Submitted to the Annals of Applied Statistics
arXiv: arXiv:0000.0000

SUPPLEMENTARY MATERIALS: QUANTIFYING TIME-VARYING SOURCES IN MAGNETOENCEPHALOGRAPHY – A DISCRETE APPROACH

By Zhigang YAO*, Zengyan FAN*, Masahito Hayashi† and William F. Eddy‡

*National University of Singapore**, *Nagoya University*† *and Carnegie Mellon University* ‡

We provide a brief illustration of the switch procedure and dynamic procedure, as well as numerical results for the simulated data and real data sets from MEG/EEG recordings.

1. Illustration of the Switch Procedure. In Section 2.4 of Yao et al. (2018a), the switch procedure was proposed to reduce the computational complexity of calculating the discrete posterior distribution when multiple sources presented. In Figure 1, we show the number of possible states that we need to include for calculating the discrete posterior distribution, where we choose equal mesh grids for each dimension of the ROI. As seen in Figure 1, the number of possible states increases dramatically when the one-dimensional mesh gird K_i increases.

2. Illustration of the Dynamic Procedure. In Section 2.5 of Yao et al. (2018a), the dynamic procedure was developed when no information on the ROI was available. Figure 2 shows the dynamic procedure in one dimension. In each iteration of the EM algorithm, the intermediate posterior distribution was calculated using the current ROI (the area between the black dash lines in Figure 2). After which, we obtained a shrunken ROI (the area between the blue dash lines in Figure 2) from the intermediate posterior distribution. When the EM algorithm converged, a shrunken ROI would be available.

3. Numerical Results for Simulated Case 2. In this section, we compare the simulation performance in Section 3.3 of Yao et al. (2018a) with the result from a non-dynamic non-switch procedure. Due to the computational complexity, we were only able to implement the non-dynamic non-switch procedure with mesh grids $K_i = 5$, $i = 1, 2, 3$. Thus, the discrete posterior distribution was calculated over 15625 possible states during the

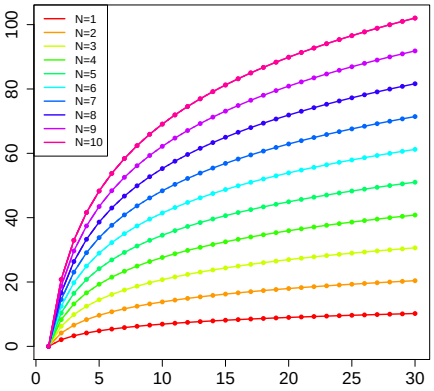


FIG 1. *Comparison of possible states for the case with N sources. The horizontal axis represents the one-dimensional mesh grid and the vertical axis represents the logarithm of possible states for N sources, where $1 \leq N \leq 10$.*

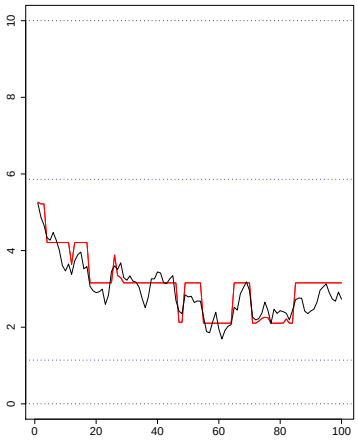
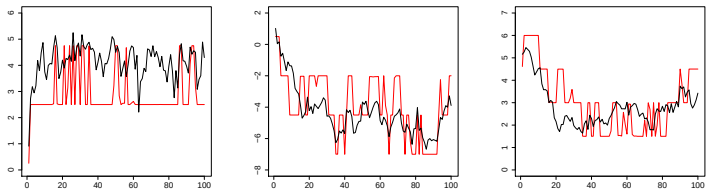
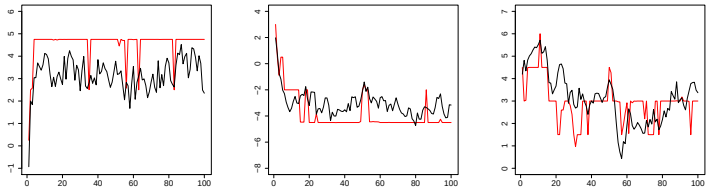


FIG 2. *Illustration of the dynamic procedure in one dimension. Black line: true trajectory of the source; black dash line: original one-dimensional ROI; red line: intermediate posterior means of the location parameter; blue dash line: shrunken one-dimensional ROI.*



(a) x dimension (b) y dimension (c) z dimension



(d) x dimension (e) y dimension (f) z dimension

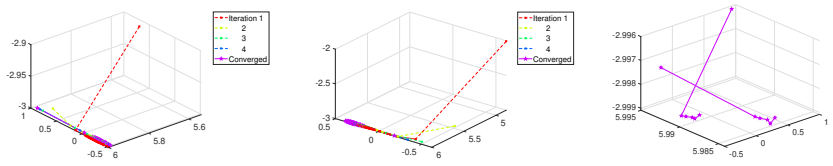
FIG 3. *Marginal posterior means for location parameter* $\mathbf{p}_t = (p_{t1}, p_{t2}, p_{t3})^{\mathrm{T}}$ *of two sources with the non-dynamic non-switch procedure in* 100 *time points. Top row: results for source* 1*; bottom row: results for source* 2*. The simulated location parameters are plotted in a black line, and the estimated posterior means are plotted in a red line.*

non-switch EM iterations. Figure 3 shows the posterior means of the sources obtained from the non-dynamic non-switch procedure. As we can see in Figure 3, the posterior means over some of the time points are quite flat, and did not capture the variation of the location parameter well.

4. Numerical Results from the BCI Data.

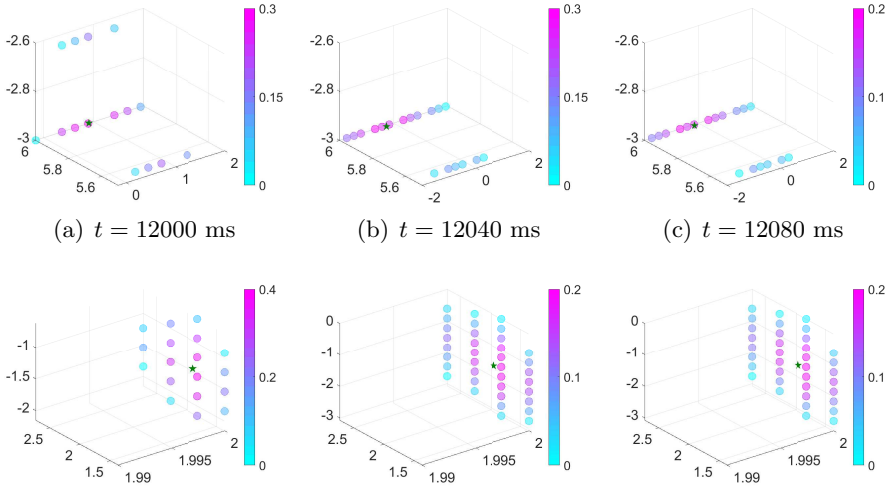
4.1. *Numerical Results from Activity for Short Time Frame.* In this section, we present the numerical results from the short frame analysis in the time window $[20000, 20099]$ ms, where two and three sources are investigated respectively. Figure 4 to Figure 6 show the results for the case with two sources, followed by the results for the three sources in Figure 7 to Figure 9.

To visualize the dynamics of the source distribution, we plot the trajectory of posterior means for the first sub-analysis in the BCI data with the two selected short time windows $[12000, 12099]$ ms and $[20000, 20099]$ ms, see Figure 10 to Figure 13.



(a) (b) (c)

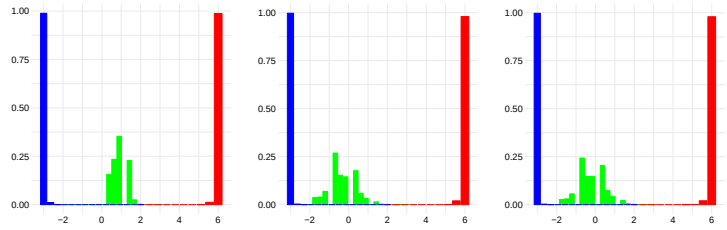
FIG 4. *Trajectory of posterior means for location parameter* (x, y, z) *of two sources in the time window* $[20000, 20099]$ *ms. (a) and (b) Trajectories of the two sources during EM iterations. (c) Trajectories of the two converged sources are highlighted at six selected time points* 12000, 12020, 12040, 12060, 12080, *and* 12099 *ms.*



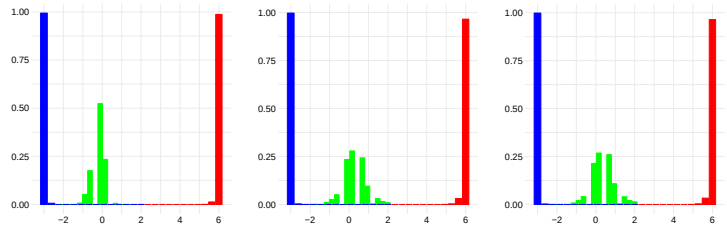
(a) $t = 12000$ ms (b) $t = 12040$ ms (c) $t = 12080$ ms

(d) $t = 12000$ ms (e) $t = 12040$ ms (f) $t = 12080$ ms

FIG 5. *Posterior distribution for location parameter* (x, y, z) *of two sources in the time window* $[20000, 20099]$ *ms. Top row: results for source* 1*; bottom row: results for source* 2*. Green star: posterior mean for location parameter at the selected time point.*

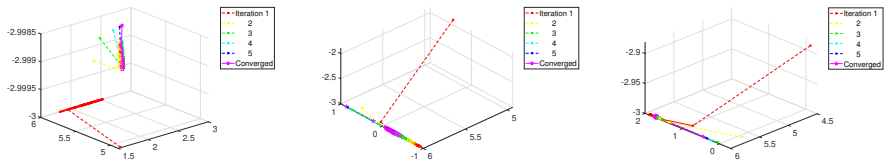


(a) $t = 12000$ ms (b) $t = 12040$ ms (c) $t = 12080$ ms

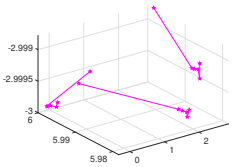


(d) $t = 12000$ ms (e) $t = 12040$ ms (f) $t = 12080$ ms

FIG 6. *Marginal posterior distribution for location parameter* (x, y, z) *of two sources in the time window* $[20000, 20099]$ *ms. Top row: results for source* 1*; bottom row: results for source* 2*. Green bar: marginal posterior distribution for parameter* x*; red bar: marginal posterior distribution for parameter* y*; blue bar: marginal posterior distribution for parameter* z*.*

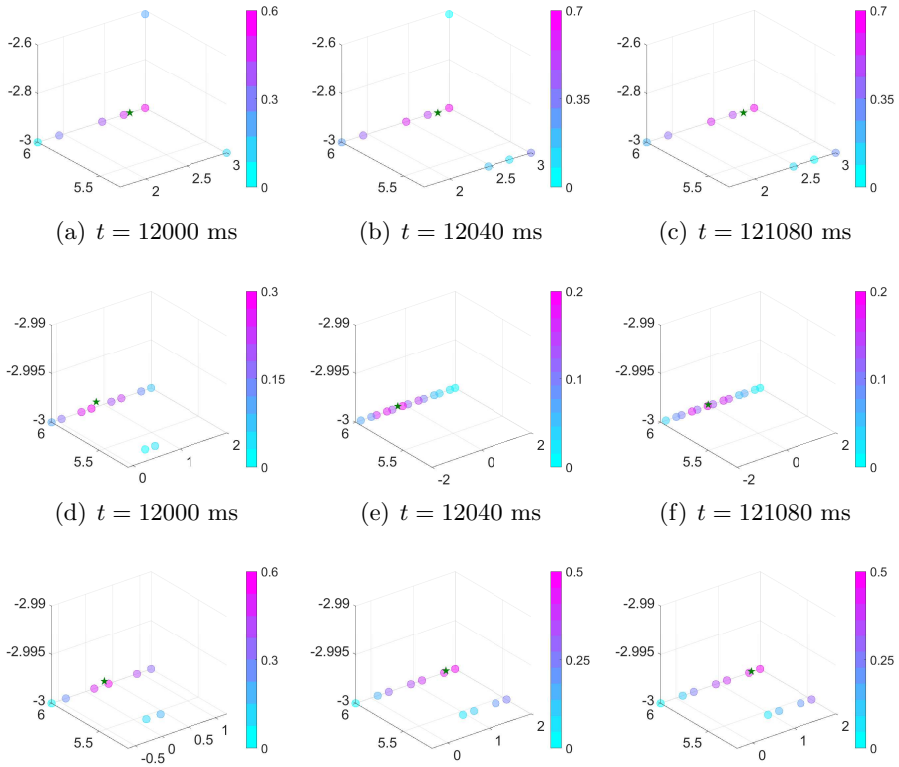


(a) (b) (c)



(d)

FIG 7. *Trajectory of posterior means for location parameter* (x, y, z) *of three sources in the time window* $[20000, 20099]$ *ms. (a), (b) and (c) Trajectories of the three sources during EM iterations. (d) Trajectories of the three converged sources are highlighted at six selected time points* 12000*,* 12020*,* 12040*,* 12060*,* 12080*, and* 12099 *ms.*

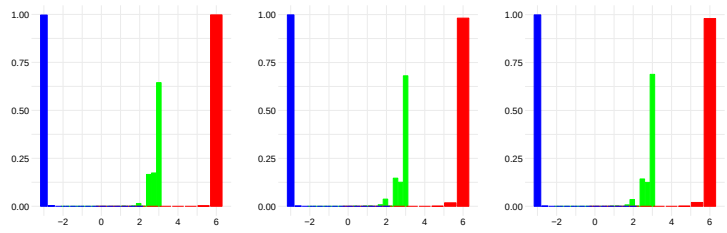


(a) $t = 12000$ ms (b) $t = 12040$ ms (c) $t = 121080$ ms

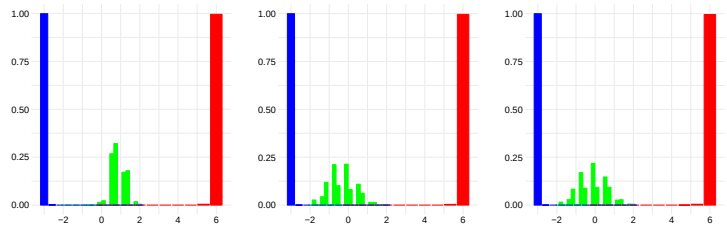
(d) $t = 12000$ ms (e) $t = 12040$ ms (f) $t = 121080$ ms

(g) $t = 12000$ ms (h) $t = 12040$ ms (i) $t = 121080$ ms

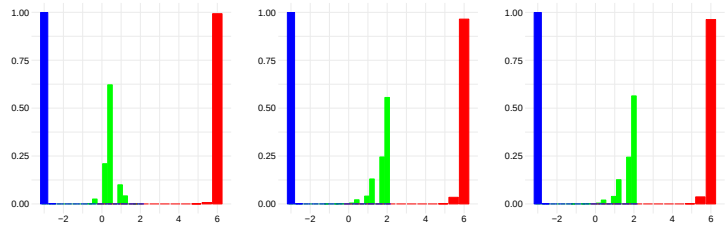
FIG 8. *Posterior distribution for location parameter* (x, y, z) *of three sources in the time window* $[20000, 20099]$ *ms. Top row: results for source* 1*; middle row: results for source* 2*; bottom row: results for source* 3*. Green star: posterior mean for location parameter at the selected time point.*



(a) $t = 12000$ ms (b) $t = 12040$ ms (c) $t = 12080$ ms

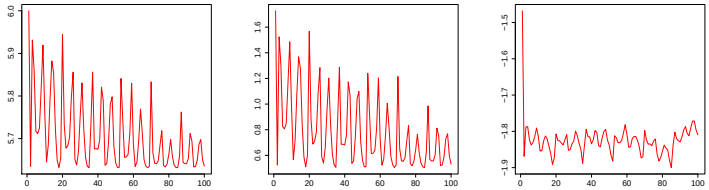


(d) $t = 12000$ ms (e) $t = 12040$ ms (f) $t = 12080$ ms



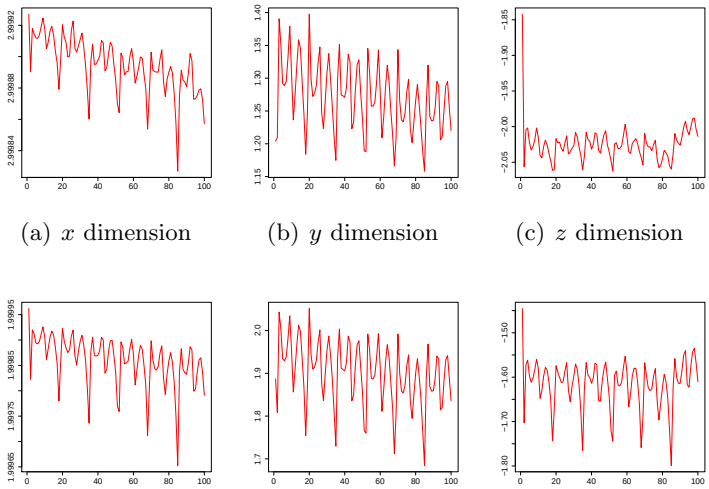
(g) $t = 12000$ ms (h) $t = 12040$ ms (i) $t = 12080$ ms

FIG 9. *Marginal posterior distribution for source location parameter* (x, y, z) *of three sources in the time window* $[20000, 20099]$ *ms. Top row: results for source* 1*; middle row: results for source* 2*; bottom row: results for source* 3*. Green bar: marginal posterior distribution for parameter* x*; red bar: marginal posterior distribution for parameter* y*; blue bar: marginal posterior distribution for parameter* z*.*



(a) x dimension (b) y dimension (c) z dimension

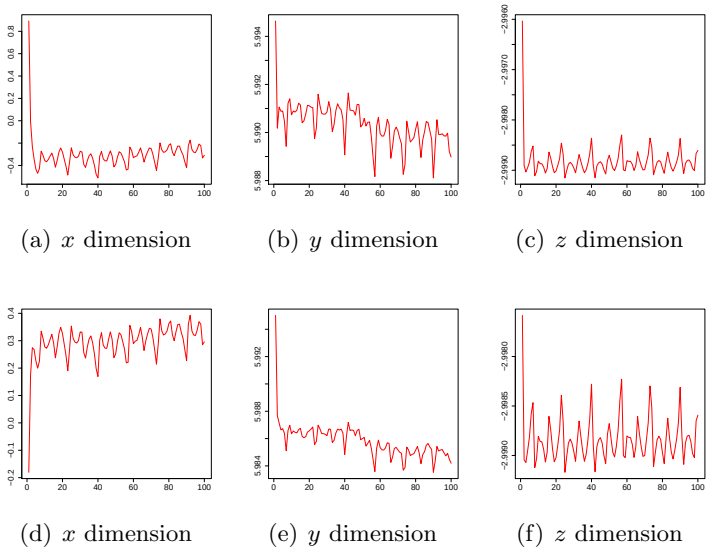
FIG 10. *Marginal posterior means for location parameter* (x, y, z) *of a single source in the time window* $[12000, 12099]$ *ms with ROI* $[-1, 6]$ *cm* $\times[0.5, 5]$ *cm* $\times[-2.5, 2]$ *cm.*



(a) x dimension (b) y dimension (c) z dimension

(d) x dimension (e) y dimension (f) z dimension

FIG 11. *Marginal posterior means for location parameter* (x, y, z) *of two sources in the time window* $[12000, 12099]$ *ms with ROIs* $[-2, 3]$ *cm* $\times[0, 5]$ *cm* $\times[-4, 0]$ *cm and* $[-3, 2]$ *cm* $\times[0, 5]$ *cm* $\times[-4, 0]$ *cm.*



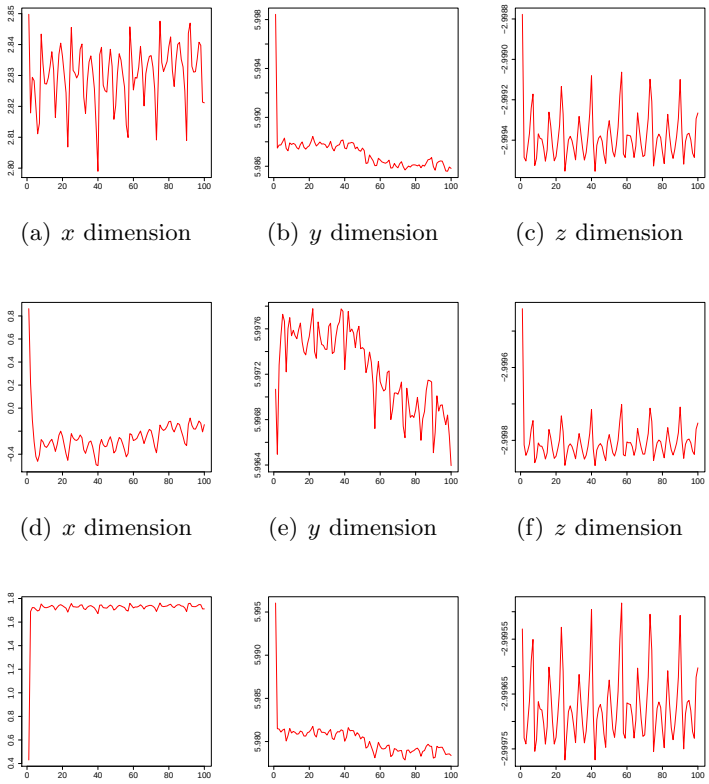
(a) x dimension (b) y dimension (c) z dimension

(d) x dimension (e) y dimension (f) z dimension

FIG 12. *Marginal posterior means for location parameter* (x, y, z) *of two sources in the time window* $[20000, 20099]$ *ms with ROIs* $[-2, 3]$ *cm* $\times[2, 6]$ *cm* $\times[-3, 2]$ *cm and* $[-3, 2]$ *cm* $\times[2, 6]$ *cm* $\times[-3, 2]$ *cm.*

4.2. *Numerical Results from Activity for Long Time Frame.* In Section 4.2 of Yao et al. (2018a), we also investigated the source distribution from the activity for long time frame in the BCI data. Figure 14 and Figure 15 show the results from the first time window $[12000, 12k^*99]$ ms, $k^* = 0, 1, \ldots, 4$. In a parallel manner, the plot of marginal distribution for source 2 and the trajectory of posterior means for two sources from the second time window $[20000, 20k^*99]$ ms, $k^* = 0, 1, \ldots, 4$, are illustrated in Figure 16 to Figure 18.

5. Illustration of the EEG Recording. As described in the Section 5 of Yao et al. (2018a), a set of EEG recordings was investigated for source localization. For each trial, the retention phase lasts for 4192 ms, as shown in Figure 19. The first 200 ms, the time window $[-200, 0]$ ms before the vertical line in Figure 19, serves as a baseline duration. To do the EEG source analysis for the period $[0, 3992]$ ms after event onset, the number of estimated sources from Yao et al. (2018b) is demonstrated in Figure 20.

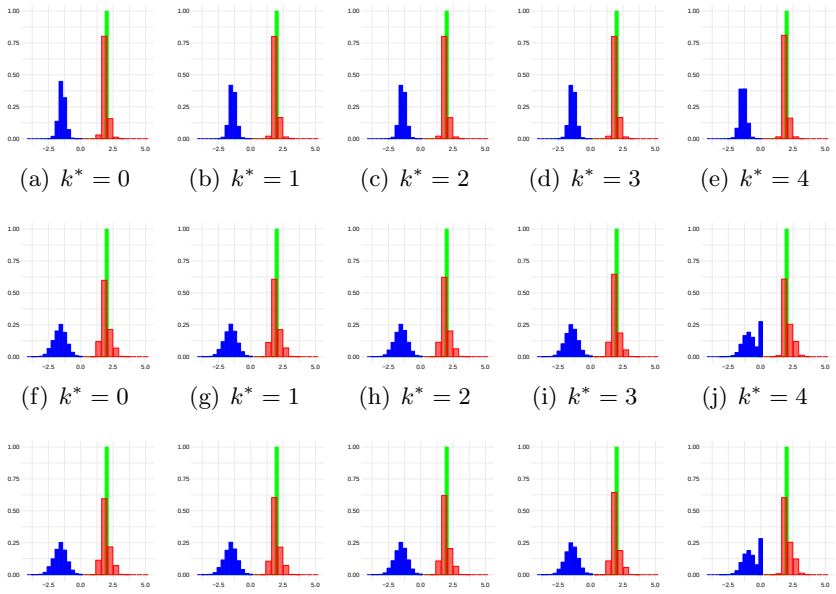


(a) x dimension (b) y dimension (c) z dimension

(d) x dimension (e) y dimension (f) z dimension

(g) x dimension (h) y dimension (i) z dimension

FIG 13. *Marginal posterior means for location parameter* (x, y, z) *of three sources in time window* $[20000, 20099]$ *ms with ROIs* $[-2, 3]$ *cm* $\times[0, 6]$ *cm* $\times[-3, 2]$ *cm,* $[-2, 2]$ *cm* $\times[0, 6]$ *cm* $\times[-3, 2]$ *cm and* $[-3, 2]$ *cm* $\times[0, 6]$ *cm* $\times[-3, 2]$ *cm.*

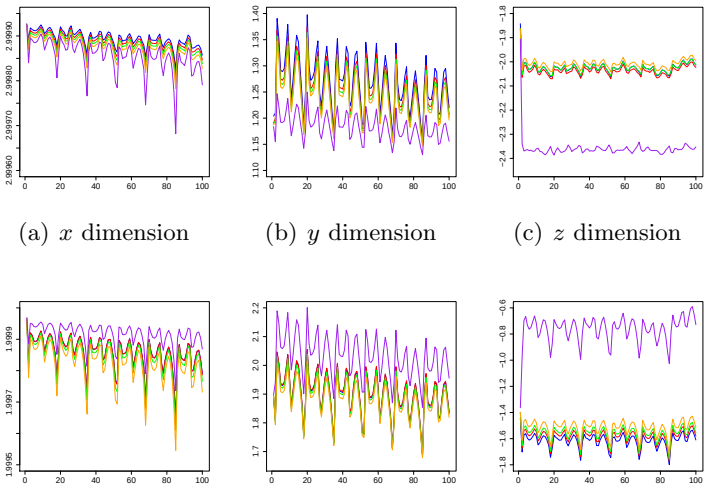


(a) $k^* = 0$ (b) $k^* = 1$ (c) $k^* = 2$ (d) $k^* = 3$ (e) $k^* = 4$

(f) $k^* = 0$ (g) $k^* = 1$ (h) $k^* = 2$ (i) $k^* = 3$ (j) $k^* = 4$

(k) $k^* = 0$ (l) $k^* = 1$ (m) $k^* = 2$ (n) $k^* = 3$ (o) $k^* = 4$

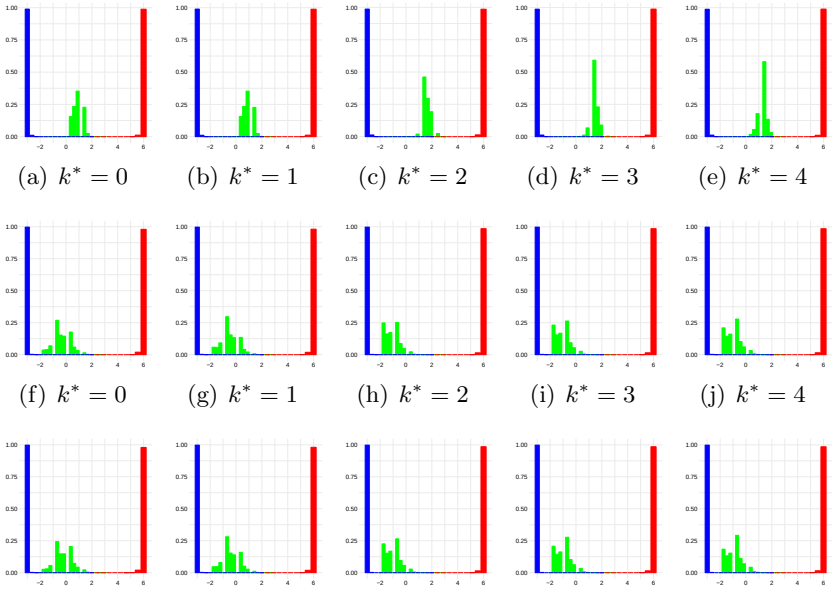
FIG 14. *Marginal posterior distribution for location parameter* (x, y, z) *of source* 2 *in the time window* $[12000, 12k^*99]$ *ms, where* $k^* = 0, 1, 2, 3, 4$. *Top row:* $t = 12000$ *ms; middle row:* $t = 12040$ *ms; bottom row:* $t = 12080$ *ms. Green bar: marginal posterior distribution for parameter* x*; red bar: marginal posterior distribution for parameter* y*; blue bar: marginal posterior distribution for parameter* z.



(a) x dimension (b) y dimension (c) z dimension

(d) x dimension (e) y dimension (f) z dimension

FIG 15. *Marginal posterior means for location parameter* (x, y, z) *of two sources in the time window* $[12000, 12k^*99]$ *ms, where* $k^* = 0, 1, 2, 3, 4$. *Top row: source* 1*; bottom row: source* 2. *Blue line:* $k^* = 0$*; red line:* $k^* = 1$*; green line:* $k^* = 2$*; orange line:* $k^* = 3$*; purple line:* $k^* = 4$.

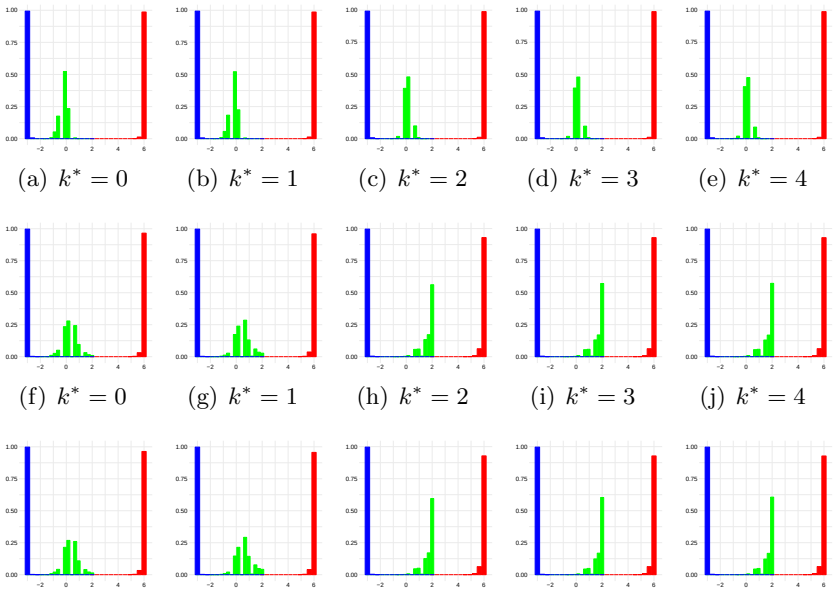


(a) $k^* = 0$ (b) $k^* = 1$ (c) $k^* = 2$ (d) $k^* = 3$ (e) $k^* = 4$

(f) $k^* = 0$ (g) $k^* = 1$ (h) $k^* = 2$ (i) $k^* = 3$ (j) $k^* = 4$

(k) $k^* = 0$ (l) $k^* = 1$ (m) $k^* = 2$ (n) $k^* = 3$ (o) $k^* = 4$

FIG 16. *Marginal posterior distribution for location parameter* (x, y, z) *of source* 1 *in the time window* $[20000, 20k^*99]$ *ms, where* $k^* = 0, 1, 2, 3, 4$. *Top row:* $t = 20000$ *ms; middle row:* $t = 20040$ *ms; bottom row:* $t = 20080$ *ms. Green bar: marginal posterior distribution for parameter* x*; red bar: marginal posterior distribution for parameter* y*; blue bar: marginal posterior distribution for parameter* z.

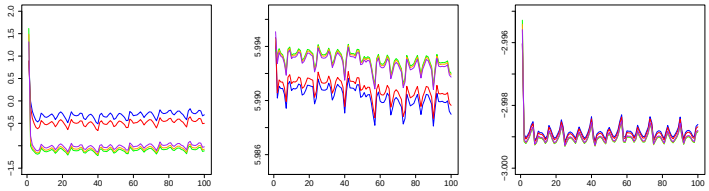


(a) $k^* = 0$ (b) $k^* = 1$ (c) $k^* = 2$ (d) $k^* = 3$ (e) $k^* = 4$

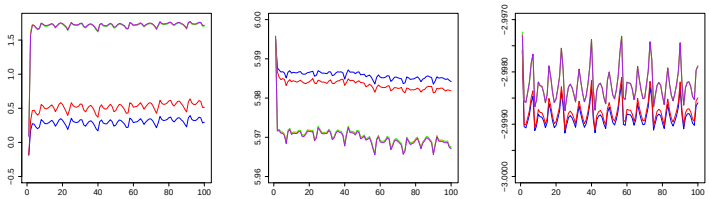
(f) $k^* = 0$ (g) $k^* = 1$ (h) $k^* = 2$ (i) $k^* = 3$ (j) $k^* = 4$

(k) $k^* = 0$ (l) $k^* = 1$ (m) $k^* = 2$ (n) $k^* = 3$ (o) $k^* = 4$

FIG 17. *Marginal posterior distribution for location parameter* (x, y, z) *of source* 2 *in the time window* $[20000, 20k^*99]$ *ms, where* $k^* = 0, 1, 2, 3, 4$. *Top row:* $t = 20000$ *ms; middle row:* $t = 20040$ *ms; bottom row:* $t = 20080$ *ms. Green bar: marginal posterior distribution for parameter* x*; red bar: marginal posterior distribution for parameter* y*; blue bar: marginal posterior distribution for parameter* z.



(a) x dimension (b) y dimension (c) z dimension



(d) x dimension (e) y dimension (f) z dimension

FIG 18. *Marginal posterior means for location parameter* (x, y, z) *of two sources in the time window* $[20000, 20k^*99]$ *ms, where* $k^* = 0, 1, 2, 3, 4$. *Top row: source* 1*; bottom row: source* 2. *Blue line:* $k^* = 0$*; red line:* $k^* = 1$*; green line:* $k^* = 2$*; orange line:* $k^* = 3$*; purple line:* $k^* = 4$.

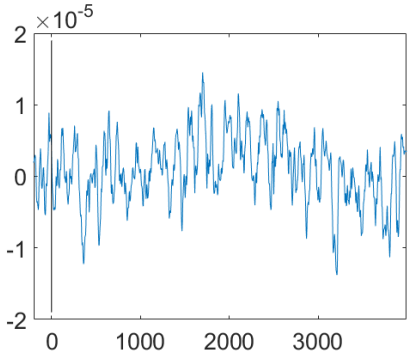


FIG 19. *EEG recording from one channel during retention under load condition* 1. *The horizontal axis represents time (ms) and the vertical axis represents the potential field (volt).*

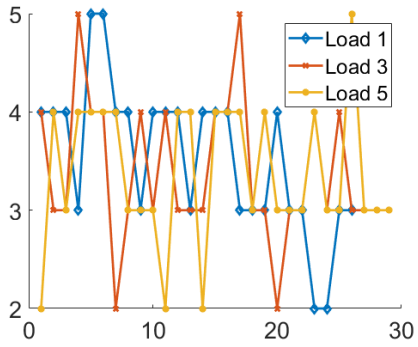


FIG 20. *Estimated number of sources in the SWM task after event onset during retention. The horizontal axis represents the index of trials and the vertical axis represents the number of estimated sources.*

References.

YAO, Z., FAN, Z., HAYASHI, M. and EDDY, F. W. (2018a). Quantifying the time-varying source in Magnetoencephalography – a discrete approach. *Unpublished manuscript* xx-xx.

YAO, Z., ZHANG, Y., BAI, Z. and EDDY, W. F. (2018b). Estimating the number of sources in Magnetoencephalography using spiked population eigenvalues. *Journal of the American Statistical Association* **113** 505–518.

DEPARTMENT OF STATISTICS AND APPLIED PROBABILITY
NATIONAL UNIVERSITY OF SINGAPORE,
21 LOWER KENT RIDGE ROAD, SINGAPORE 117546
E-MAIL: zhigang.yao@nus.edu.sg

DEPARTMENT OF STATISTICS AND APPLIED PROBABILITY
NATIONAL UNIVERSITY OF SINGAPORE,
21 LOWER KENT RIDGE ROAD, SINGAPORE 117546
E-MAIL: stafz@nus.edu.sg

GRADUATE SCHOOL OF MATHEMATICS
NAGOYA UNIVERSITY
NAGOYA, JAPAN 464-8602
E-MAIL: masahito@math.nagoya-u.ac.jp

DEPARTMENT OF STATISTICS
CARNEGIE MELLON UNIVERSITY
PITTSBURGH, PENNSYLVANIA 15213
E-MAIL: bill@stat.cmu.edu