S Supporting Information

S.1 Dichotomization of FMRI Data

We derive a binary representation of fMRI voxel time series by considering for any given value the deviation from its median. The resulting binary, i.e., dichotomous, time series can be exploited to speed up subsequent analyses, as processing of binary data can be efficiently implemented based on bits. Furthermore, through dichotomization the data become addressable by specialized methods that are not applicable to the data in their original continuous-valued form [Loewe et al., 2013].

FMRI data are series of periodically acquired 3D images of continuous-valued intensities. One such series is denoted by \( \mathbf{i} = (i_1, i_2, \ldots, i_T) \), where \( T \) is the number of points in time at which the intensity images \( i_k \), \( k \in \{1, \ldots, T\} \), are recorded. The individual images are organized in a regular 3D voxel grid of size \( X \times Y \times Z \).

In order to simplify the processing, the voxel coordinates \((x, y, z)\) are mapped (in an essentially arbitrary, but fixed fashion) to a linear index \( v \) with \( 1 \leq v \leq V = X \cdot Y \cdot Z \). In this way a data set can be represented by a data matrix \( \mathbf{S}^{V \times T} = (\mathbf{s}_{v,k}) \).

By \( \mathbf{s}_v = s_{v,*} = (s_{v,1}, s_{v,2}, \ldots, s_{v,T}) \), that is, the \( v \)-th row of \( \mathbf{S} \), we denote the time series of voxel \( v \).

We use a simple binary discretization, i.e., a dichotomization, in order to assign to each value in a time series either ‘1’ or ‘0’. Formally, the dichotomized time series \( \mathbf{d}_v \) of a voxel \( v \) is given by \( \mathbf{d}_v = (d_{v,1}, d_{v,2}, \ldots, d_{v,T}) = (d(s_{v,1}), d(s_{v,2}), \ldots, d(s_{v,T})) \in \{0, 1\}^T \), induced by the function \( d(s_{v,k}) = \mathbf{H}(s_{v,k} - \tilde{s}_v) \), \( k \in \{1, \ldots, T\} \), where \( \tilde{s}_v \) denotes the median of the values in \( s_v \) and \( \mathbf{H} \) is the Heaviside step function, defined as \( \mathbf{H}(x) = 0 \) if \( x < 0 \) and \( \mathbf{H}(x) = 1 \) otherwise. In other words, \( d_{v,k} \) is equal to ‘1’ if the signal intensity value \( s_{v,k} \) amounts at least to the median of its corresponding time series \( s_v \).

Analogous to the matrix \( \mathbf{S}^{V \times T} \), the corresponding dichotomized matrix is denoted by \( \mathbf{D}^{V \times T} = (d_{v,k}) \), and \( d_v = d_{v,*} \) denotes its \( v \)-th row.

S.2 Mathematical Derivation

Due to the dichotomization scheme used, the lower integration limits of

\[
p_{11} = \int_{\Phi^{-1}(p_{\mathbf{s}*})}^{\infty} \int_{\Phi^{-1}(p_{\mathbf{b}*})}^{\Phi^{-1}(p_{\mathbf{x}*})} f(z_x, z_y) dz_x dz_y.
\]

become zero, as \( p_{\mathbf{b}*} = p_{\mathbf{s}*} = p_{\mathbf{x}*} = 0.5 \) and thus \( \Phi^{-1}(p_{\mathbf{b}*}) = \Phi^{-1}(p_{\mathbf{s}*}) = \Phi^{-1}(0.5) = 0 \). Transformation from cartesian to polar coordinates through the substitutions

\[
z_x = r \cos \varphi
\]

\[
z_y = r \sin \varphi
\]

\[
z_x^2 + z_y^2 - 2r_t z_x z_y = r^2(1 - 2r_t \cos \varphi \sin \varphi) = r^2(1 - r_t \sin(2\varphi))
\]

\[
dz_x dz_y = rdrd\varphi
\]

yields

\[
p_{11} = \frac{1}{2\pi \sqrt{1 - r_t^2}} \int_0^{\infty} \int_0^{\infty} r \exp \left( -\frac{r^2(1 - r_t \sin(2\varphi))}{2(1 - r_t^2)} \right) dr d\varphi
\]

\[
= \frac{1}{2\pi \sqrt{1 - r_t^2}} \int_0^{\infty} \left[ \frac{-(1 - r_t^2)}{1 - r_t \sin(2\varphi)} \exp \left( -\frac{r^2(1 - r_t \sin(2\varphi))}{2(1 - r_t^2)} \right) \right]_0^{\infty} d\varphi
\]

\[
= \frac{1}{2\pi} \int_0^{\infty} \frac{d\varphi}{1 - r_t \sin(2\varphi)}.
\]

Using

\[
\int \frac{d\varphi}{1 - r_t \sin(2\varphi)} = \frac{1}{\sqrt{1 - r_t^2}} \arctan \left( \frac{\tan \varphi - r_t}{\sqrt{1 - r_t^2}} \right) + \text{const}.
\]
we proceed with
\[
\begin{align*}
  p_{11} &= \frac{\sqrt{1-r_t^2}}{2\pi} \left[ \frac{1}{\sqrt{1-r_t^2}} \arctan \left( \frac{\tan \varphi - r_t}{\sqrt{1-r_t^2}} \right) \right]_0^\pi \\
  &= \frac{\sqrt{1-r_t^2}}{2\pi} \left[ \frac{\pi}{2\sqrt{1-r_t^2}} - \frac{1}{\sqrt{1-r_t^2}} \arctan \left( -\frac{r_t}{\sqrt{1-r_t^2}} \right) \right] \\
  &= \frac{1}{2\pi} \arccos(-r_t).
\end{align*}
\]
(1)

By solving for \( r_t \), we obtain
\[
r_t = -\cos(2\pi p_{11}).
\]
(2)

The same result also arises as a by-product of Pearson’s holistic examination of the problem within the framework of series expansion [Pearson, 1900] (page 7, 20th equation). In contrast, the derivation we presented here focuses only on the special case at hand (where all marginal probabilities are equal) and may thus be more easily comprehensible.

S.3 Supplementary Figures

![Figure S.1](image1.png)

The first three splits into smaller and smaller blocks in the recursion of our cache-oblivious algorithm.

![Figure S.2](image2.png)

Exploiting SIMD instructions for the correlation computation (SSE2 or AVX, four sums in parallel).

![Figure S.3](image3.png)

Both \( r \) and \( r_t \) are biased estimators of \( \rho \). MSD(\( r, \rho \)) and MSD(\( r_t, \rho \)) for \( T = 100 \) and \( T = 300 \).
References
