non-weights-normalised estimate, and the weights-normalised estimate. Note that the weights-normalised estimate of the standard deviation in *session 2* is bigger not smaller than in *session 1* (i.e. the apparent effect direction has reversed).

We can see why by considering the weights vectors, $\mathbf{w}_{A,1}$ and $\mathbf{w}_{A,2}$, from *sessions 1* and *2* for *dipole A*, plotted as dashed cyan and magenta vectors alongside the lead-field vectors in Fig. 1B. The unity pass band constraint (shown in Fig. 1B as a grey dotted line) limits the weights vectors to a specific subspace. In *session 1*, *dipole B* has the greatest variance. The beamformer tries to minimise the projected variance and so finds the optimum weights to be mostly orthogonal to the lead-fields of *dipole B*. In *session 2*, *dipole B*'s variance greatly reduces and the beamformer adapts by finding a set of weights that are now more orthogonal to the lead-fields of *dipole C*, whose projected variance is relatively greater in *session 2*. This has the unintended consequence of changing the Euclidean length of the weights vectors by a greater proportion than the change in standard deviation of *dipole A*. As such, the weights-normalised estimate of the change in standard deviation for *dipole A* actually reverses sign.

A real data example

This change in the valence of the estimated difference in standard deviation or variance between sessions is not limited to simulations but can be observed in real MEG data. Here, we show such an example. Specifically, we show that when attempting to detect the well-documented increase in oscillatory alpha power in eyes-closed rest compared with continuous visual stimulation, the weights normalisation confound actually leads to an inferred decrease in oscillatory alpha power in the *eyes-closed* condition (Berger, 1929; Buzsáki, 2006).

Data

Participants

Ten healthy volunteers were recruited. The cohort comprised 7 males (all right-handed) and 3 females (2 right-handed) with a mean age of 27 years and standard error of 0.48 years. The study was approved by the University of Nottingham Medical School Research Ethics Committee.

Data acquisition

Each participant underwent a scan which included a 10-min block where the participant was scanned at rest with their eyes closed and instructed not to fall asleep, and a 10-min block where the participant watched a movie, which was projected through a waveguide in the magnetically shielded room onto a screen placed 40 cm in front of the subject. Each participant was scanned in a supine position.

MEG data were acquired by using a CTF 275 channel whole-head system. The data were sampled at 600 Hz and synthetic third order

[2] The Matlab script for performing this simulation can be downloaded from www.fmrib.ox.ac.uk/~woolrich/weights_normalisation_paper_simulations.m.

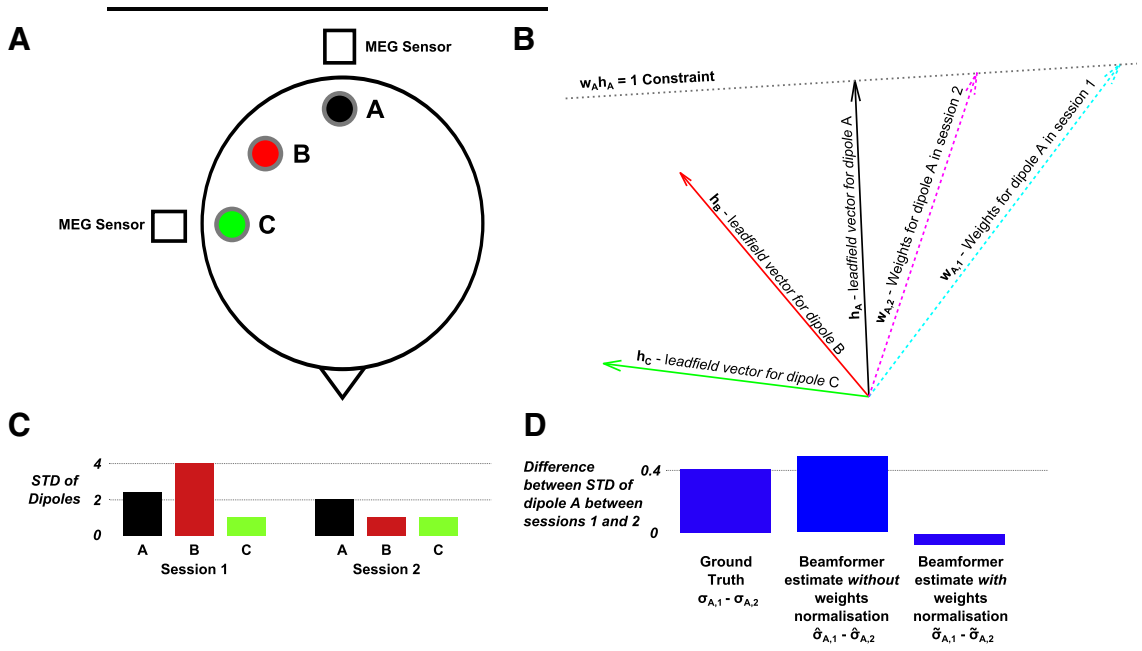

Fig. 1. A simulation showing one example of how weights normalisation can reverse the apparent direction of an effect. In this simulation, the effect we measure is the standard deviation of *dipole A*, σ_A, between two sessions (1 & 2) where $\sigma_{A,1} > \sigma_{A,2}$. A. The schematic of our 2-dimensional simulation, consisting of 3 dipoles and 2 MEG sensors. The dipole time courses are uncorrelated and normally distributed. B. The lead-field vectors for our three dipoles (solid black, red and green arrows) and the beamformer weights vectors for *dipole A* from *session 1* and *session 2* (dashed cyan and magenta arrows). C. The standard deviations of the three dipoles in *sessions 1* and *2*. D. The ground truth and beamformer estimated differences (without and with weights normalisation) between the standard deviation of *dipole A* in *sessions 1* and *2*.

synthetic gradiometer correction was applied to reduce external interference. Head localisation within the MEG helmet was achieved by using three electromagnetic head position indicator (HPI) coils. Prior to data acquisition, the HPI coil locations and the subject's head shape were digitised by using a Polhemus Isotrack system. Structural MR images for each subject were acquired by using a Philips Achieva 3T MRI system (MPRAGE; 1 mm isotropic resolution, 256 × 256 × 160 matrix, TR = 8.1 ms, TE = 3.7 ms, TI = 960 ms, shot interval = 3 s, flip angle = 8° and SENSE factor 2). Two subjects were discarded due to problems with co-registration.

Methods

Each session of data was analysed in the following way. Independent component analysis (ICA) was used to decompose the data into 150 temporally independent time courses and associated sensor topographies. Artefact components corresponding to eye-blink, cardiac and mains interference were manually identified by the combined inspection of the spatial topography, time course, kurtosis of the time course and frequency spectrum for all components. Eye-blink artefacts typically exhibited high kurtosis (>20), a repeated blink structure in the time course and very structured spatial topographies. Cardiac component time courses strongly resembled the typical ECG signals, as well as having high kurtosis (>20). Mains interference had extremely low kurtosis (typically <−1) and a frequency spectrum dominated by 50 Hz line noise. These artefact components were subtracted from the sensor space data using the montage function in SPM8 (FIL, UCL) in order to correct the lead fields at the same time as denoising the data. High variance channels were then identified and discarded. Any periods of corrupted data (for example by transient muscle activity) were visually identified and flagged. These "bad epochs" were not discarded but were excluded from specific stages of the analysis (such as the estimation of the data covariance matrix). The "bad epochs" were included in stages where data continuity was required (such as band-pass filtering).

The data were then band-pass filtered into the alpha band (8–13 Hz). The data covariance matrix for each session was then estimated by using the band-limited data. A linearly-constrained minimum variance beamformer was used to estimate the source space neural activity at every vertex of 6 mm grid spanning the whole brain. Dipole orientations were estimated as those that projected the maximum power (Sekihara et al., 2001). Raw beamformer time courses were estimated at each voxel by multiplying the beamformer weights with the band-limited data. At each voxel, the oscillatory amplitude envelope of the band-limited neural activity was calculated by taking the absolute of the *analytic signal*, computed via the Hilbert transform. The envelopes were estimated both on the raw beamformer time courses and the weights-normalised beamformer time courses. It should be noted that the beamformer source reconstruction includes a dipole sign ambiguity. However, as we analysed the amplitude envelopes of the beamformed data, this ambiguity is inconsequential. For each envelope time course, the mean and variance were computed as measures of the mean of oscillatory alpha power and the variability of oscillatory alpha power associated with that voxel.

We applied a paired *t*-test between *eyes-closed* and *activate-state* sessions to the mean and variance of the envelope at each voxel over all subjects. We did this test by using both the non-weights-normalised and weights-normalised envelopes to demonstrate the confound that *weights normalisation* introduces. Voxel-wise multiple comparisons in the statistical maps were accounted for using threshold-free cluster enhanced (TFCE) permutation testing in FSL (FMRIB, Oxford).

Results

The corrected t-statistical maps are shown in Fig. 2. All super-threshold voxels survived TFCE correction. Without weights normalisation, we found a significant ($p_{corrected} \leq 0.05$) increase in the mean and variance of the alpha band envelope in the visual cortex. However, when weights normalisation was applied, we observed the reverse: a significant *decrease* in alpha power in *eyes-closed* sessions relative to

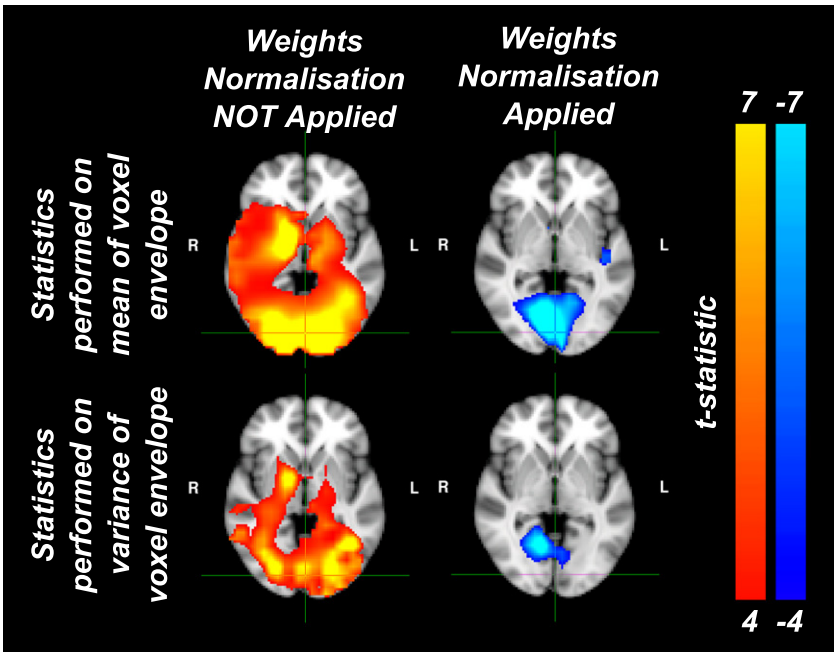

Fig. 2. Axial slices through the visual cortex showing results of a voxel-wise paired *t*-test performed on the mean and variance of the envelope of the alpha (8–13 Hz) oscillations of the *eyes-closed* and *active-state* sessions before and after applying weights-normalisation. For each map, t-statistics were thresholded at ± 4, with positive values shown in red/yellow and negative values shown in blue. Without weights normalisation, we correctly infer an increase in the mean and variance of alpha power in the *eyes-closed* condition compared with the *active-state* condition. With weights normalisation, we incorrectly infer a decrease in the mean and variance of alpha power in the eyes-closed condition compared with the *active-state* condition, demonstrating that the weights normalisation confound can be so severe as to actually reverse the underlying effect direction. FSL's RANDOMISE was used to perform threshold-free cluster enhanced (TFCE) permutation testing to account for multiple comparisons. All t-statistics shown are members of significant ($P_{corrected} < 0.05$) clusters.

the *active-state* sessions. In addition, the spatial location of this power decrease was deeper and more medial.

Discussion

The critical message of this technical note is that beamformer weights normalisation can completely distort the estimate of variance, or any metric that requires an unbiased estimate of variance. Conceptually, the weights normalisation causes the estimated variance to be dependent upon the beamformer weights, and the weights can be different for different scanning sessions (for example, due to there being different distributions and magnitudes of activity across the wider brain) even if the power of the brain activity in the location of interest is the same. We have demonstrated this effect in three ways: theoretically, with a simulation, and with a practical demonstration of the effect in real data.

We have shown an extreme case of this confound, where the apparent sign of the effect we are trying to measure can actually be reversed. A key question is whether this sign reversal is always a consequence of weights normalisation or whether our examples presented here represent a worst case. When setting up our simulation, we found that the reversal of inferred effect direction depended heavily on the specific relative changes in variance of the dipoles and their lead-field orientations. As such, the reversal of effect direction is not an automatic consequence of weights normalisation. In other circumstances, the weights normalisation confound may simply introduce an additional source of error that degrades the statistical significance of a given effect. Furthermore, in other cases it is possible that performing statistics on weights normalised estimates will have little effect on the group-level statistics.

Considering our simulation in more detail, the beamformer tries to minimise the projected variance and so finds the optimum weights to be mostly orthogonal to the lead-fields of *dipole B*. In *session 2*, *dipole B*'s variance greatly reduces and the beamformer adapts by finding a set of weights that are now more orthogonal to the lead-fields of *dipole C*, whose projected variance is relatively greater in *session 2*. This has the unintended consequence of changing the Euclidean length of the weights vectors by a greater proportion than the change in standard deviation of *dipole A*.

In this paper, we demonstrated the sign reversal in both simulations and real data. However, we should ask whether both these results represent the same underlying issue. The mechanism by which we generated a change in effect valence in our simulations could plausibly be occurring in our *eyes-closed*/*active-state* analysis. If we consider the medial visual cortex where we observed a reversal in the direction of alpha power change, it is possible that this region experienced a small increase in alpha power during the *eyes-closed* rest compared with the *active-state*. However, other visual areas may have experienced much larger alpha power increases in the *eyes-closed* condition compared with the active-state. Furthermore, it is likely that other brain areas showed no difference in alpha power between the two conditions. As such, the conditions that lead to a reversal in effect direction in our simulation could quite plausibly be occurring in our *eyes-closed*/*active-state* analysis (be it in 275 dimensions rather than 2).

In our theoretical exposition, we made the assumption that contributions to the beamformer estimate of variance due to signal leakage and sensor noise were approximately constant across multiple sessions. In our data set it is reasonable to expect that the noise is constant between *eyes-closed* and *active-state* sessions as these sessions were acquired during a single scan for each subject. However, the contributions due to signal leakage depend on the activity of all the (neuronal and artefactual) sources. As such, this contribution will vary between different conditions. Whilst this is an important issue to note, it is not relevant to the key point of this paper.

Finally, this technical note has exclusively considered the weights normalisation confound in the context of beamforming. However, other source reconstruction methods include data-dependant normalisation terms, such as minimum norm estimates with corrections for superficiality bias (Hauk et al., 2011). It will be very important to assess whether these alternatives to beamforming are also exposed to equivalent confounds and whether the effects can be as drastic as we have demonstrated here.

Conclusion

Beamforming is a powerful tool for performing source reconstruction in MEG. However, most beamformer implementations only consider single session analyses and, as such, are not designed with group-level analyses in mind. In this technical note, we have shown that weights normalisation should not be used when performing multi-session statistics. However, it is a necessary analysis stage for other analyses, such as resting-state ICAs where the depth bias of raw data must be removed (Brookes et al., 2011). Consequently, much care must be taken to decide when weights normalisation should and should not be applied in multi-session MEG analyses.

Acknowledgments

The authors would like to thank Sofia Palazzo Corner for her assistance in acquiring the MEG data used, the Centre for Doctoral Training in Healthcare Innovation and the RCUK Digital Economy Programme for funding Dr Henry Luckhoo, the Leverhulme Trust for fellowship support for Dr Matthew Brookes, the University of Nottingham who funded the CTF MEG scanner. Dr Mark Woolrich was supported by the Wellcome Trust (092753), the MRC/EPSRC UK MEG Partnership award, and by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre based at Oxford University Hospitals Trust Oxford University.

Conflict of interest

There are no conflicts of interest.

References

Berger, H., 1929. Über das Elektrenkephalogramm des Menschen. Arch. Psychiatr. Nervenkr. 87, 527–570 (URL: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Uber+das+Elektrenkephalogramm+des+Menschen.#0).

Borgiotti, G., Kaplan, L., 1979. Super resolution of uncorrelated interference sources by using adaptive array techniques. IEEE Trans. Antennas Propag. AP-27, 842–845 (URL: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1142176).

Brookes, M.J., Stevenson, C.M., Barnes, G.R., Hillebrand, A., Simpson, M.I.G., Francis, S.T., Morris, P.G., 2007. Beamformer reconstruction of correlated sources using a modified source model. NeuroImage 34, 1454–1465. http://dx.doi.org/10.1016/j.neuroimage.2006.11.012 (URL: http://www.ncbi.nlm.nih.gov/pubmed/17196835).

Brookes, M.J., Woolrich, M.W., Luckhoo, H., Price, D., Hale, J.R., Stephenson, M.C., Barnes, G.R., Smith, S.M., Morris, P.G., 2011. Investigating the electrophysiological basis of resting state networks using magnetoencephalography. Proc. Natl. Acad. Sci. U. S. A. 108, 16783–16788. http://dx.doi.org/10.1073/pnas.1112685108 (URL: http://www.ncbi.nlm.nih.gov/pubmed/21930901).

Buzsáki, G., 2006. Rhythms of the Brain, 1st ed. Oxford University Press, Oxford.

Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G., Mattout, J., 2008. Multiple sparse priors for the M/EEG inverse problem. NeuroImage 39, 1104–1120. http://dx.doi.org/10.1016/j.neuroimage.2007.09.048 (URL: http://www.ncbi.nlm.nih.gov/pubmed/17997111).

Hall, E.L., Woolrich, M.W., Thomaz, C.E., Morris, P.G., Brookes, M.J., 2013. Using variance information in magnetoencephalography measures of functional connectivity. NeuroImage 67, 203–212. http://dx.doi.org/10.1016/j.neuroimage.2012.11.011 (URL: http://www.ncbi.nlm.nih.gov/pubmed/23165323).

Hämäläinen, M., Ilmoniemi, R.J., 1984. Interpreting measured magnetic fields of the brain: minimum norm estimates of current distributions. Med. Biol. Eng. Comput. 32, 35–42.

Hämäläinen, M., Hari, R., Ilmoniemi, R.J., 1993. Magnetoencephalography — theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev. Mod. Phys. 65, 413–497.

Hauk, O., Wakeman, D.G., Henson, R., 2011. Comparison of noise-normalized minimum norm estimates for MEG analysis using multiple resolution metrics. NeuroImage 54, 1966–1974. http://dx.doi.org/10.1016/j.neuroimage.2010.09.053 (URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3018574&tool=pmcentrez&rendertype=abstract).

Huang, M.X., Shih, J.J., Lee, R.R., Harrington, D.L., Thoma, R.J., Weisend, M.P., Hanlon, F., Paulson, K.M., Li, T., Martin, K., Millers, G.A., Canive, J.M., 2004. Commonalities and differences among vectorized beamformers in electromagnetic source imaging. Brain Topogr. 16, 139–158 (URL: http://www.ncbi.nlm.nih.gov/pubmed/15162912).

Robinson, S.E., Vrba, J., 1998. Functional neuroimaging by synthetic aperture magnetometry (SAM). Recent Advances in Biomagnetism, pp. 302–305.

Sekihara, K., Nagarajan, S.S., Poeppel, D., Marantz, A., Miyashita, Y., 2001. Reconstructing spatio-temporal activities of neural sources using an MEG vector beamformer technique. IEEE Trans. Biomed. Eng. 48, 760–771. http://dx.doi.org/10.1109/10.930901 (URL: http://www.ncbi.nlm.nih.gov/pubmed/11442288).

Van Veen, B.D., van Drongelen, W., Yuchtman, M., Suzuki, A., 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans. Biomed. Eng. 44, 867–880. http://dx.doi.org/10.1109/10.623056 (URL: http://www.ncbi.nlm.nih.gov/pubmed/9282479).

Vrba, J., Robinson, S.E., 2001. Signal processing in magnetoencephalography. Methods 25, 249–271. http://dx.doi.org/10.1006/meth.2001.1238 (URL: http://www.ncbi.nlm.nih.gov/pubmed/11812209).