

УДК 612.822

БАЙЕСОВСКИЙ ПАРАЛЛЕЛЬНЫЙ ФАКТОРНЫЙ АНАЛИЗ ДЛЯ ИССЛЕДОВАНИЯ СВЯЗАННЫХ С СОБЫТИЯМИ ПОТЕНЦИАЛОВ

© 2020 г. В. А. Пономарев^{1,*}, Ю. Д. Кропотов¹

¹ Федеральное государственное бюджетное учреждение науки Институт мозга человека им. Н.П. Бехтеревой
Российской академии наук, Санкт-Петербург, Россия

*e-mail: valery_ponomarev@mail.ru

Поступила в редакцию 24.04.2020 г.

После доработки 01.06.2020 г.

Принята к публикации 01.06.2020 г.

Целью данной работы являлась разработка байесовской вероятностной модели для параллельного факторного анализа связанных с событиями потенциалов (ПСС) мозга человека. Предложено 12 статистических моделей, учитывающих специфику сигналов источников ПСС. Для этих моделей были разработаны процедуры построения выборки случайных значений их параметров, основанные на методах Монте-Карло с цепями Маркова. Эффективность этих процедур оценивалась, используя как синтезированные данные с различным соотношением сигнал–шум, так и массив записей ПСС, полученных у 351 человека в *Go/NoGo* тесте. Была выбрана процедура, позволяющая получить оценки параметров модели с наилучшей точностью. Анализ зависимости сигналов модели от вида выполняемой деятельности человеком показал, что байесовский параллельный факторный анализ позволяет выделить функционально различные компоненты ПСС.

Ключевые слова: связанные с событиями потенциалы, параллельный факторный анализ, *Go/NoGo* тест

DOI: 10.31857/S004446772006009X

ВВЕДЕНИЕ

Потенциалы, связанные с событиями (ПСС), широко используются в исследованиях процессов мозга [Кропотов, 2010; Luck, Carreman, 2011]. В них отражается протекание различных психических процессов, таких как восприятие стимулов и их категоризация, принятие решения о последующем действии, переключение с одного вида выполняемого задания на другой, подготовка ответа или его подавление, обнаружение несоответствия между доминирующей моделью поведения и текущим контекстом и многие другие. Это позволяет получить объективную информацию о мозговых механизмах, лежащих в основе этих процессов.

Типичная и широко распространенная схема исследования ПСС выглядит следующим образом. Электроэнцефалограмма регистрируется с помощью электродов, расположенных на поверхности кожи головы у человека, который выполняет определенный психологический тест. Причем, как правило,

такие исследования выполняются с участием представительного числа испытуемых. В психологическом тесте выполняются несколько видов деятельности, случайным образом чередующейся в зависимости от типа предъявляемых стимулов. Методом усреднения вычисляются ПСС для каждого из видов деятельности отдельно и сравниваются между собой. Среднегрупповые ПСС вычисляются для выявления общих для группы особенностей. Подбирая подходящие виды деятельности, такой подход позволяет выявить составляющие ПСС, связанные с различными психическими процессами, оценить временной интервал, в течение которого они протекают, и определить зоны мозга, которые участвуют в их реализации.

В качестве альтернативного подхода для выявления общих характеристик ПСС может рассматриваться параллельный факторный анализ, использующий тензорное разложение [Cichocki et al., 2015; Cong et al., 2015; Zhou et al., 2015] и позволяющий представить

ПСС в виде суммы компонент с различными сигналами и топографиями [Пономарев и др., 2019а]. Применение параллельного факторного анализа может оказаться перспективным, поскольку с его помощью удастся выделить функционально различные компоненты ПСС [Пономарев, et al., 2019б].

Чтобы оценить компоненты модели параллельного факторного анализа, часто используется метод наименьших квадратов, а для поиска минимума целевой функции применяются различные итеративные процедуры [Kolda, Bader, 2009; Cichocki et al., 2015]. Такой подход имеет ряд ограничений.

Во-первых, при наличии шума целевая функция может оказаться невыпуклой и иметь несколько локальных минимумов. В таких случаях итеративные процедуры (в частности, метод стохастического градиентного спуска) могут сходиться только к одному из локальных минимумов в зависимости от начального приближения.

Во-вторых, аналогично традиционному подходу, необходимо выполнить сравнение компонент ПСС при разных видах деятельности. Для этого необходимо оценить доверительные интервалы для компонент ПСС. Теоретически эту задачу можно решить с помощью многократной генерации искусственных выборок записей ПСС методом Монте-Карло (бутстрэп), но на практике это приводит к необходимости много раз выполнять тензорное разложение, что требует значительных временных затрат. Кроме того, из-за невыпуклости целевой функции такой метод может давать слишком широкие доверительные интервалы.

Наконец, в-третьих, число компонент в модели неизвестно, но при этом метод наименьших квадратов не подразумевает возможности получить оценку этого параметра.

Для решения перечисленных проблем широко используется байесовское иерархическое моделирование [Gelman et al., 2014], которое, в частности, применялось для решения задач параллельного факторного анализа [Schmidt, Mohamed, 2009; Xiong et al., 2010; Natarajan et al., 2014; Nakatsuji et al., 2016; Chen et al., 2019 a, b]. Однако предложенные статистические модели были разработаны для анализа многомерных данных другой природы и не учитывают специфику сигналов источников ПСС. Простое использование таких известных моделей может приво-

дить к противоречивым или трудно интерпретируемым результатам.

Целью данной работы являлась разработка байесовской модели для параллельного факторного анализа ПСС.

МЕТОДИКА

Пусть матрица \mathbf{X} размера $K \times T$, K – число электродов и T – число отсчетов времени, описывает ПСС. Обычно в экспериментах с регистрацией ПСС у человека предъявляется несколько проб (N), по которым вычисляются средние значения ПСС. В этом случае имеется возможность получить оценки дисперсии ПСС для каждого канала и каждого отсчета времени в отдельности, которые будем обозначать как \mathbf{Z} . Если имеется несколько записей ПСС у разных людей, полученных в одинаковых условиях, то они описываются множеством матриц \mathbf{X}^j и \mathbf{Z}^j размера $K \times T$, с числом проб N_j , где $j = 1 \dots J$ и J – число участников исследования. И в этом случае множества матриц \mathbf{X}^j и \mathbf{Z}^j удобно объединить в трехмерные массивы данных $\mathbf{Y} = \{\mathbf{X}^1 \dots \mathbf{X}^J\}'$ и $\mathbf{V} = \{\mathbf{Z}^1 \dots \mathbf{Z}^J\}'$ размера $J \times K \times T$ с элементами $y_{j,k,t}$ и $v_{j,k,t}$ соответственно.

Допустим, что (в соответствии с квазистационарной аппроксимацией уравнений Максвелла) эти ПСС описываются моделью линейного мгновенного смешивания [Hallez et al., 2007]:

$$x_{k,t} = \sum_{d=1}^D a_{d,k} s_{d,t} + e_{k,t} \quad \text{или} \quad \mathbf{X} = \mathbf{A}'\mathbf{S} + \mathbf{E}, \quad (1)$$

где $\mathbf{S} \equiv [s_1, s_2 \dots s_T]$ – матрица размера $D \times T$, описывающая сигналы источников ПСС, D – число источников (компонент), $\mathbf{A} \equiv [a_1, a_2 \dots a_K]$ – матрица смешивания размера $D \times K$, строки матрицы описывают топографии источников, и \mathbf{E} – матрица размера $D \times K$, описывающая аддитивный шум.

Допустим также, что справедливо приближение, при котором топографии и сигналы источников ПСС одинаковы у всех людей, включая полярность, а изменяется только их величина. Тогда ПСС у всех участников исследования описываются следующей моделью:

$$y_{j,k,t} = \sum_{d=1}^D u_{d,j} a_{d,k} s_{d,t} + e_{j,k,t} \equiv r_{j,k,t} + e_{j,k,t}, \quad (2)$$

где $\mathbf{U} \equiv [\mathbf{u}_1, \mathbf{u}_2 \dots \mathbf{u}_J]$ – матрица размера $D \times J$ с элементами $u_{d,j} \geq 0$, содержащими величину сигнала с индексом d у участника исследования с индексом j , а $r_{j,k,t} = \sum_{d=1}^D u_{d,j} a_{d,k} s_{d,t}$ – так называемая каноническая полиадическая модель (*canonical polyadic model*), или модель параллельного факторного анализа (*parallel factor analysis* или *PARAFAC*), которую для краткости будем обозначать как “CP-модель” [Kolda, Bader, 2009; Cichocki et al., 2015; Cong et al., 2015; Zhou et al., 2015].

Исходно матрицы \mathbf{U} , \mathbf{A} и \mathbf{S} неизвестны, и необходимо их определить. Причем это возможно сделать только с точностью до произвольных масштабов и перестановок компонент [Kolda, Bader, 2009]. Для оценки матриц \mathbf{U} , \mathbf{A} и \mathbf{S} используется метод Монте-Карло [Gelman et al., 2014]. Используя подходы, применяемые в байесовском иерархическом моделировании, в настоящем исследовании была построена статистическая модель сигналов, основанная на следующих предположениях.

1) Шум представляет собой последовательность независимых нормально распределенных величин с нулевым средним и дисперсией $\sigma_{j,k,t}^2$. Тогда функцию максимального правдоподобия можно представить в следующем виде

$$p(\mathbf{Y}|\mathbf{U}, \mathbf{A}, \mathbf{S}, \vartheta) = \prod_{j,k,t} N_1(y_{j,k,t} | r_{j,k,t}, \vartheta_{j,k,t}^{-1}), \quad (3)$$

где $N_1(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$, и

$\vartheta_{j,k,t} = 1/\sigma_{j,k,t}^2$ называют “точностью”.

2) Векторы-столбцы матрицы \mathbf{A} взаимно независимы и имеют многомерное нормальное распределение с произвольными средними μ_k^A и матрицами точности Λ_k^A , где $\Lambda_k^A = \Sigma_k^{A^{-1}}$, и Σ_k^A – матрица ковариаций [Salakhutdinov, Mnih, 2008; Natarajan et al., 2014; Nakatsuji et al., 2016; Chen et al., 2019 a, b]

$$p(\mathbf{a}_k | \mu_k^A, \Lambda_k^A) = N_D(\mathbf{a}_k | \mu_k^A, \Lambda_k^{A^{-1}}), \quad (4)$$

где $N_D(\mathbf{x}|\mu, \Lambda^{-1}) = (2\pi)^{-\frac{D}{2}} |\Lambda| \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^T \Lambda (\mathbf{x}-\mu)\right)$ – многомерное нормальное распределение.

3) Векторы-столбцы матрицы \mathbf{U} взаимно независимы и имеют усеченное в интервале от 0 до $+\infty$ многомерное нормальное распре-

деление с произвольными μ_j^U и Λ_j^U [Alquier, Guedj, 2017]

$$p(\mathbf{u}_j | \mu_j^U, \Lambda_j^U) = R_D(\mathbf{u}_j | \mu_j^U, \Lambda_j^{U^{-1}}) \equiv \frac{1}{c_j} N_D(\mathbf{u}_j | \mu_j^U, \Lambda_j^{U^{-1}}) \prod_d \theta(u_{d,j}), \quad (5)$$

где $c = \int_0^\infty N_D(\mathbf{x}|\mu, \Lambda^{-1}) d\mathbf{x}$ – константа интегрирования, которая, используя аппроксимацию, может быть вычислена численными методами [Trinh, Genz, 2015], $\theta(x)$ – единичная ступенчатая функция $\theta(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$.

4) Сигналы источников таковы, что их значения в момент времени t зависят только от предыдущих значений $(t-1)$ и имеют многомерное нормальное распределение с произвольными μ^S и Λ^S [Xiong et al., 2010]

$$p(\mathbf{s}_1 | \mu^S, \Lambda^S) = N_D(\mathbf{s}_1 | \mu^S, \Lambda^S), \quad t = 1, \quad (6)$$

$$p(\mathbf{s}_t | \mathbf{s}_{t-1}, \Lambda^S) = N_D(\mathbf{s}_t | \mathbf{s}_{t-1}, \Lambda^S), \quad t = 2 \dots T. \quad (7)$$

5) Точность $\vartheta_{j,k,t}$ является случайной величиной с распределением

$$p(\vartheta_{j,k,t} | N_j, v_{j,k,t}) = \Gamma(\vartheta_{j,k,t} | N_j/2, N_j v_{j,k,t}/2), \quad (8)$$

где $\Gamma(x|a, b)$ – Гамма-распределение.

6) Скрытые параметры модели $\mu_k^A, \Lambda_k^A, \mu_j^U, \Lambda_j^U, \mu^S$ и Λ^S являются случайными величинами со следующими распределениями [Xiong et al., 2010; Natarajan et al., 2014; Nakatsuji et al., 2016; Chen et al., 2019 a, b]

$$p(\mu_k^A, \Lambda_k^A) = N_D(\mu_k^A | \mu_0^A, (\beta_0^A \Lambda_k^A)^{-1}) \cdot W_D(\Lambda_k^A | \Lambda_0^A, v_0^A), \quad (9)$$

$$p(\mu_j^U, \Lambda_j^U) = c_j \cdot R_D(\mu_j^U | \mu_0^U, (\beta_0^U \Lambda_j^U)^{-1}) \cdot W_D(\Lambda_j^U | \Lambda_0^U, v_0^U), \quad (10)$$

$$p(\mu^S, \Lambda^S) = N_D(\mu^S | \mu_0^S, (\beta_0^S \Lambda^S)^{-1}) \cdot W_D(\Lambda^S | \Lambda_0^S, v_0^S), \quad (11)$$

$$p(\Lambda_t^S | \Lambda_0^S, v_0^S) = W_D(\Lambda_t^S | \Lambda_0^S, v_0^S), \quad (12)$$

где $W_D(\Lambda | \Lambda_0, v_0)$ – распределение Уишарта для матриц размера $D \times D$ с v_0 степенями свободы ($v_0 > D - 1$), и $\mu_0^U, \beta_0^U, \Lambda_0^U, v_0^U, \mu_0^A, \beta_0^A, \Lambda_0^A, v_0^A, \mu_0^S, \beta_0^S, \Lambda_0^S, v_0^S$ – параметры, отражающие наши априорные знания относительно специфики конкретной задачи и являющиеся константами.

Таблица 1. Свойства моделей для разных алгоритмов
Table 1. Model properties for different algorithm

Алгоритм	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
Компоненты взаимно независимые	+	+	+	+	–	–	–	–	–	–	–	–
Λ_j^U, Λ_k^A и Λ_t^S диагональные	+	+	+	+	+	+	+	+	–	–	–	–
Учитывается автокорреляция сигналов	–	+	–	+	–	+	–	+	–	+	–	+
$\mu_j^U, \mu_k^A, \mu_t^S, \Lambda_j^U, \Lambda_k^A$ и Λ_t^S зависят от индексов k, j и t	+	+	–	–	+	+	–	–	+	+	–	–

* – в соответствии с формулой (6) для моделей, учитывающих автокорреляцию сигналов, скрытый параметр μ^S является вектором, и не имеет индекса t .

* – in accordance with formula (6) for models that take into account the autocorrelation of signals, the hidden parameter μ^S is a vector and does not have an index t .

Наконец, допустим, что совместное условное распределение для всех неизвестных параметров модели может быть представлено в следующем виде

$$\begin{aligned}
 & p(\mathbf{U}, \mathbf{A}, \mathbf{S}, \mu^U, \Lambda^U, \mu^A, \Lambda^A, \mu^S, \Lambda^S, \vartheta | \mathbf{Y}, \mathbf{V}, \mathbf{N}) = \\
 & = p(\mathbf{Y} | \mathbf{U}, \mathbf{A}, \mathbf{S}, \vartheta) \prod_j p(\mathbf{u}_j | \mu_j^U, \Lambda_j^U) \times \\
 & \times \prod_k p(\mathbf{a}_k | \mu_k^A, \Lambda_k^A) p(\mathbf{s}_1 | \mu^S, \Lambda_1^S) \times \\
 & \times \prod_{t=2}^T p(\mathbf{s}_t | \mathbf{s}_{t-1}^S, \Lambda_t^S) \prod_j p(\mu_j^U, \Lambda_j^U) \times \\
 & \times \prod_k p(\mu_k^A, \Lambda_k^A) p(\mu^S, \Lambda_1^S) \prod_{t=2}^T p(\Lambda_t^S) \times \\
 & \times \prod_{j,k,t} p(\vartheta_{j,k,t} | N_{j,k,t}, v_{j,k,t}).
 \end{aligned} \tag{13}$$

Рассматривались различные модификации этой модели, в которых предполагалось, что:

1) Компоненты модели взаимно независимые [Schmidt, Mohamed, 2009]. Тогда формулы (4) и (9) принимают следующий вид:

$$\begin{aligned}
 & p(\mathbf{a}_k | \mu_k^A, \Lambda_k^A) = \prod_d N_1(a_{d,k} | \mu_{d,k}^A, \alpha_{d,k}^{A-1}) \text{ и } p(\mu_k^A, \Lambda_k^A) = \\
 & = \prod_d N_1(\mu_{d,k}^A | \mu_0^A, (\beta_0^A \alpha_{d,k}^A)^{-1}) \Gamma(\alpha_{d,k}^A | a_0^A, b_0^A), \text{ где } \mu_0^A, \\
 & a_0^A \text{ и } b_0^A \text{ – скалярные величины. Аналогично для формул (5)–(7) и (10)–(12).}
 \end{aligned}$$

2) Компоненты модели взаимно зависимые, но матрицы Λ_j^U, Λ_k^A и Λ_t^S диагональные [Alquier, Guedj, 2017]. В этом случае изменялись только формулы (9–12) аналогично тому, как это описано выше.

3) Сигналы не имеют автокорреляции [Natarajan et al., 2014; Alquier, Guedj, 2017]. В этом случае μ^S заменяется на μ_t^S , формулы (6), (7), (11) и (12) – на формулы, аналогичные (4) и (9) соответственно, а произведения $p(\mathbf{s}_1 | \mu^S, \Lambda_1^S) \prod_{t=2}^T p(\mathbf{s}_t | \mathbf{s}_{t-1}^S, \Lambda_t^S)$ и $p(\mu^S, \Lambda_1^S) \prod_{t=2}^T p(\Lambda_t^S)$ в формуле (13) – на $\prod_t p(\mathbf{s}_t | \mu_t^S, \Lambda_t^S)$ и $\prod_t p(\mu_t^S, \Lambda_t^S)$ соответственно.

4) Скрытые параметры $\mu_j^U, \mu_k^A, \mu_t^S, \Lambda_j^U, \Lambda_k^A$ и Λ_t^S не зависят от индексов k, j и t соответственно [Xiong et al., 2010; Natarajan et al., 2014; Alquier, Guedj, 2017]. При этом в формуле (13) произведения $\prod_j p(\mu_j^U, \Lambda_j^U), \prod_k p(\mu_k^A, \Lambda_k^A)$ и $p(\mu^S, \Lambda_1^S) \prod_{t=2}^T p(\Lambda_t^S)$ (или $\prod_t p(\mu_t^S, \Lambda_t^S)$ – см. выше) заменяются на $p(\mu^U, \Lambda^U), p(\mu^A, \Lambda^A)$ и $p(\mu^S, \Lambda^S)$.

Поскольку заранее неизвестно, какие из предположений лучше соответствуют свойствам ПСС, рассматривались 12 различных моделей, перечисленных в табл. 1.

Условное распределение для каждого из неизвестных параметров модели несложно получить, используя формулу (13). Во всех случаях эти условные распределения могут быть приведены к стандартным, таким как гамма-распределение, распределение Уишарта, одномерные или многомерные нормальное и усеченное нормальное распределения, для которых предложены эффективные алгоритмы генерации случайных чисел [Smith, Носк-

ing, 1972; Press et al., 2007; Li, Ghosh, 2015; Koch, Vorr, 2019]. Вывод окончательных формул для 12 моделей можно найти в приложении к Руководству пользователя для программного обеспечения WinEEG (<https://mit-sar-eeg.com/>).

Для построения выборки случайных значений параметров CP-модели использовался метод Гиббса – “Gibbs sampler”, аналогичный тем, которые описаны ранее [Xiong et al., 2010; Natarajan et al., 2014; Nakatsuji et al., 2016; Alquier, Guedj, 2017; Chen et al., 2019a].

В качестве примера рассмотрим алгоритм A10, в основе которого лежат предположения относительно распределений случайных величин, описываемые формулами (3–12), а входными данными являются трехмерные массивы \mathbf{Y} и \mathbf{V} , одномерный массив \mathbf{N} и матрицы \mathbf{U} , \mathbf{A} и \mathbf{S} , элементам которых присвоены случайные значения. На каждом шаге этой итеративной процедуры выполнялись следующие действия.

1) Для всех значений d выполняется перенормировка строк матриц \mathbf{U} , \mathbf{A} и \mathbf{S} следующим образом: $u_{d,j} = u_{d,j} \frac{J}{\|\mathbf{u}_{d,\bullet}\|_1}$, $a_{d,k} = a_{d,k} \frac{K}{\|\mathbf{a}_{d,\bullet}\|_F}$ и $s_{d,t} = s_{d,t} \frac{\|\mathbf{u}_{d,\bullet}\|_1 \|\mathbf{a}_{d,\bullet}\|_F}{JK}$, $\|\cdot\|_1$ – сумма абсолютных величин элементов вектора, $\|\cdot\|_F$ – норма Фробениуса.

2) Для всех значений k, j и t выбираются $\vartheta_{j,k,t}$ из $\Gamma(\vartheta_{j,k,t} | \overline{a_j}, \overline{b_{j,k,t}})$, где $\overline{a_j} = \frac{N_j + 1}{2}$ и $\overline{b_{j,k,t}} = \frac{1}{2}[(y_{j,k,t} - r_{j,k,t})^2 + N_j v_{j,k,t}]$.

3) Для всех значений j выбираются матрицы Λ_j^U из $W_D(\Lambda_j^U | \Lambda_j^{*U}, v^{*U})$, где $(\Lambda_j^{*U})^{-1} = (\Lambda_0^U)^{-1} + \frac{\beta_0^U}{\beta_0^U + 1}(\mu_0^U - \mathbf{u}_j)(\mu_0^U - \mathbf{u}_j)'$ и $v^{*U} = v_0^U + 1$, и далее – векторы μ_j^U из $R_D(\mu_j^U | \mu_j^{*U}, (\beta^{*U} \Lambda_j^U)^{-1})$, где $\mu_j^{*U} = \frac{\beta_0^U \mu_0^U + \mathbf{u}_j}{\beta_0^U + 1}$ и $\beta^{*U} = \beta_0^U + 1$.

4) Для всех значений k выбираются матрицы Λ_k^A из $W_D(\Lambda_k^A | \Lambda_k^{*A}, v^{*A})$, где $(\Lambda_k^{*A})^{-1} = (\Lambda_0^A)^{-1} + \frac{\beta_0^A}{\beta_0^A + 1}(\mu_0^A - \mathbf{a}_k)(\mu_0^A - \mathbf{a}_k)'$ и $v^{*A} = v_0^A + 1$, и далее –

векторы μ_k^A из $N_D(\mu_k^A | \mu_k^{*A}, (\beta^{*A} \Lambda_k^A)^{-1})$, где $\mu_k^{*A} = \frac{\beta_0^A \mu_0^A + \mathbf{a}_k}{\beta_0^A + 1}$ и $\beta^{*A} = \beta_0^A + 1$.

5) Для $t = 1$ выбирается матрица Λ_1^S из $W_D(\Lambda_1^S | \Lambda_1^{*S}, v^{*S})$, где $(\Lambda_1^{*S})^{-1} = (\Lambda_0^S)^{-1} + \frac{\beta_0^S}{\beta_0^S + 1}(\mu_0^S - \mathbf{s}_1)(\mu_0^S - \mathbf{s}_1)'$ и $v^{*S} = v_0^S + 1$, потом – вектор μ^S из $N_D(\mu^S | \mu^{*S}, (\beta^{*S} \Lambda_1^{*S})^{-1})$, где $\mu^{*S} = \frac{\beta_0^S \mu_0^S + \mathbf{s}_1}{\beta_0^S + 1}$ и $\beta^{*S} = \beta_0^S + 1$, и далее – матрицы Λ_t^S из $W_D(\Lambda_t^S | \Lambda_t^{*S}, v^{*S})$ для всех значений t последовательно, начиная с $t = 2$, где $(\Lambda_t^{*S})^{-1} = (\Lambda_0^S)^{-1} + (\mathbf{s}_t - \mathbf{s}_{t-1})(\mathbf{s}_t - \mathbf{s}_{t-1})'$.

6) Для всех значений j выбираются \mathbf{u}_j^U из $R_D(\mathbf{u}_j^U | \mu_j^{+U}, \Lambda_j^{+U})$, где $\mu_j^{+U} = (\Lambda_j^{+U})^{-1}(\Lambda_j^U \mu_j^U + \mathbf{b}_j^U)$, $\Lambda_j^{+U} = \Lambda_j^U + \mathbf{G}_j^U$, $\mathbf{b}_{d,j}^U = \sum_{k,t} \vartheta_{j,k,t} y_{j,k,t} \mathbf{q}_{d,k,t}^U \mathbf{g}_{d1,d2,j}^U = \sum_{k,t} \vartheta_{j,k,t} \mathbf{q}_{d1,k,t}^U \mathbf{q}_{d2,k,t}^U$ и $\mathbf{q}_{d,k,t}^U = a_{d,k} s_{d,t}$.

7) Для всех значений k выбираются \mathbf{a}_k^A из $N_D(\mathbf{a}_k^A | \mu_k^{+A}, \Lambda_k^{+A})$, где $\mu_k^{+A} = (\Lambda_k^{+A})^{-1}(\Lambda_k^A \mu_k^A + \mathbf{b}_k^A)$, $\Lambda_k^{+A} = \Lambda_k^A + \mathbf{G}_k^A$, $\mathbf{b}_{d,k}^A = \sum_{j,t} \vartheta_{j,k,t} y_{j,k,t} \mathbf{q}_{d,j,t}^A \mathbf{g}_{d1,d2,k}^A = \sum_{j,t} \vartheta_{j,k,t} \mathbf{q}_{d1,j,t}^A \mathbf{q}_{d2,j,t}^A$ и $\mathbf{q}_{d,j,t}^A = u'_{d,j} s_{d,t}$.

8) Для всех значений t (последовательно, начиная с $t = 1$) выбираются \mathbf{s}_t^S из $N_D(\mathbf{s}_t^S | \mu_t^{+S}, \Lambda_t^{+S})$, где $\mu_t^{+S} = (\Lambda_t^{+S})^{-1}(\Lambda_t^S \mu_t^S + \Lambda_2^S \mathbf{s}_2 + \mathbf{b}_t^S)$, $\Lambda_t^{+S} = \Lambda_1^S + \Lambda_2^S + \mathbf{G}_t^S$, $\mu_t^{+S} = (\Lambda_t^{+S})^{-1}(\Lambda_t^S \mathbf{s}_{t-1}^S + \Lambda_{t+1}^S \mathbf{s}_{t+1}^S + \mathbf{b}_t^S)$, $\Lambda_t^{+S} = \Lambda_t^S + \Lambda_{t+1}^S + \mathbf{G}_t^S$, $\mu_T^{+S} = (\Lambda_T^{+S})^{-1}(\Lambda_T^S \mathbf{s}_{T-1}^S + \mathbf{b}_T^S)$, $\Lambda_T^{+S} = \Lambda_T^S + \mathbf{G}_T^S$, для $t = 1, 1 < t < T$ и $t = T$ соответственно, $\mathbf{b}_{d,t}^S = \sum_{j,k} \vartheta_{j,k,t} y_{j,k,t} \mathbf{q}_{d,j,k}^S \mathbf{g}_{d1,d2,t}^S = \sum_{j,k} \vartheta_{j,k,t} \mathbf{q}_{d1,j,k}^S \mathbf{q}_{d2,j,k}^S$ и $\mathbf{q}_{d,j,k}^S = u'_{d,j} a'_{d,k}$.

9) Запоминаются вновь полученные значения: $\mathbf{U} = \mathbf{U}'$, $\mathbf{A} = \mathbf{A}'$ и $\mathbf{S} = \mathbf{S}'$.

Постоянные параметры были следующими: $\mu_0^U = [1, \dots, 1]'$, $\beta_0^U = 1$, $\Lambda_0^U = \mathbf{I}_D$, $v_0^U = D$, $\mu_0^A = [0, \dots, 0]'$, $\beta_0^A = 1$, $\Lambda_0^A = \mathbf{I}_D$, $v_0^A = D$, $\mu_0^S = [0, \dots, 0]'$, $\beta_0^S = 1$, $\Lambda_0^S = \mathbf{I}_D$ и $v_0^S = D$, где \mathbf{I}_D – единичная матрица размера D .

Всего выполнялось 15000 шагов. Первые 10000 значений параметров отбрасывались, чтобы избежать влияния начальных условий, а последующие 5000 использовались для

оценки средних значений и доверительных интервалов для элементов матриц U , A и S .

Выборка случайных значений параметров в других алгоритмах выполнялась аналогично.

Типичной задачей исследования ПСС является выявление интервалов времени, в которых наблюдается зависимость ПСС от условий. Однако если матрицы U , A и S оцениваются для каждого условия отдельно, то вследствие принципиальной неопределенности масштабов компонент в CP -модели сравнение сигналов при различных условиях становится невозможным. Чтобы преодолеть это ограничение, можно предположить, что топографии компонент (матрица A) и величины компонент (матрица U) одинаковы для всех условий. Тогда становится возможным объединить ПСС, полученные при различных условиях, в единый временной ряд, и получить оценки U , A и S для всей совокупности данных одновременно. Дисперсии ПСС объединялись аналогично, а массив N имел дополнительную размерность, поскольку число проб могло зависеть от условий. Отметим также, что если статистическая модель учитывала автокорреляцию сигналов, то в соответствующих алгоритмах $A2$, $A4$, $A6$, $A8$, $A10$ и $A12$ выборка случайных значений для Λ_t^S , μ^S и s_t (операции 4 и 7) выполнялась для каждого условия отдельно.

Поскольку параметр D — число компонент — в модели неизвестен, его необходимо оценить. Для сравнения байесовских моделей можно использовать коэффициент Байеса, который определяется как

$$\frac{p(D_1|Y)}{p(D_2|Y)} = \frac{p(D_1)p(Y|D_1)}{p(D_2)p(Y|D_2)}, \quad (14)$$

где $p(D_1)$ и $p(D_2)$ — априорные вероятности моделей с числом компонент D_1 и D_2 , а $p(Y|D_1)$ и $p(Y|D_2)$ — функции предельного правдоподобия

$$p(Y|D_m) = \int p(Y|U, A, S, D_m) d\{U, A, S\}. \quad (15)$$

В том случае, когда $p(D_m)$ неизвестны, предполагается, что все $p(D_m)$ равны, тогда модель с максимальным $p(Y|D_m)$ считается наиболее вероятной.

Чтобы оценить $p(Y|D_m)$, необходимо вычислить интеграл в формуле (15), но получить аналитическое решение крайне сложно. Оценку $p(Y|D_m)$ можно получить, используя

метод, предложенный в работе Чибя [Chib, 1995]. В нашем случае при реализации метода Чибя определялись три подмножества параметров CP -модели, соответствующих матрицам U , A и S , и для каждого из них выполнялось 5000 шагов итеративной процедуры.

Для оценки эффективности 12 алгоритмов использовались массив записей ПСС, полученных в *Go/NoGo* тесте, и синтезированные данные.

Методика регистрации массива данных ПСС была подробно описана ранее [Пономарев, Кропотов, 2013; Kropotov et al., 2016; Пономарев и др., 2019 а, б]. Поэтому здесь приводится краткое описание. Массив включал в себя 351 запись, полученных у здоровых людей в возрасте от 18 до 55 лет (150 мужчин). Запись ЭЭГ выполнялась с помощью компьютерного электроэнцефалографа “Мицар-ЭЭГ” (полоса пропускания — 0.53–50 Гц, режкторный фильтр 45–55 Гц, частота квантования — 250 Гц). Электроды располагались в соответствии с международной системой 10–20, референт — объединенные электроды, расположенные на мочках ушей, заземляющий — в отведении *Frz*. Перед вычислением ПСС выполнялись коррекция артефактов моргания и удаление артефактов [Пономарев и др., 2019а].

Зрительный *Go/NoGo* тест состоял из 400 проб длительностью 3000 мс, по 100 проб каждого типа. В пробах предъявлялись пары зрительных стимулов (изображения животных — “Ж”, растений — “Р” и людей — “Ч”), для каждого типа проб пары были различные: “Ж–Ж”, “Ж–Р”, “Р–Р” и “Р–Ч”. Длительность предъявления стимула — 100 мс, интервал между стимулами в пробе — 1000 мс. Первый стимул в пробе предъявлялся через 300 мс после ее начала. В пробах “Р–Ч” также предъявлялся звуковой стимул для поддержания уровня внимания. Пробы различного типа предъявлялись в псевдослучайном порядке. Участнику исследования давалось задание нажимать на кнопку только после предъявления пары “Ж–Ж”. ПСС вычислялись отдельно для каждого из типов проб (условий). Пробы, в которых задание выполнялось неправильно, исключались. Для анализа использовались ПСС, полученные в пробах “Ж–Ж”, “Ж–Р” и “Р–Р”, а частота квантования уменьшалась до 125 Гц путем прореживания.

Все исследования проведены в соответствии с принципами биомедицинской этики, сформулированными в Хельсинкской декларации 1964 г. и ее последующих обновлениях, и одобрены Комитетом по этике Института мозга человека им. Н.П. Бехтерева Российской академии наук (ИМЧ РАН), Санкт-Петербург, Россия, и местным комитетом по этике кантона Граубюнден, округ Гризон, Швейцария. Каждый участник исследования представил добровольное письменное информированное согласие, подписанное им после разъяснения ему потенциальных рисков и преимуществ, а также характера предстоящего исследования.

Синтезированные данные были созданы следующим образом. Было выбрано несколько тестовых *CP*-моделей с различным числом компонент D от 1 до 8, которые были ранее получены для описанного выше массива записей ПСС. Для каждой из этих моделей вычислялись \mathbf{R} (см. формулу 2) и \mathbf{Y} ($J = 351$, $K = 19$, $T = 1125$) в соответствии с формулой

$$\mathbf{Y} = \mathbf{R} + w\mathbf{E} \frac{\|\mathbf{R}\|_F^2}{\|\mathbf{E}\|_F^2},$$

где \mathbf{E} – трехмерный массив случайных нормально распределенных чисел, полученных из распределения $N(0,1)$, $\|\cdot\|_F$ – норма Фробениуса и w – положительная величина, определяющая соотношение сигнал–шум, которая варьировалась. Для определения массивов \mathbf{V} и \mathbf{N} также использовались записи ПСС. Были вычислены среднегрупповые значения дисперсии ПСС. Далее, элементам трехмерного массива \mathbf{V} присваивались средние значения среднегрупповых дисперсий в интервале, предшествующем первому стимулу в пробе (первые 300 мс), различные для индекса k и одинаковые для индексов j и t . Массив \mathbf{N} содержал число проб для каждого условия и индивидуальной записи ПСС отдельно.

Для оценки точности описания исходных данных с помощью *CP*-модели использовались количественные показатели: относительная ошибка аппроксимации ε_{app} , инвариантное по отношению к произвольному масштабу расстояние между двумя *CP*-моделями $\hat{\Delta}$ и средние значения относительной ширины доверительного интервала $\overline{CI^U}$, $\overline{CI^A}$ и $\overline{CI^S}$ для элементов каждой из матриц \mathbf{U} , \mathbf{A} и \mathbf{S} отдельно. ε_{app} вычислялась в соответствии с

формулой $\varepsilon_{app} = \|\mathbf{Y} - \mathbf{R}\|_F^2 / \|\mathbf{Y}\|_F^2$. $\hat{\Delta} = \frac{1}{D} \sum_{d=1}^D \Delta_{d,d}$, где матрица $\{\Delta_{r,q}\}$ с элементами $\Delta_{r,q} = \delta(U_{r,\bullet}^1, U_{q,\bullet}^2) + \delta(A_{r,\bullet}^1, A_{q,\bullet}^2) + \delta(S_{r,\bullet}^1, S_{q,\bullet}^2)$ получена путем перестановок столбцов таким образом, чтобы максимальные значения $\Delta_{r,q}$ были расположены на диагонали, $\delta(\mathbf{b}_{r,\bullet}^1, \mathbf{b}_{q,\bullet}^2) = 1 - \frac{\langle \mathbf{b}_{r,\bullet}^1, \mathbf{b}_{q,\bullet}^2 \rangle}{\|\mathbf{b}_{r,\bullet}^1\| \|\mathbf{b}_{q,\bullet}^2\|}$, $\mathbf{b}_{r,\bullet}$ – строка матрицы \mathbf{B} с индексом r , $\langle \mathbf{u}, \mathbf{v} \rangle$ – скалярное произведение векторов \mathbf{u} и \mathbf{v} , $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$. $\overline{CI^U} = \frac{1}{D} \sum_{d=1}^D CI_d^U$, где CI_d^U – отношение суммы разностей верхней и нижней границ доверительного интервала к сумме абсолютных значений элементов строки матрицы \mathbf{U} с индексом d , аналогично для $\overline{CI^A}$ и $\overline{CI^S}$.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ

При использовании синтезированных данных, прежде всего, оценивалась способность двенадцати алгоритмов восстановить тестовую *CP*-модель при выборе соответствующего числа компонент D . Оказалось, что если тестовая *CP*-модель состоит из небольшого числа компонент ($D \leq 3$), то все 12 алгоритмов способны ее восстановить с удовлетворительной точностью ($\hat{\Delta} < 0.02$) независимо от соотношения сигнал–шум. При большем числе компонент для большинства алгоритмов величина $\hat{\Delta}$ была также меньше 0.02, исключая $A2$ и $A4$, для которых величина $\hat{\Delta}$ могла быть больше, чем 0.3.

При произвольном выборе числа компонент в *CP*-модели для большинства алгоритмов независимо от соотношения сигнал–шум типичная картина была следующей: относительная ошибка аппроксимации ε_{app} монотонно уменьшалась, а функция предельного правдоподобия $p(\mathbf{Y}|D)$ монотонно увеличивалась с ростом D . Но для алгоритмов $A2$, $A4$, $A7$ и $A8$ зависимости ε_{app} и $p(\mathbf{Y}|D)$ от D могли быть немонотонные.

Примеры монотонной зависимости $p(\mathbf{Y}|D)$ от D для сильно зашумленных синтезированных данных ($w = 1$), полученные с помощью алгоритма *A10*, представлены на рис. 1. Видно, что $p(\mathbf{Y}|D)$ быстро увеличивалась с ростом D , пока оно не достигает числа компо-

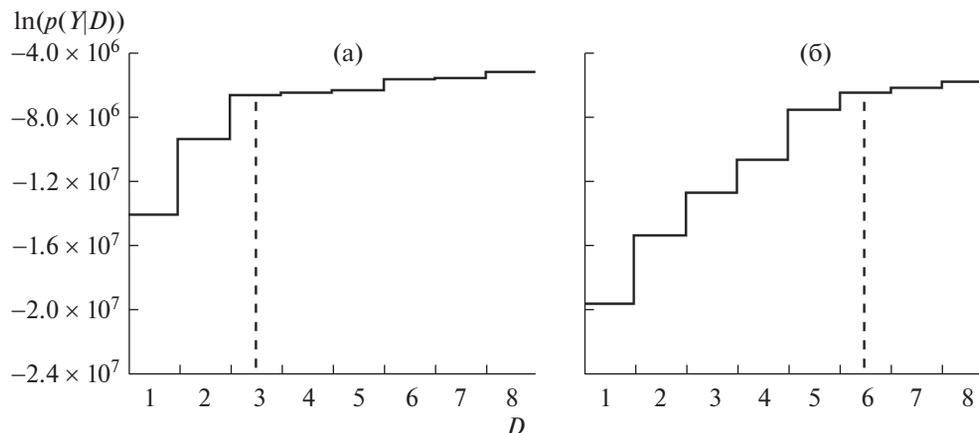


Рис. 1. Результаты оценки числа компонент в CP -модели для синтезированных данных. По оси абсцисс – число компонент (D), по оси ординат – логарифм функции предельного правдоподобия. Вертикальная линия показывает правильное D : (а) – $D = 3$, (б) – $D = 6$.

Fig. 1. The results of estimating the number of components in the CP -model for the synthetic data. Abscissa, the number of components (D); ordinate, the logarithm of marginal likelihood. The vertical line show the true value of D : (а) – $D = 3$, (б) – $D = 6$.

нент в тестовой CP -модели, но далее скорость роста $p(\mathbf{Y}|D)$ заметно уменьшается. Такое свойство зависимости $p(\mathbf{Y}|D)$ от числа компонент подтверждает возможность использования коэффициента Байеса для определения числа компонент. Причем оптимальное число компонент в CP -модели может соответствовать точке излома функции $p(\mathbf{Y}|D)$.

Для синтезированных данных величина $\overline{CI^U}$, $\overline{CI^A}$ и $\overline{CI^S}$ относительно мало зависела от соотношения сигнал–шум. При этом показатели $\overline{CI^U}$, $\overline{CI^A}$ и $\overline{CI^S}$ увеличивались с ростом D . Иными словами, чем больше параметров имела CP -модель (больше компонент), тем больше была дисперсия их оценки. Несмотря на то что эти наблюдения приводят к тривиальным выводам, тем не менее, они подтверждают адекватность использованных количественных показателей для оценки относительной ширины доверительных интервалов. Не столь очевидным оказался характер их зависимости от используемого алгоритма. В частности, если в статистических моделях предполагается, что скрытые параметры μ_j^U , μ_k^A , μ_t^S , Λ_j^U , Λ_k^A и Λ_t^S зависят от индексов k, j и t (алгоритмы $A1$, $A2$, $A5$, $A6$, $A9$ и $A10$), то в этих моделях количество скрытых параметров больше по сравнению с другими соответствующими (см. табл. 1). Тем не менее оказалось, что при использовании этих алгоритмов ве-

личина $\overline{CI^U}$, $\overline{CI^A}$ и $\overline{CI^S}$ была в 1.2–1.5 раза меньше по сравнению с алгоритмами $A3$, $A4$, $A7$, $A8$, $A11$ и $A12$ соответственно. Если же статистическая модель учитывала автокорреляцию сигналов (алгоритмы с четными номерами), это мало отражалось на величине $\overline{CI^U}$ и $\overline{CI^A}$, но для алгоритмов $A6$ и $A10$ величина $\overline{CI^S}$ была в 1.3–2 раза меньше по сравнению с десятью другими.

Представленные выше результаты указывают на то, что алгоритмы $A6$ и $A10$ являются наиболее эффективными. Во-первых, они обеспечивают приемлемую точность восстановления тестовой CP -модели. Во-вторых, дисперсия оценки параметров CP -модели относительно небольшая, по сравнению с другими алгоритмами. В-третьих, зависимость функции предельного правдоподобия от числа компонент является монотонной.

Анализ массива записей ПСС также выполнялся, используя все 12 алгоритмов. В целом результаты оценки эффективности алгоритмов совпадали с описанными выше. Поэтому здесь будут представлены результаты, полученные с помощью алгоритмов $A6$ и $A10$.

Результаты оценки оптимального числа компонент в CP -модели, полученные с помощью алгоритма $A10$, показаны на рис. 2. Причем оценка CP -модели выполнялась 8 раз для каждого значения D , что позволяло оценить зависимость оценки функции предельного

правдоподобия от начального приближения. Видно, что функция предельного правдоподобия монотонно увеличивалась с ростом D . Но скорость роста заметно уменьшается, если $D \geq 6$. Иными словами, точка излома функции $p(Y|D)$ соответствует $D = 6$. Аналогичные результаты были получены с помощью алгоритма $A6$.

При оптимальном D оценки параметров CP -модели относительно мало зависят от начального приближения. В то же время в результате сравнительного анализа использования алгоритмов $A6$ и $A10$ было выявлено, что последний показывает лучшую повторяемость. В частности, среднее значение \hat{D} , полученное при попарном сравнении CP -моделей, была равной 0.035 для $A10$, по сравнению с 0.232 для $A6$. Тем не менее среднее значение \hat{D} для ПСС больше, чем для синтезированных данных, что показывает необходимость выполнять тензорное разложение несколько раз для его верификации.

Пример одной компоненты CP -модели с доверительными интервалами (с надежностью 0.95), полученной с помощью алгоритма $A10$, показан на рис. 3. Отметим, что на графиках и топограммах все величины показаны в условных единицах, поскольку CP -модель может быть оценена только с точностью до произвольного масштаба. Ширина доверительного интервала для элементов матрицы U значительно меньше, чем разброс значений величин компоненты от человека к человеку (рис. 3 (а)). В то же время величина этой компоненты в среднем увеличивается с возрастом (рис. 3 (б)). Ширина доверительного интервала для элементов матрицы A относительно невелика (рис. 3 (в)), т.е. оценка топографии этой компоненты была получена с удовлетворительной точностью. Наконец, ширина доверительного интервала для сигналов (рис. 3 (г)) значительно меньше величины их отклонения от изолинии, и можно выделить интервалы времени, в течение которых эти доверительные интервалы для сигналов, соответствующих различным типам проб, не перекрываются. Таким образом, оценка сигналов была получена с удовлетворительной точностью для того, чтобы выявить различия сигналов при выполнении разных видов деятельности.

Топографии и сигналы шести компонент CP -модели показаны на рис. 4. Сигналы компонент нормированы так, что среднее значе-

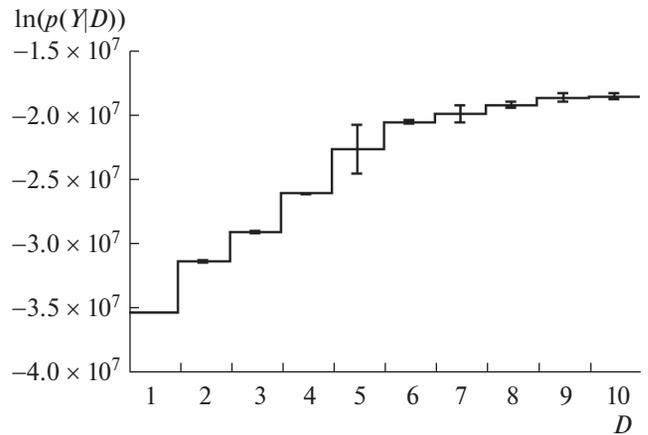


Рис. 2. Результаты оценки числа компонент в CP -модели для ПСС. Отрезки на кривых — плюс-минус два стандартных отклонения. Обозначения те же, что на рис. 1.

Fig. 2. The results of estimating the number of components in the CP -model for the ERPs. Spreads — plus or minus two standard deviations. For other designation, see Fig. 1.

ние величины компонент (элементов строк матрицы U) и абсолютное значение максимума на топографиях компонент (в строках матрицы A) равны 1. Ниже под графиками показаны интервалы времени, в течение которых при парном сравнении эти сигналы различались, т.е. соответствующие доверительные интервалы не перекрывались. Эта модель с хорошей точностью аппроксимирует среднegrupповые ПСС ($\epsilon_{app} = 0.014$) и описывает 71.8% дисперсии исходных данных.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Результаты исследования показали, что байесовский параллельный факторный анализ может оказаться эффективным средством для выявления общих характеристик сигналов мозговых источников, отражающихся в ПСС, регистрируемых на поверхности кожи головы. В то же время этот подход имеет ряд ограничений, связанных с тем, что CP -модель является определенным приближением. Рассмотрим особенности этого приближения подробнее.

Во-первых, предполагается, что топографии компонент одинаковы у всех людей. Это, безусловно, является существенным приближением, поскольку, несмотря на то, что анатомическое строение мозга сходно у различных людей, индивидуальные особенности

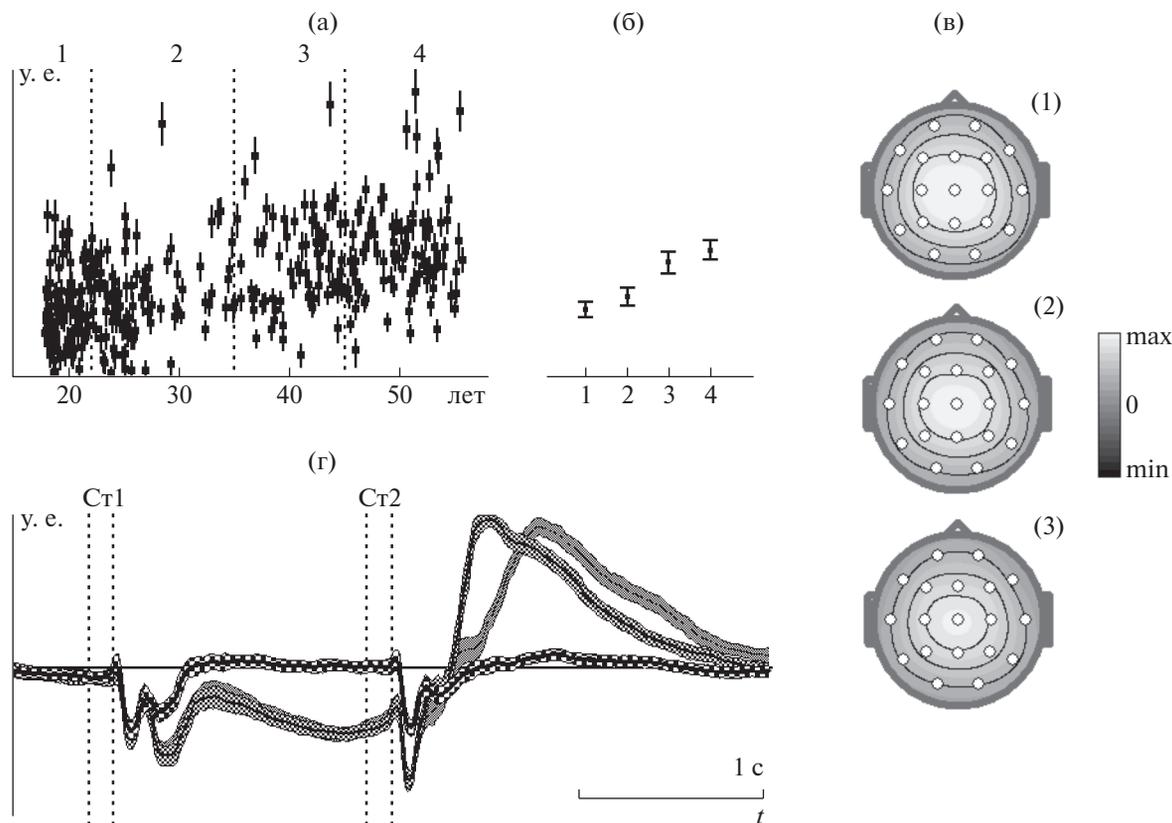


Рис. 3. Доверительные интервалы для параметров компоненты *CP*-модели. (а) – величина компоненты у разных испытуемых, по оси абсцисс – возраст, по оси ординат – величина компоненты в условных единицах, вертикальные пунктирные линии – границы четырех возрастных групп. (б) – среднее значение величины компоненты в четырех возрастных группах. (в) – топография компоненты (2) и нижняя (3) и верхняя (1) границы доверительных интервалов. (г) – сигналы компоненты в пробах *a* – “Ж–Ж”, *b* – “Ж–Р”, *c* – “Р–Р”. По оси абсцисс – время, по оси ординат – величина сигнала в условных единицах. Вертикальные пунктирные линии – моменты включения и выключения стимулов в пробе.

Fig. 3. Confidence intervals for the parameters of the component of the *CP*-model. (a) – magnitude of the components in different subjects; abscissa, age; ordinate, magnitude in conventional units; vertical dashed lines – borders of four age groups. (б) – average magnitude of the component in four age groups. (в) – topography of components (2) and lower (3) and upper (1) boundaries of confidence intervals. (г) – component signals in trials *a* – “A–A”, *b* – “A–P”, *c* – “P–P”. Abscissa, time; ordinate, signal magnitude in conventional units. The time points when the stimuli were turned on and off in a trial are shown with vertical dashed lines.

все-таки существуют. Кроме того, положение электродов относительно мозговых источников может варьировать от человека к человеку. Но, во-первых, используя метод независимых компонент, было показано, что для спонтанной ЭЭГ в пределах точности измерения топографии компонент приблизительно одинаковы [Ponomarev et al., 2014]. Во-вторых, считается, что источники электрического потенциала для ЭЭГ и длиннотентных компонент ПСС находятся в коре головного мозга [Hallez et al., 2007]. И, наконец, физические законы распространения электрических полей одинаковы для ЭЭГ и ПСС. Все эти факты указывают на то, что, несмот-

ря на то, что индивидуальные отличия топографий компонент ПСС существуют, они, вероятно, относительно невелики, и в первом приближении можно принять, что эти топографии одинаковы.

Во-вторых, предполагается, что сигналы компонент одинаковы у всех людей. Строго говоря, это предположение несправедливо для произвольной группы участников исследования. В частности, хорошо известно, что время реакции человека и пиковая латентность некоторых волн ПСС варьируют от человека к человеку и зависят от возраста. Эти свойства являются косвенным подтверждением того, что сигналы компонент также не-

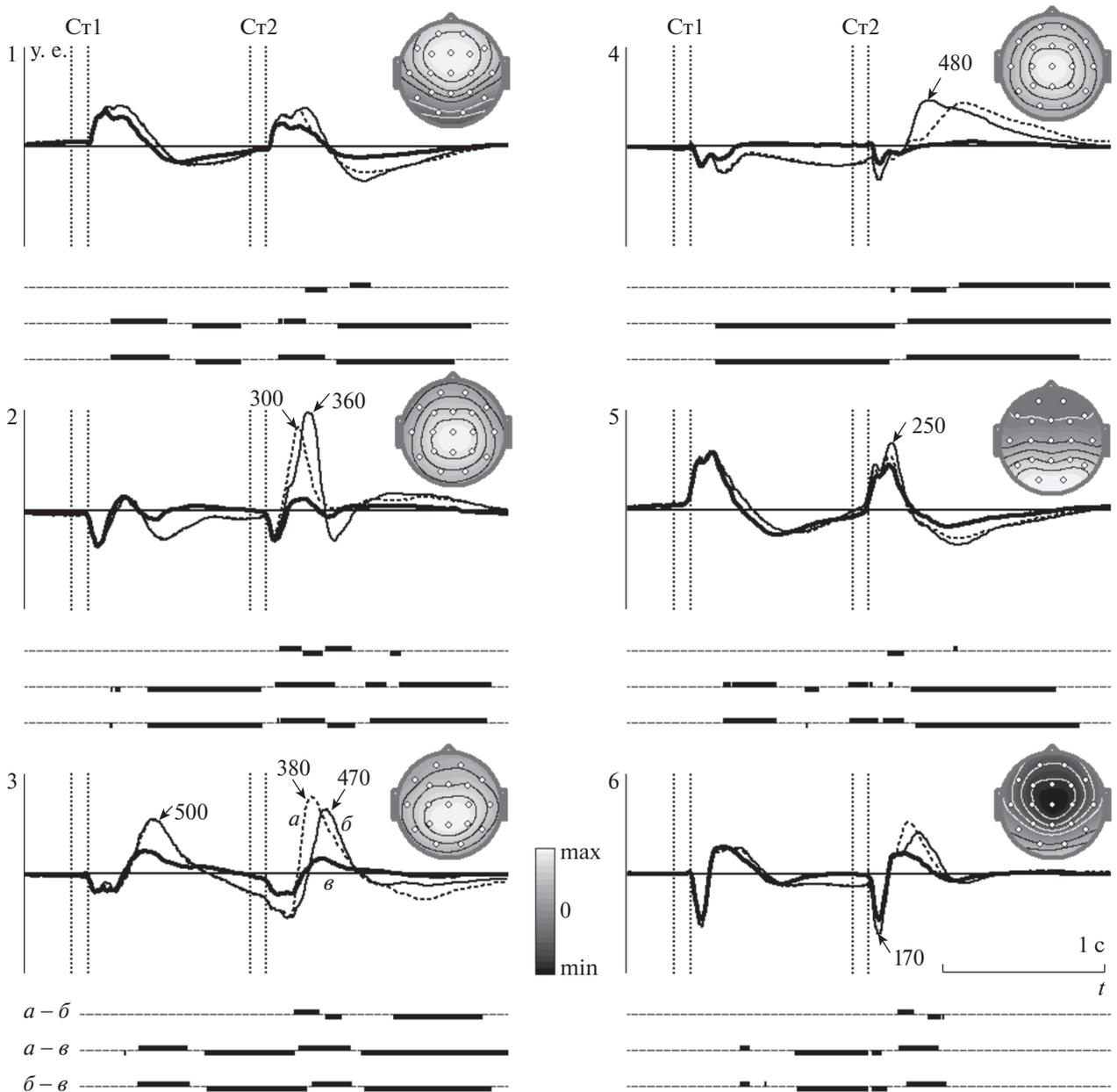


Рис. 4. Сигналы и топографии компонент в *CP*-модели. *a* – пробы “Ж–Ж” (*Go* условие), *b* – пробы “Ж–Р” (*NoGo* условие), *ε* – пробы “Р–Р”. По оси абсцисс – время, по оси ординат – величина сигнала в условных единицах. Вертикальные пунктирные линии – моменты включения и выключения стимулов в пробе. Среднегрупповое время реакции – 400 мс. Латентности пиков, отмеченных стрелками, – в миллисекундах. Интервалы времени, в течение которых доверительные интервалы (с надежностью 0.95) не перекрывались, выделены под графиками с помощью прямоугольников, расположенных в три строки для пар кривых *a* – *b*, *a* – *ε* и *b* – *ε* соответственно.

Fig. 4. Waveforms and topographies of the components in the *CP*-model. *a* – “A–A” trials (*Go* condition), *b* – “A–P” (*NoGo* condition), and *ε* – “P–P” trials. Abscissa, time; ordinate, signal magnitude in conventional units. The time points when the stimuli were turned on and off in a trial are shown with vertical dashed lines. The group average reaction time – 400 ms. The latencies of the peaks indicated with arrows are shown in milliseconds. The time intervals during which the confidence intervals (with level 0.95) did not overlap are highlighted under the graphs using rectangles located in three rows for pairs of waveforms *a* – *b*, *a* – *ε* and *b* – *ε*, respectively.

одинаковы. С другой стороны, широко используемый метод усреднения ПСС по группе участников исследования позволяет выде-

лить хорошо выраженные волны ПСС. Но это было бы невозможно, если бы индивидуальные различия ПСС были бы большими.

Кроме того, анализ зависимости пиковой латентности волн ПСС от возраста показывает, что существует возможность выбрать возрастной диапазон, в котором эта зависимость относительно невелика [Kropotov et al., 2016; Пономарев и др., 2019a]. Эти наблюдения позволяют считать, что при подходящем выборе возрастного диапазона в первом приближении можно принять, что сигналы компонент одинаковы.

В-третьих, предполагается, что топографии и величины компонент не зависят от условий. Такое предположение может показаться слишком строгим. Тем не менее, *CP*-модель допускает зависимость сигналов компонент от условия исследования. Более того, допустимыми являются даже такие случаи, когда при одном условии сигнал источника существует, а при другом он отсутствует. Эти аргументы указывают на то, что это предположение может оказаться справедливым. И это подтверждается результатами исследований, представленными выше. Тем не менее точность оценки сигналов уменьшается с ростом числа компонент в *CP*-модели. Поэтому включение в *CP*-модель большого числа условий, при которых выполняется различная деятельность, может сказываться на точности измерений, поскольку для описания таких ПСС может потребоваться большое число компонент.

Наконец, следует иметь в виду, что не для каждого массива данных и *CP*-модели с произвольным числом компонент можно получить единственное решение (подробнее см. [Kolda, Bader, 2009]). Это может накладывать ограничения на максимально допустимое число компонент в *CP*-модели. Тем не менее при анализе достаточно большого массива записей ПСС это ограничение не является слишком строгим.

Следует особо подчеркнуть, что предположения о схожести топографий и сигналов компонент могут быть приняты в первом приближении для здоровых испытуемых в ограниченном возрастном диапазоне и, возможно, для некоторых групп пациентов. Но у людей с различными поражениями мозга топографии и сигналы могут сильно различаться, и это ограничивает использование предложенного подхода.

Результаты сравнительного анализа эффективности работы 12 алгоритмов показали, что лучшую точность оценки параметров *CP*-

моделей можно получить, если, во-первых, в статистических моделях используются многомерные распределения, во-вторых, выборка значений скрытых параметров выполняется для каждой индивидуальной записи ПСС, каждого момента времени и каждого электрода отдельно, и, наконец, статистическая модель учитывает автокорреляцию сигналов. По-видимому, это непосредственно связано со спецификой ПСС, которые, во-первых, довольно сильно варьируют от человека к человеку, во-вторых, представляют собой нестационарный и относительно быстро изменяющийся во времени переходной процесс, имеющий значительную автокорреляцию, и, наконец, пространственное распределение электрического потенциала по скальпу имеет хорошо выраженные экстремумы, положение которых меняется во времени.

Отметим некоторые особенности данной статистической модели. Во-первых, в ней считается, что только элементы матрицы U не могут быть отрицательными. Напротив, в ранее предложенных моделях считалось, что элементы всех матриц могут быть либо произвольными [Xiong et al., 2010; Natarajan et al., 2014; Nakatsuji et al., 2016; Chen et al., 2019a], либо неотрицательными [Schmidt, Mohamed, 2009, Alquier, Guedj, 2017]. Во-вторых, при моделировании автокорреляции сигналов считается, что скрытый параметр точности Λ_t^S зависит от индекса t (времени), тогда как ранее предполагалось, что он постоянный [Xiong et al., 2010]. И, наконец, в данной модели учитывается дисперсия исходных данных.

Сравнивая *CP*-модель, представленную на рис. 4, с полученной ранее [Пономарев и др., 2019б, рис. 2] для тех же записей ПСС, видно, что большинство сигналов и топографий компонент довольно схожи, несмотря на то, что использовались различные методы оценки. Это косвенно подтверждает адекватность предлагаемого в данной работе подхода.

С другой стороны, использование байесовских статистических моделей предоставляет дополнительные возможности. Во-первых, такой подход позволяет оценить не только параметры *CP*-модели, но и соответствующие им доверительные интервалы, и, следовательно, выполнить статистические сравнения сигналов при различных условиях. Во-вторых, используя коэффициент Байеса, имеется возможность получить объективную оценку числа компонент в *CP*-модели. Наконец, сравнивая

метод градиентного спуска, использованный ранее [Пономарев и др., 2019 а,б], и предлагаемый в этой работе подход, необходимо отметить еще одно преимущество последнего. В частности, в зависимости от начального приближения метод градиентного спуска сходится к одному из локальных минимумов невыпуклой целевой функции. Поэтому для выбора наилучшей модели было необходимо многократно выполнять тензорное разложение, варьируя начальное приближение. Предложенный в данной работе метод оценки параметров *СР*-модели, использующий байесовские статистические модели, работает более стабильно. И, несмотря на то, что для верификации необходимо выполнить тензорное разложение несколько раз, использование этого подхода значительно уменьшает объемы вычислений.

Предложенный подход был разработан, прежде всего, для группового анализа ПСС. Он также подходит для анализа индивидуальных данных, но при этом следует учитывать ряд ограничений. Во-первых, соотношение сигнал–шум в единичных пробах меньше, чем у усредненных по пробам ПСС, что может сказаться на точности оценки топографий и сигналов компонент. Во-вторых, вариабельность величины различных компонентов ПСС от пробы к пробе может быть коррелирована, что может приводить к уменьшению ранга матрицы *U* и, как следствие, ограничивает возможность разделения сигналов этих компонентов. В-третьих, латентность некоторых компонентов ПСС и времени реакции человека коррелированы, причем время реакции варьирует от пробы к пробе. Если в зависимости от задания дисперсия времени реакции окажется слишком большой, то предположение о схожести сигналов в пробах станет несправедливым, и точность оценки сигналов будет снижена. Наконец, сигналы компонент могут быть оценены только с точностью до произвольного масштаба. Если оценки этих сигналов получены при анализе индивидуальных данных, то будет невозможно выполнить их количественное сравнение при последующем групповом анализе.

Хотя исследование функциональной роли компонент *СР*-модели не входило в задачу данного исследования, кратко остановимся на этом вопросе. Величина ранних пиков в сигналах компонент 1 и 6 после второго стимула больше в пробах “Ж–Ж” и “Ж–Р” по сравнению с “Р–Р”, что, по-видимому, свя-

зано с различным уровнем внимания [Kropotov, Ponomarev, 2015]. Величина пика в сигналах компоненты 5 с латентностью 250 мс больше в пробах “Ж–Р” по сравнению с “Ж–Ж” и “Р–Р”. Ранее было показано, что подобный эффект, вероятно, связан с процессом сравнения стимула с образом, хранящимся в рабочей памяти [Kropotov, Ponomarev, 2015]. Особенности сигналов компонент 2, 3 и 4 подробно обсуждались ранее, и был высказан ряд гипотез [Пономарев и др., 2019б]. В частности, возможно, что в сигналах компоненты 2 после второго стимула отражаются процессы, связанные с принятием решения о выполнении ответной реакции или ее подавлении; в сигналах компоненты 3 после первого и второго стимулов – с переключением с одного вида деятельности на другой; а в сигналах компоненты 4 после второго стимула – с восстановлением состояния нейронных сетей после действия. Иными словами, сигналы компонент по-разному зависят от типа проб, и в них отражается протекание различных мозговых процессов. Эти результаты указывают на то, что байесовский параллельный факторный анализ позволяет выделить функционально различные компоненты ПСС.

ЗАКЛЮЧЕНИЕ

Предложена байесовская вероятностная модель для параллельного факторного анализа связанных с событиями потенциалов мозга человека. Для этой модели была разработана процедура построения выборки случайных значений ее параметров, основанная на методах Монте-Карло с цепями Маркова. Использование этой процедуры позволяет, во-первых, получить объективную оценку числа компонент в модели и, во-вторых, выполнить статистические сравнения сигналов при выполнении различных видов деятельности. Результаты использования этой модели показали, что с помощью байесовского параллельного факторного анализа удастся выделить функционально различные компоненты ПСС, что, в свою очередь, позволяет рассматривать динамику процессов, протекающих в мозге, с нетрадиционной точки зрения.

Авторы выражают благодарность Dr. Andreas Müller, Детский центр, Кур, Швейцария за предоставленные записи ПСС в *Go/NoGo* тесте.

Работа выполнена в рамках государственного задания Минобрнауки по теме № АААА-А19-119101890066-2.

СПИСОК ЛИТЕРАТУРЫ

- Кропотов Ю.Д.* Количественная ЭЭГ, когнитивные вызванные потенциалы мозга человека и нейротерапия. Донецк: Издатель Заславский Ю.А., 2010. 512 с.
- Пономарев В.А., Кропотов Ю.Д.* Уточнение локализации источников вызванных потенциалов в GO/NOGO тесте с помощью моделирования структуры их взаимной ковариации. Физиология человека. 2013. 39 (1): 36–50.
- Пономарев В.А., Пронина М.В., Кропотов Ю.Д.* Параллельный факторный анализ в исследовании связанных с событиями потенциалов. Физиология человека. 2019а. 45 (3): 5–15.
- Пономарев В.А., Пронина М.В., Кропотов Ю.Д.* Скрытые компоненты связанных с событиями потенциалов в зрительном Go/NoGo тесте с предупреждающим стимулом. Физиология человека. 2019б. 45 (5): 20–29.
- Alquier P., Guedj B.* An Oracle Inequality for Quasi-Bayesian Non-Negative Matrix Factorization. *Mathematical Methods of Statistics*. 2017. 26 (1): 55–67.
- Chen X., He Z., Sun L.* A Bayesian tensor decomposition approach for spatiotemporal traffic data imputation. *Transp. Res. Part C Emerg. Technol.* 2019а. 98: 73–84.
- Chen X., He Z., Chen Y., Lu Y., Wang J.* Missing traffic data imputation and pattern discovery with a Bayesian augmented tensor factorization model. *Transp. Res. Part C Emerg. Technol.* 2019б. 104: 66–77.
- Chib S.* Marginal Likelihood from the Gibbs Output. *J. Am. Stat. Assoc.* 1995. 90 (432): 1313–1321.
- Cichocki A., Mandic D., Phan A.-H., Caiafa C.F., Zhou G., Zhao Q., De Lathauwer L.* Tensor decompositions for signal processing applications from two-way to multiway component analysis. *IEEE Signal Proc. Mag.* 2015. 32 (2): 145–163.
- Cong F., Lin Q.H., Kuang L.D., Gong X.F., Astikainen P., Ristaniemi T.* Tensor decomposition of EEG signals: a brief review. *J. Neurosci. Methods*. 2015. 248: 59–69.
- Gelman A., Carlin J.B., Stern H.S., Dunson D.B., Vehtari A., Rubin D.B.* Bayesian Data Analysis, Third Edition. New-York, London: CRC Press, 2014. 675 p.
- Hallez H., Vanrumste B., Grech R., Muscat J., De Clercq W., Vergult A., D'Asseler Y., Camilleri K.P., Fabri S.G., Van Huffel S., Lemahieu I.* Review on solving the forward problem in EEG source analysis. *J. Neuroeng. Rehabil.* 2007. 4: 46.
- Koch H., Bopp G.P.* Fast and Exact Simulation of Multivariate Normal and Wishart Random Variables with Box Constraints. arXiv preprint. 2019. arXiv: 1907.00057.
- Kolda T., Bader B.* Tensor decompositions and applications. *SIAM Review*. 2009. 51 (3): 455–500.
- Kropotov J.D., Ponomarev V.A.* Differentiation of neuronal operations in latent components of event-related potentials in delayed match-to-sample tasks. *Psychophysiology*. 2015. 52 (6): 826–838.
- Kropotov J., Ponomarev V., Tereshchenko E.P., Muller A., Jancke L.* Effect of aging on ERP components of cognitive control. *Front. Aging Neurosci.* 2016. 8: 69.
- Li Y., Ghosh S.K.* Efficient sampling methods for truncated multivariate normal and student-t distributions subject to linear inequality constraints. *J. Stat. Theory Pract.* 2015. 9: 712–732.
- Luck S.J., Kappenman E.S.* (Eds.) The Oxford handbook of event-related potential components. Oxford: Oxford University Press. 2011. 642 p.
- Nakatsuji M., Toda H., Sawada H., Zheng J.G., Hender J.A.* Semantic sensitive tensor factorization. *Artif. Intell.* 2016. 230: 224–245.
- Natarajan R., Banerjee A., Shan H.* Multiple imputation of missing data in multi-dimensional retail sales data sets via tensor factorization. US Patent: US08818919, Aug. 26, 2014/13/204,237
- Ponomarev V.A., Mueller A., Candrian G., Grin-Yatsenko V.A., Kropotov J.D.* Group independent component analysis (gICA) and current source density (CSD) in the study of EEG in ADHD adult. *Clin. Neurophysiol.* 2014. 125 (1): 83–97.
- Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P.* Numerical recipes. The art of scientific computing. Third edition. New York: Cambridge University Press. 2007. 1256 p.
- Salakhutdinov R., Mnih A.* Bayesian probabilistic matrix factorization using Markov chain Monte Carlo. *Proc. of 25th Int. Conf. on Machine Learning*. 2008. 880–887 pp.
- Schmidt M.N., Mohamed S.* Probabilistic non-negative tensor factorization using Markov chain Monte Carlo. *Proc. of 17th European Signal Proc. Conf.* 2009. 1918–1922 pp.
- Smith W.B., Hocking R.R.* Algorithm AS 53: Wishart variate generator. *J. R. Stat. Soc. Ser. C*. 1972. 21 (3): 341–345.
- Trinh G., Genz A.* Bivariate conditioning approximations for multivariate normal probabilities. *Stat. Comput.* 2015. 25: 989–996.
- Xiong L., Chen X., Huang T.K., Schneider J., Carbonell J.G.* Temporal collaborative filtering with Bayesian probabilistic tensor factorization. *Proc. of 2010 SIAM Int. Conf. on Data Mining*. 2010. 211–222 pp.

Zhou G., Zhao Q., Zhang Y., Adali T., Xie S., Cichocki A.
Linked component analysis from matrices to high

order tensors: applications to biomedical data.
Proc. IEEE. 2016. 104 (2): 310–331.

BAYESIAN PARALLEL FACTOR ANALYSIS FOR THE STUDY OF EVENT-RELATED POTENTIALS

V. A. Ponomarev^{a,#} and Yu. D. Kropotov^a

^a N.P. Bechtereva Institute of the Human Brain, Russian Academy of Sciences, St. Petersburg, Russia

[#] e-mail: valery_ponomarev@mail.ru

The aim of this study was to develop a Bayesian probabilistic model for parallel factor analysis of event-related potentials (ERP) of the human brain. Twelve statistical models are proposed that take into account the characteristics of the signals of ERP sources. For these models, Bayesian inference procedures were developed based on Markov chains Monte Carlo sampling. The effectiveness of these procedures was evaluated using both synthetic data with different signal-to-noise ratios and an array of ERPs obtained from 351 subjects in the *Go/NoGo* task. The procedure for obtaining estimates of the model parameters of the best accuracy was chosen. An analysis of the dependence of the model signals on the type of human activity performed showed that Bayesian parallel factor analysis is able to identify functionally different components of the ERP.

Keywords: event-related potentials, parallel factor analysis, *Go/NoGo* task