Comparison of whole-brain task-modulated functional connectivity methods for fMRI task connectomics

Ruslan Masharipov*, Irina Knyazeva, Alexander Korotkov, Denis Cherednichenko, Maxim Kireev

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

*Corresponding author: Ruslan Masharipov
E-mail: masharipov@ihb.spb.ru

Abstract

Higher brain functions require flexible integration of information across widely distributed brain regions depending on the task context. Resting-state functional magnetic resonance imaging (fMRI) has provided substantial insight into large-scale intrinsic brain network organisation, yet the principles of rapid context-dependent reconfiguration of that intrinsic network organisation are much less understood. A major challenge for task connectome mapping is the absence of a gold standard for deriving whole-brain task-modulated functional connectivity matrices. Here, we performed biophysically realistic simulations to control the ground-truth task-modulated functional connectivity over a wide range of experimental settings. We revealed the best-performing methods for different types of task designs and their fundamental limitations. Importantly, we found that rapid (100 ms) modulations of oscillatory neuronal synchronisation can be recovered from sluggish haemodynamic fluctuations even at typically low fMRI temporal resolution (2 s). Finally, we provide practical recommendations on task design and statistical analysis to foster task connectome mapping.

Main

Building a comprehensive map of human brain connections, called the human connectome[1], can be considered one of the largest and most challenging scientific projects in the field of human neuroscience over the past two decades. In the past decade, the trend in fMRI studies has shifted from functional segregation to functional integration, which is reflected in an increasing number of publications on functional connectivity (FC) compared to task activations (Fig. 1a). Despite the initial scepticism about resting-state FC (RSFC)[2,3], it has become the most popular fMRI approach and has profoundly advanced our understanding of the intrinsic functional organisation of the brain in health[4,5] and disease[6,7]. Furthermore, there is a growing awareness of the importance of whole-brain FC dynamics modulated extrinsically by various task demands, called the *task connectome*[8].

Different FC types can be assessed from the task-state blood oxygen-level-dependent (BOLD) signal, ranging from more intrinsic to more extrinsic FC (Fig. 1b). The simplest approach is to correlate the whole time series, similar to the RSFC calculation. This type of FC can be called task-state FC (TSFC). Three neuronal sources of variability underlie TSFC: spontaneous task-independent (intrinsic) oscillations, task-modulated (extrinsic) oscillations, and co-activations caused by simultaneous activations without communication between brain regions. As co-activations can spuriously increase FC estimates, it has been proposed to regress out task activations from the task-state BOLD signal and correlate the residuals, which is referred to as background FC (BGFC)[9]. TSFC and BGFC are typically very similar to RSFC since all are driven mainly by spontaneous oscillations[10]. In contrast, task-modulated FC (TMFC) reflects dynamic changes in FC during one condition compared to another, eliminating the influence of spontaneous task-independent oscillations and co-activations. Several TMFC methods have been proposed, including direct correlation difference (CorrDiff)[11], standard[12], generalised[13] and correlational[14] forms of psychophysiological interaction (sPPI, gPPI, cPPI) with and without deconvolution procedure[15], and beta-series correlations (BSC) based on least-squares all

(LSA)[16] and least-squares separate (LSS)[17] procedures. Each TMFC approach has its own advantages and limitations. The lack of a gold standard for deriving whole-brain TMFC matrices and limited knowledge about the fundamental limitations of TMFC estimation related to sluggishness of the BOLD signal hinder the process of task connectome mapping.

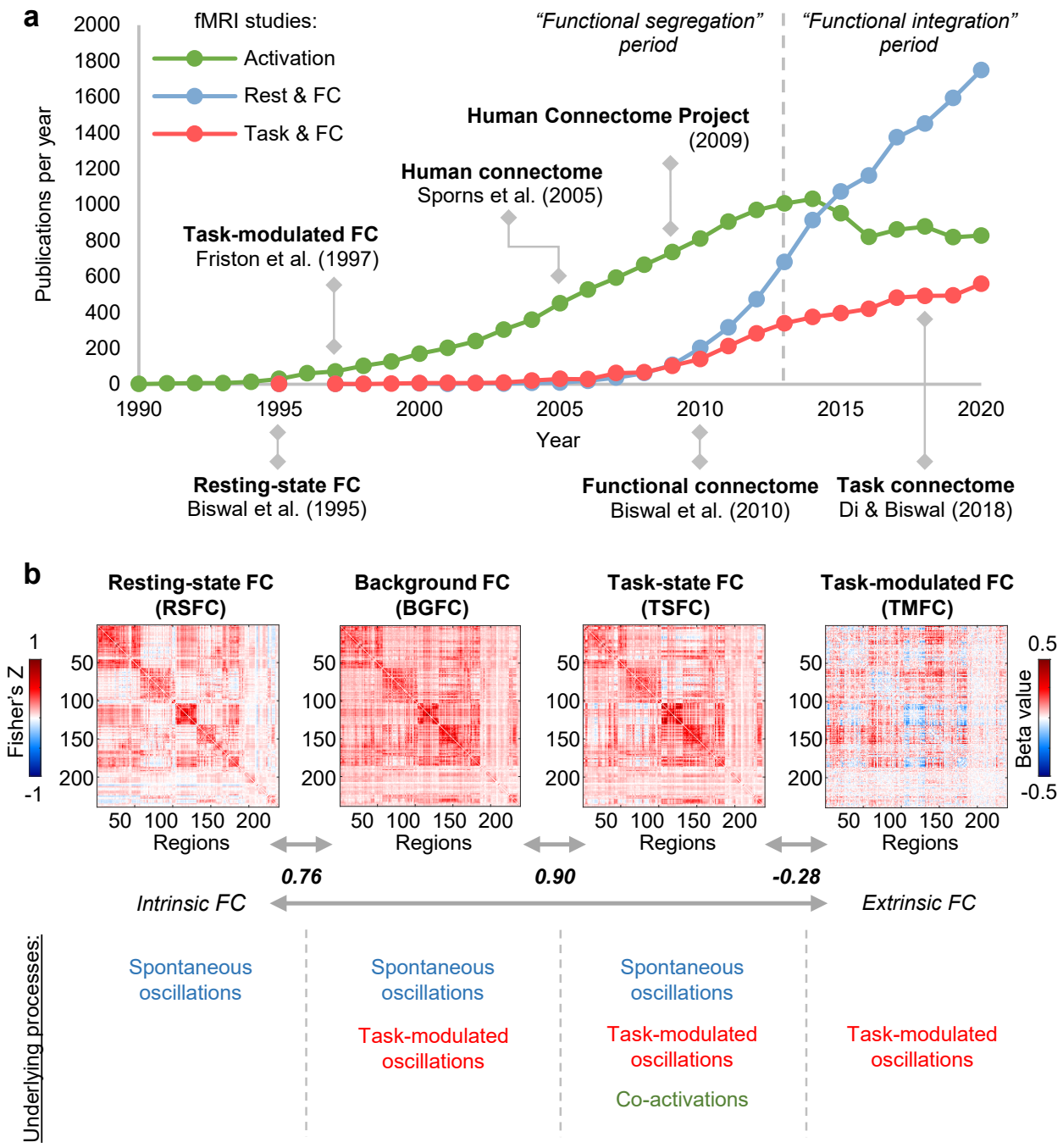


Fig. 1. Overview of FC types. (**a**) fMRI publications per year mentioning "activation" (green), "functional connectivity" and "rest" (red), "functional connectivity" and "task" (blue). Data obtained via PubMed search from 2000 to 2020. Here, we note the trend shift in the fMRI field from studies of functional segregation (activation studies) to functional integration (connectivity studies) since 2013. We also observe the dominance of RSFC studies since 2015. (**b**) Illustration of various types of FC matrices along the intrinsic-extrinsic axis. To calculate FC matrices, we used resting-state and working memory task data from the Human Connectome Project. To estimate TMFC, we applied the gPPI method with the deconvolution procedure ("2-back > 0-back" contrast). To evaluate the similarity between matrices, we used Pearson's r correlation.

Previous simulation studies evaluating TMFC methods have two main limitations. First, they often ignore the dynamic, oscillatory nature of neuronal population activity arising from interactions between excitatory and inhibitory neurons by using simple delta or boxcar functions to simulate neuronal activity

in a pair of regions[13,15,18,19]. Second, studies with biophysically realistic simulations based on neural mass models consider only block designs and a limited range of TMFC methods[20,21]. Meanwhile, one of the key unresolved questions is the performance of different TMFC methods for fast event-related designs with slow data acquisition, which are typical for the majority of fMRI studies. Thus, to date, there are no biophysically realistic simulation studies systematically comparing existing TMFC methods in different experimental fMRI setups.

Here, using biophysically realistic simulations and empirical data, we determined the best-performing whole-brain TMFC methods for different fMRI task designs and identified their limitations. To simulate the neuronal dynamics of multiple interconnected brain regions, we applied a large-scale Wilson-Cowan neural mass model consisting of 100 excitatory-inhibitory units[22,23]. The BOLD signals were generated from simulated neuronal activity using the Balloon-Windkessel haemodynamic model[24]. To control the ground-truth TMFC, we manipulated synaptic weights between neural mass units depending on the task context, which corresponded to short-term synaptic plasticity[25,26,27].

First, we demonstrate that one popular TMFC method, cPPI, is not capable of estimating TMFC. Second, we establish that TMFC methods are susceptible to spurious inflation of FC due to co-activations in event-related designs even more than in block designs[21]. Third, we find that the most sensitive methods for rapid event-related designs and block designs are sPPI and gPPI with a deconvolution procedure, while for all other designs, the best method is BSC-LSS. Despite the scepticism regarding the deconvolution procedure[28], we demonstrate that deconvolution prominently increases the sensitivity of the PPI methods in both event-related and block designs. Fourth, while some authors classify PPI as a FC method[28,29] and others see PPI as a simple regression model of effective connectivity[30,31,32,33,34,35], our results explicitly demonstrate that PPI with deconvolution can, in principle, provide information about the direction of causal influence at high signal-to-noise ratios (SNRs) and long scan times. However, in most studies, the asymmetry of PPI matrices is likely to spuriously arise from a low SNR, short event duration, small number of events per condition, small sample size, and long repetition time (TR).

Finally, we demonstrate that rapid (100 ms) task-related modulation of gamma-band neuronal synchronisation can be uncovered from ultra-slow BOLD-signal oscillations, even with typically slow data acquisition (TR = 2 s). Meanwhile, recently developed fast FMRI sequences (TR < 1 s) yield increased sensitivity of TMFC methods not only by increasing the amount of data but also due to more precise insights into fast neuronal dynamics hidden behind the sluggish haemodynamic processes.

Results

To compare different TMFC methods, we first used empirical fMRI data from the Human Connectome Project (HCP)[36] and the Consortium for Neuropsychiatric Phenomics (CNP)[37]. Next, we performed a series of simulations to compare the sensitivity and specificity of TMFC methods in experiments with block and event-related designs, different SNRs, sample sizes, duration of events, mean inter-stimulus intervals (ISIs) and number of events. For block designs, we considered the CorrDiff approach and PPI methods, and for event-related designs, we considered the PPI methods and BSC methods. Then, we compared the effectiveness of fast data acquisition for a fixed number of scans and fixed total scan time. Finally, we performed simulations with symmetric (undirected) and asymmetric (directed) ground-truth synaptic matrices to identify sources of PPI matrix asymmetry.

Empirical comparison of TMFC methods

We found that all TMFC methods, except cPPI, produce similar unthresholded matrices for the block design (working memory task, Fig. 2a). The correlations between CorrDiff matrices and symmetrised sPPI and gPPI matrices ranged from 0.66 to 1. We also found that the correlations between these TMFC matrices and RSFC, BGFC, TSFC and cPPI matrices ranged from -0.13 to -0.28.

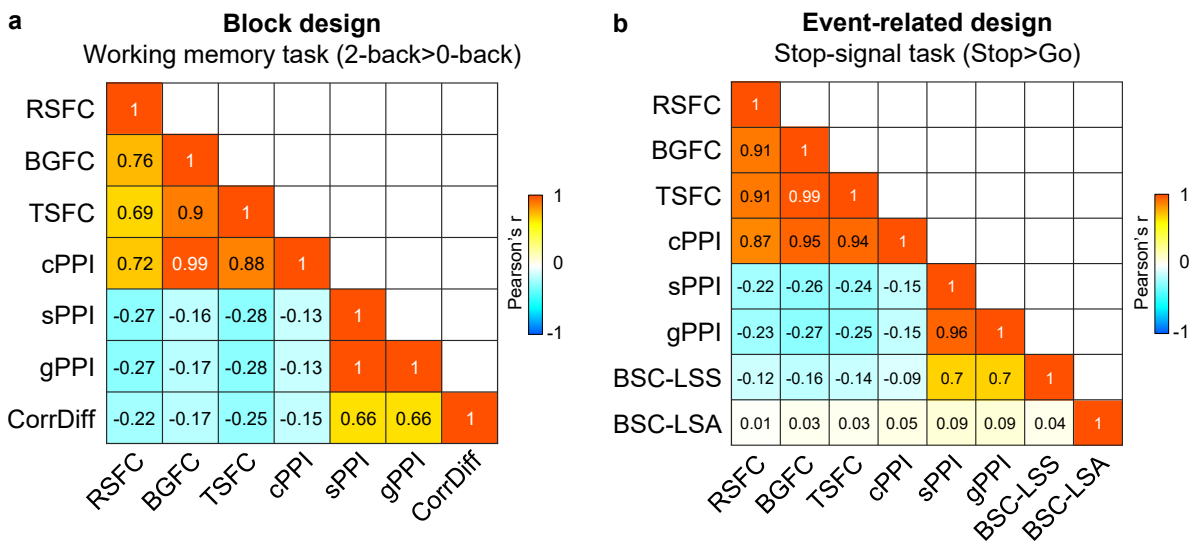


Fig. 2. Correlations between unthresholded RSFC, BGFC, and TSFC matrices and FC matrices obtained by different TMFC methods. To evaluate the similarity between the raw FC matrices, we calculated Pearson's r correlations between lower diagonal elements. sPPI and gPPI matrices were symmetrised. All PPI terms were calculated with the deconvolution step. **(a)** Results for the block design: working memory task. **(b)** Results for the event-related design: stop-signal task.

For the event-related design (stop-signal task), similar TMFC matrices were produced by the sPPI, gPPI and BSC-LSS methods, with correlation coefficients ranging from 0.7 to 0.96 (Fig. 2b). At the same time, correlations between these TMFC matrices and the RSFC, BGFC, TSFC and cPPI matrices ranged from -0.09 to -0.27. Remarkably, the BSC-LSA approach produced a random-like matrix correlated with all other FC matrices with a correlation of 0.01–0.09. Similar results were obtained for empirical data from other tasks with block and event-related designs (social cognition task and task-switching task), which demonstrates that the observed effects are specific to a design type rather than a particular task (Supplementary Fig. S1).

Notably, the matrices generated by the cPPI method are similar to matrices obtained using the RSFC, BGFC, and TSFC approaches, which mainly reflect task-unrelated neuronal oscillations[10]. Correlations between the cPPI, RSFC, BGFC and TSFC matrices ranged from 0.69 to 0.99 for block and event-related design (Fig. 2a, b). The partial correlation between two PPI terms controlling for physiological (BOLD signal in both regions) and psychological variables (task regressor), as implemented in cPPI, was as high as the simple correlation between these two PPI terms and the correlation between physiological regressors (Extended Data Fig. 1). Therefore, we concluded that the cPPI approach is unable to separate task-modulated (extrinsic) from task-independent (intrinsic) sources of FC. However, comparison of FC matrices based solely on empirical data does not allow us to say which method better reflects TMFC. To answer this question, we applied a biophysically realistic simulation approach that enabled us to control the ground-truth TMFC.

Large-scale neural mass simulations

Our simulation approach was based on the coupled oscillator model for FC[38], according to which FC measured by fMRI arises from a correlation between ultra-slow BOLD oscillations caused by ultra-slow modulation of the envelopes of synchronised gamma-band oscillations (Fig. 3a, b). For details about the simulation procedures, see Methods. The matrices of gamma-band neuronal synchronisation measured by the phase-locking value were closely matched to synaptic weight matrices (Fig. 3c, d). In accordance with previous empirical observations[38], we observed strong correlations between simulated ultra-slow fluctuations of the gamma-band envelope and time-shifted BOLD signal. For the event-related design, the simulated BOLD signal was correlated with the gamma-band envelope at r = 0.79 with a 3.5 s time lag (Fig. 3e-j). Similar results were obtained for the block design, where the correlation was r = 0.81 with a 3.5 s time lag (Supplementary Fig. S2).

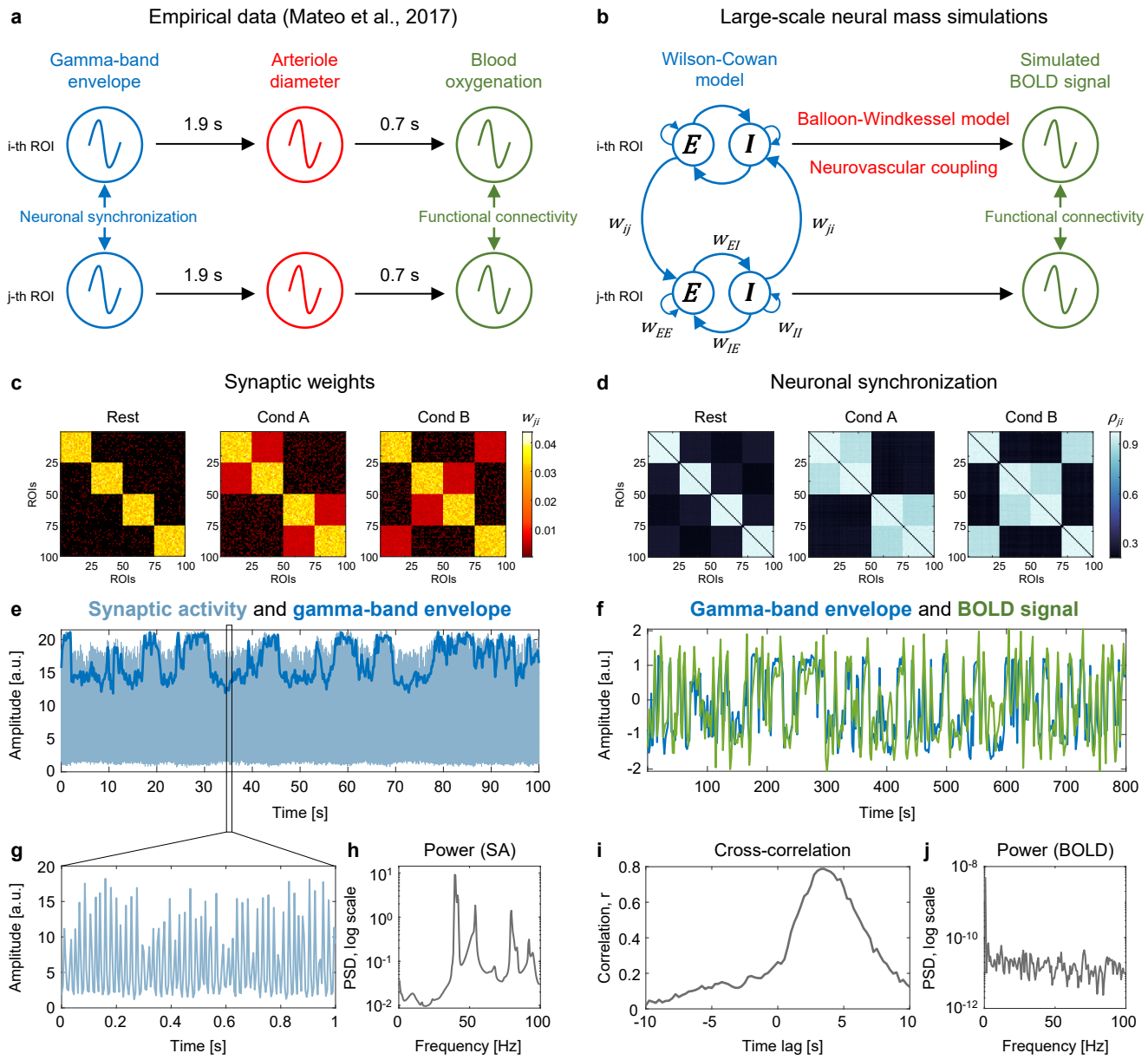


Fig. 3. The emergence of ultra-slow BOLD signal fluctuations from ultra-slow modulations of fast oscillatory activity. **(a)** Coupled oscillator model for functional connectivity[38]. **(b)** Scheme for generating the BOLD signal from gamma-band oscillations simulated by the Wilson-Cowan model. **(c)** Ground-truth symmetric synaptic weight (w_{ji}) matrices for a single subject. For the long-range connections between Wilson-Cowan units (w_{ji}), we used three synaptic weight matrices corresponding to two task conditions ("Cond A" and "Cond B") and interim "Rest" periods. Each synaptic weight matrix consisted of 100 brain regions, which corresponded to the minimum number of areas in brain atlases[39], and four functional modules, given the presence of such modules in the human brain[5]. The highest synaptic weights were within each module during the "Rest" periods[5]. Under the task conditions, we increased synaptic weights between functional modules relative to the "Rest" periods[8]. **(d)** Gamma-band neuronal synchronisation estimated by the phase-locking value (ρ_{ji}) for a single subject. The simulation was performed for the event-related design with one hundred 1 s events per condition and ISI = 6 s. **(e)** Example of simulated synaptic activity (SA) and gamma-band envelope for one of 100 connected brain regions (ROIs). **(f)** The time series of the gamma-band envelope and BOLD signal generated by the Balloon-Windkessel model based on simulated SA. We used the standard parameters of the haemodynamic model[24], which have previously been used in whole-brain RSFC simulation studies[23,40]. **(g)** One second of simulated SA. **(h)** Power spectral density (PSD) of simulated SA. The main peak at 40 Hz (gamma-band oscillations). **(i)** Cross-correlation between the gamma-band envelope and BOLD signal. The maximum correlation $r = 0.79$ corresponds to a time lag of 3.5 seconds. **(j)** Power spectral density of the BOLD signal.

The cPPI method fails to estimate TMFC

We first considered simulations without co-activations to investigate whether different TMFC methods produce FC matrices similar to ground-truth synaptic weight matrices for a sample size N=100, SNR = 0.4, and TR = 2 s. For differences between simulated conditions ("Cond A-B"), almost all tested

methods produced TMFC matrices similar to the ground truth (Fig. 4a-f). In contrast, the cPPI method produced matrices similar to BGFC and TSFC (Fig. 4g), which reproduced the above-mentioned findings from the empirical data analysis (Fig. 2). Therefore, we excluded the cPPI method from further analysis, as it is unable to assess TMFC.

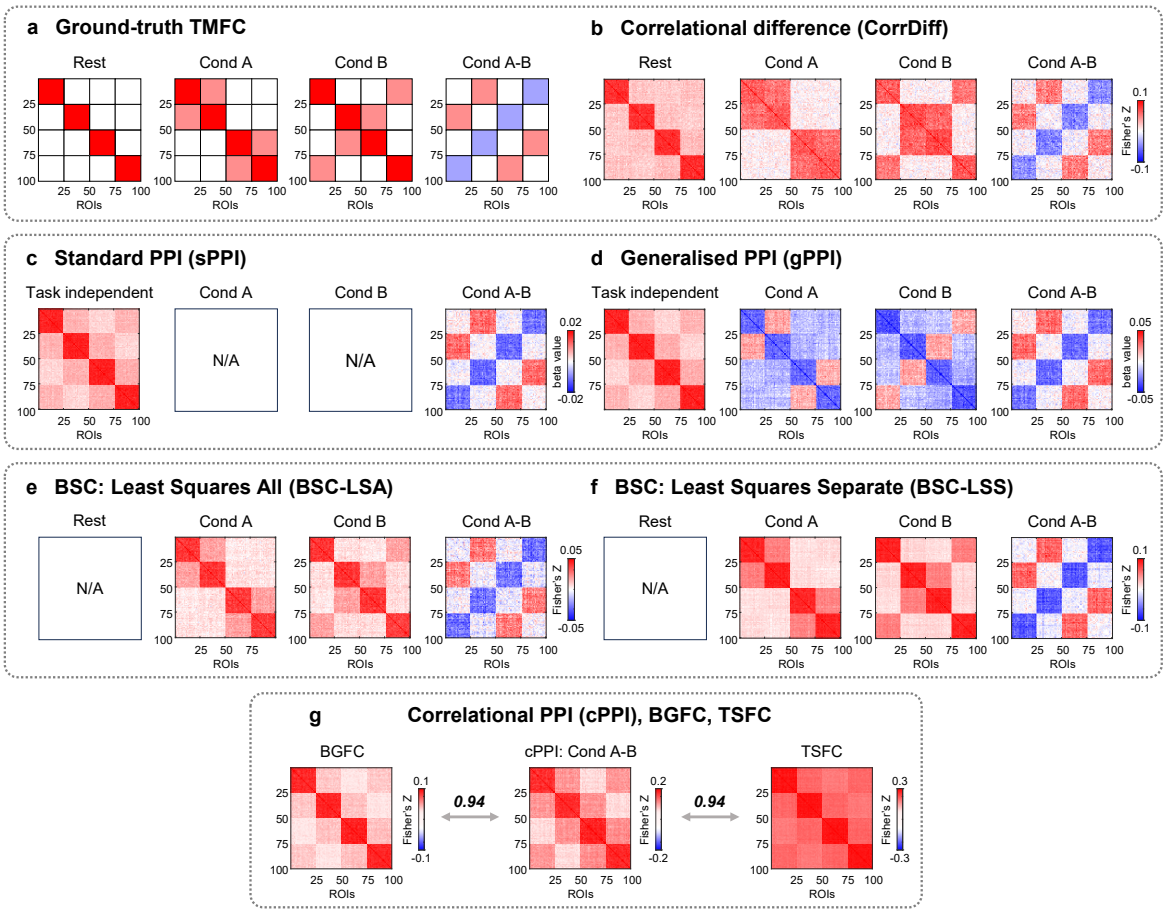


Fig. 4. Illustration of FC matrices produced by different TMFC methods based on simulated data. An example of simulation results for sample size N = 100, SNR = 0.4 and TR = 2 s. **(a)** Expected FC matrices based on ground-truth synaptic weight matrices. **(b-d)** For the CorrDiff, sPPI and gPPI methods, we considered the block design with ten 20 s blocks per condition. **(e-f)** For the BSC-LSA and BSC-LSS methods, we considered the event-related design with one hundred 1 s events per condition and mean ISI = 6 s. **(g)** The cPPI, BGFC and TSFC matrices were calculated based on the block design simulation. Analogous results were obtained for the event-related design simulation. To evaluate the similarity between matrices, we calculated Pearson's r correlations between lower diagonal elements. All PPI terms were calculated using the deconvolution step. sPPI and gPPI matrices were symmetrised.

Despite its simplicity, the CorrDiff approach allows the computation of easily interpretable FC matrices separately for task and rest blocks; these matrices adequately reflect the underlying ground-truth synaptic weight matrices (Fig. 4a, b). The sPPI and gPPI approaches enable estimation of both TMFC and task-independent FC (Fig. 4c, d). Task-independent FC is based on beta coefficients for physiological regressors delivered from selected ROIs and is similar to FC during rest periods. In addition, the gPPI approach can be used to calculate FC separately for each of the task conditions (i.e., Condition > Baseline). The BSC-LSA and BSC-LSS approaches also enable the calculation of FC matrices for each of the task conditions but not for rest periods (task-independent FC) since rest periods are usually not modelled explicitly (Fig. 4e, f).

Co-activations spuriously inflate TMFC estimates

Next, we considered simulations with co-activations to investigate how different TMFC methods address artificial inflation of TMFC estimates due to simultaneous activation of brain regions without task-related modulation of synaptic weights between them (Fig. 5a-c). Specificity (true positive rate,

TPR) and sensitivity (true negative rate, TNR) were calculated based on TMFC matrices thresholded at $\alpha = 0.001$ with false discovery rate (FDR) correction (Supplementary Information 1).

As a result, we found that if co-activations are not removed from the fMRI time series before TMFC analysis, they spuriously inflate TMFC estimates in three cases. First, the sPPI method practically does not eliminate co-activations and therefore has near zero specificity in both block and event-related designs (Fig. 5f-g). Second, the gPPI method demonstrates low specificity when applied to event-related designs without deconvolution (Fig. 5d, g). Third, the specificity of all TMFC methods decreases if the design matrix is not upsampled before convolution with the haemodynamic response function (Fig. 5f,g).

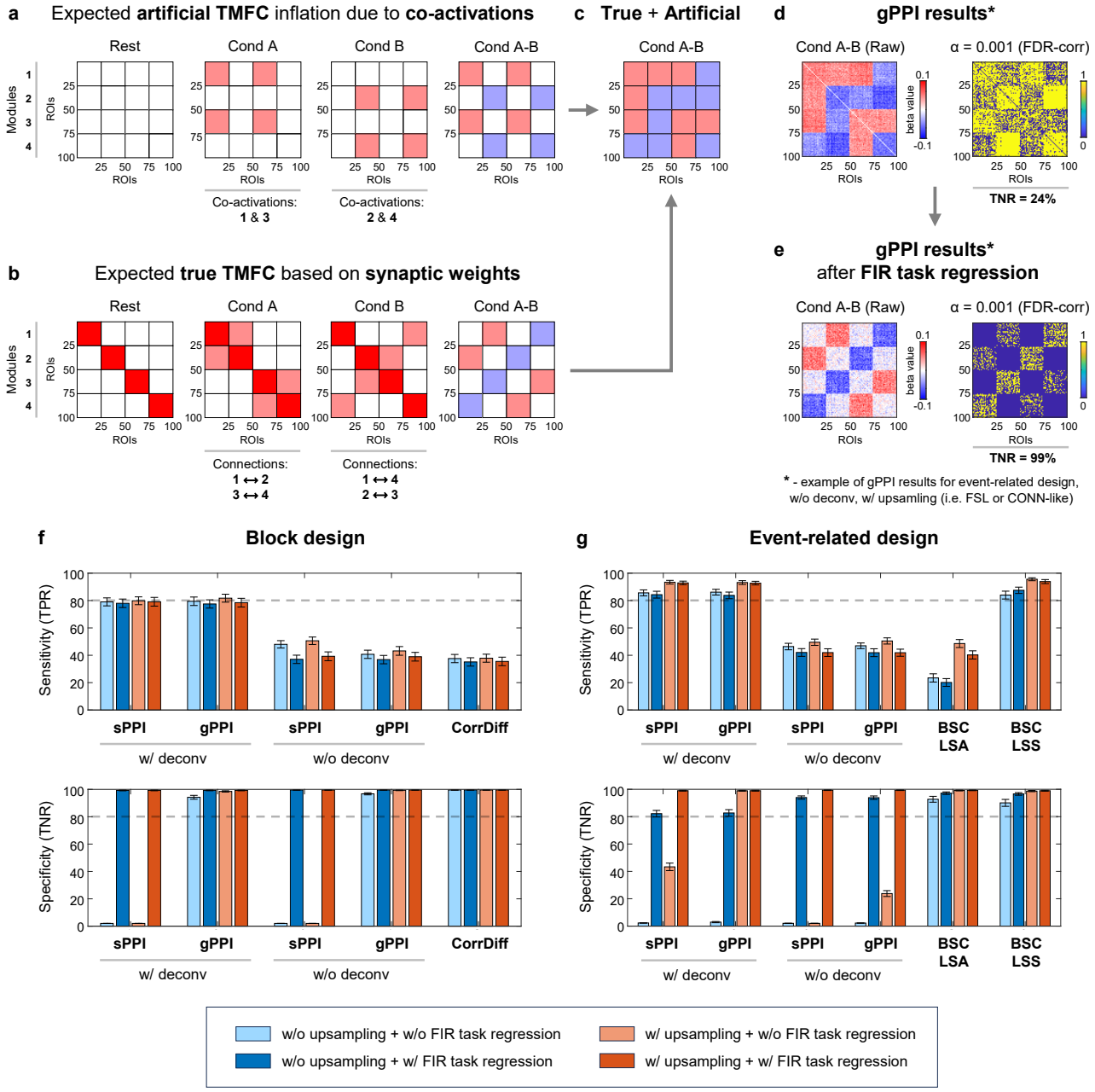


Fig. 5. Inflation of TMFC estimates due to co-activations. Simulation results for sample size N = 100, SNR = 0.4 and TR = 2 s. **(a)** Expected influence of co-activations on FC estimates (artificial TMFC). To model co-activations, we added simultaneous haemodynamic responses for different functional modules to the simulated BOLD signal without changing the synaptic weights between them. Functional modules №1 and №3 are co-activated in "Cond A", while modules №2 and №4 are co-activated in "Cond B". Artificial FC inflation was expected within and between these modules. **(b)** Expected FC matrices based on ground-truth synaptic weight matrices (true TMFC). In "Cond A", synaptic weights were increased between modules №1 and №2 and modules №3 and №4. In "Cond B", synaptic weights were increased between modules №1 and №4 and modules №2 and №3. **(c)** If the TMFC method fails to eliminate co-activations, we will observe FC changes between all ROIs (artificial + true TMFC). **(d-e)** Raw and thresholded TMFC matrices obtained using the gPPI method

without (w/o) the deconvolution step and with (w/) design matrix upsampling, similar to the gPPI implementation in the FSL or CONN toolbox. TMFC matrices were thresholded at $\alpha = 0.001$ (two-sided one-sample t test, false discovery rate (FDR) correction). **(d)** Without task activation regression, we observed FC changes between almost all ROIs (low specificity). **(e)** After finite impulse response (FIR) task regression, artificial TMFC was removed, leaving mostly true TMFC (high specificity). **(f)** Sensitivity (true positive rate, TPR) and specificity (true negative rate, TNR) of different TMFC methods for the block design with ten 20 s blocks per condition. **(g)** TPR and TNR for the event-related design with one hundred 1 s events per condition and men ISI = 6 s. All TMFC matrices were calculated with and without FIR task regression, as well as with and without upsampling of the design matrix before convolution. PPI terms were calculated with and without the deconvolution step. sPPI and gPPI matrices were symmetrised. Error bars represent SD and were computed from 1000 random resamplings with replacement.

Upsampling of the design matrix is necessary to improve the convolution procedure and is implemented in many popular neuroimaging software packages (SPM, FSL, AFNI, CONN toolbox). However, it may be absent in some in-house TMFC analysis scripts. The most prominent effect of upsampling on specificity can be seen for the gPPI method with deconvolution in event-related designs (Fig. 5g).

To better isolate TMFC from co-activation effects, it has been proposed to regress out task activations using finite impulse response (FIR) functions prior to TMFC analysis[21]. FIR task regression substantially improved the specificity of all TMFC methods in both block and event-related designs (Fig. 5f, g). For instance, the gPPI method without deconvolution had a specificity of 24% before FIR task regression and 99% after (Fig. 5e). The downside of FIR task regression is that it slightly reduces sensitivity, most notably for the sPPI and gPPI methods without deconvolution and the BSC-LSA method.

In all subsequent sections, we consider simulations with co-activations and perform TMFC analyses with design matrix upsampling and FIR task regression. We will not report specificity further since in no case did it fall below 95%.

Influence of noise and sample size on the sensitivity of TMFC methods

In this section, we describe the robustness of TMFC methods to high noise levels and low sample sizes. For the block design, the PPI methods with deconvolution were the most sensitive and robust to noise (Fig. 6a, Extended Data Table 1). The least sensitive method was CorrDiff since it needed a sample size > 100 to achieve > 80% sensitivity at high SNR = 0.5, while the PPI methods with deconvolution needed a sample size > 50 (Fig. 6c).

For the event-related design, the BSC-LSS method was the most sensitive (Fig. 6b, Extended Data Table 2). The difference between the BSC-LSS method and PPI methods with deconvolution was relatively small and more pronounced at high noise levels (SNR < 0.40). In contrast, the BSC-LSA method had the lowest sensitivity because of multicollinearity problems[17]. The BSC-LSS method, as well as the PPI methods with deconvolution, needed a sample size N > 50 to achieve > 80% sensitivity at SNR = 0.5 (for the event-related design), while the BSC-LSA method needed a sample size N > 100 (Fig. 6d).

In contrast to previous simulation studies[13,18], we did not detect a noticeable increase in the sensitivity of the gPPI method compared to the sPPI method. Additional Bayesian analysis provided evidence for the absence of differences between these methods for the block and event-related designs (Extended Data Tables 1 and 2).

It has also been previously suggested that deconvolution may benefit only event-related designs and can be omitted for block designs[8]. Indeed, some popular neuroimaging packages implement the PPI method without the deconvolution step (e.g., FSL and CONN toolbox). Here, we show that the deconvolution procedure substantially increases the sensitivity of the PPI methods in both block and event-related designs (Fig. 6a, b). Without deconvolution, these methods failed to achieve > 80% sensitivity for sample sizes N < 100 and SNR = 0.5 (Fig. 6c, d). Therefore, in the remaining sections, we will consider the sPPI and gPPI methods only with deconvolution.

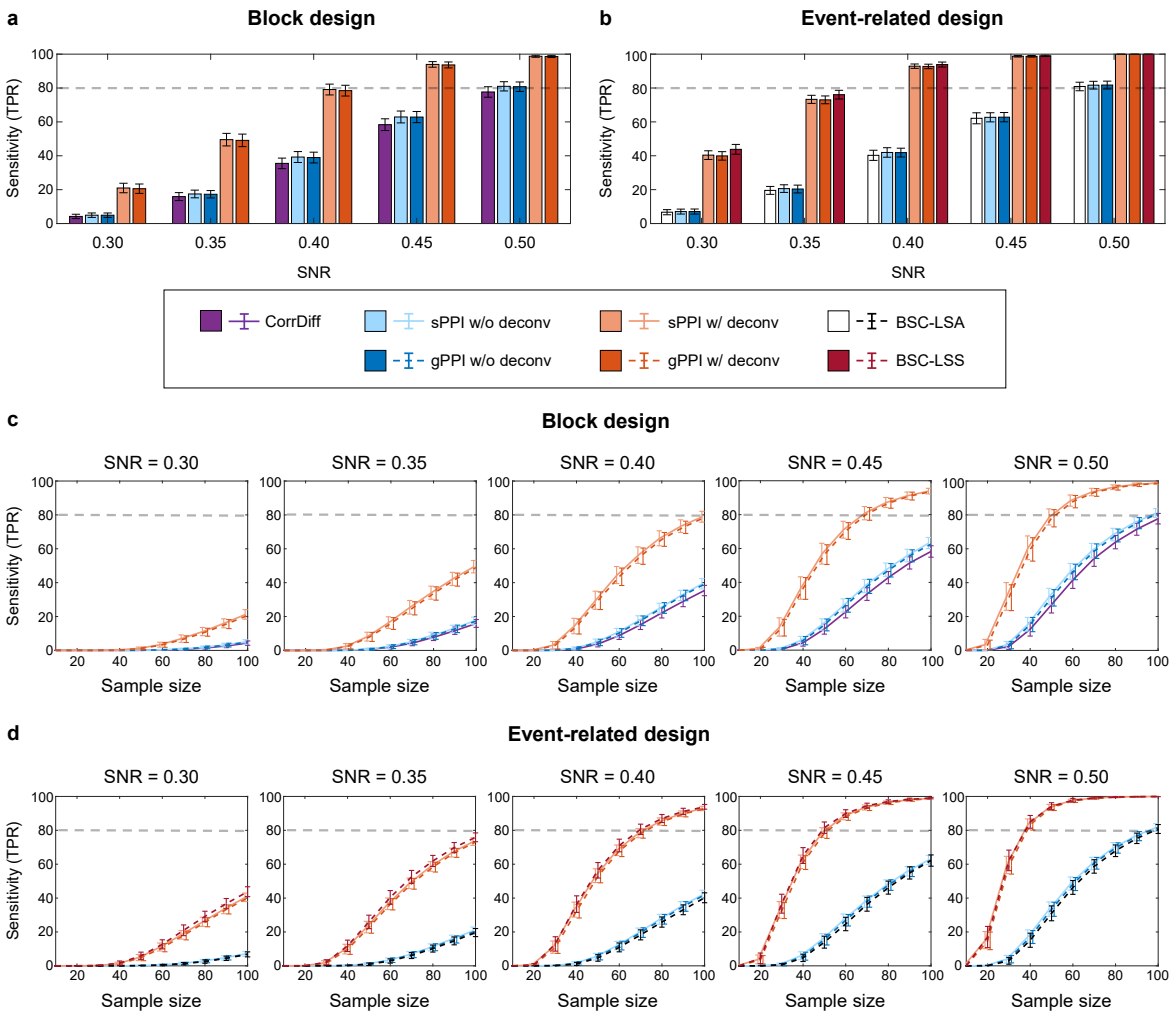


Fig. 6. Sensitivity of all TMFC methods depending on the signal-to-noise ratio (SNR) and sample size. Simulation results for TR = 2 s. **(a-b)** Sensitivity of different TMFC methods at sample size N = 100 for **(a)** the block design with ten 20 s blocks per condition and **(b)** event-related design with one hundred 1 s events per condition and mean ISI = 6 s. **(c-d)** Sensitivity of different TMFC methods depending on the sample size for **(c)** the block design and **(d)** event-related design. PPI terms were calculated with (w/) and without (w/o) the deconvolution step. sPPI and gPPI matrices were symmetrised. Error bars represent SD and were computed from 1000 random resamplings with replacement.

Sensitivity of the TMFC methods for event-related designs with different temporal structures

Next, we independently varied the temporal parameters of the event-related design, including event duration, mean ISI, number of events and data acquisition (TR), to assess their impact on the sensitivity of TMFC methods with a sample size N =100 and medium SNR = 0.40. Shortening the event duration substantially decreased the sensitivity of all TMFC methods (Fig. 7a, Extended Data Table 3). The BSC-LSS method was slightly more sensitive than the PPI methods, which was more noticeable for short event durations < 1 s; in contrast, the BSC-LSA method achieved > 80% sensitivity only for long event durations > 2 s.

Shortening the mean ISI slightly decreased the sensitivity of the BSC-LSS and PPI methods and substantially reduced the sensitivity of the BSC-LSA method (Fig. 7b, Extended Data Table 4). The BSC-LSA method had > 80% sensitivity only for slow event-related designs with mean ISI = 12 s. For the rapid event-related designs (ISI ≤ 4 s), the PPI methods were slightly more sensitive than the BSC-LSS approach. In all other cases, the BSC-LSS approach was more sensitive than the PPI methods.

Reducing the number of events also substantially decreased the sensitivity of all TMFC methods (Fig. 7c, Extended Data Table 5). The BSC-LSS approach was slightly more robust to shortening scan time than were the PPI methods. The BSC-LSS and PPI methods needed at least 80 events per condition to achieve > 80% sensitivity.

Finally, fast data acquisition (TR < 1 s) yielded the maximum sensitivity for all methods (Fig. 7d, Extended Data Table 6), while reducing the fMRI temporal resolution from the typical TR = 2 s to the frequently used TR = 3 s reduced sensitivity below 80%. Notably, the PPI methods were more sensitive than the BSC-LSS method for TR = 3 s.

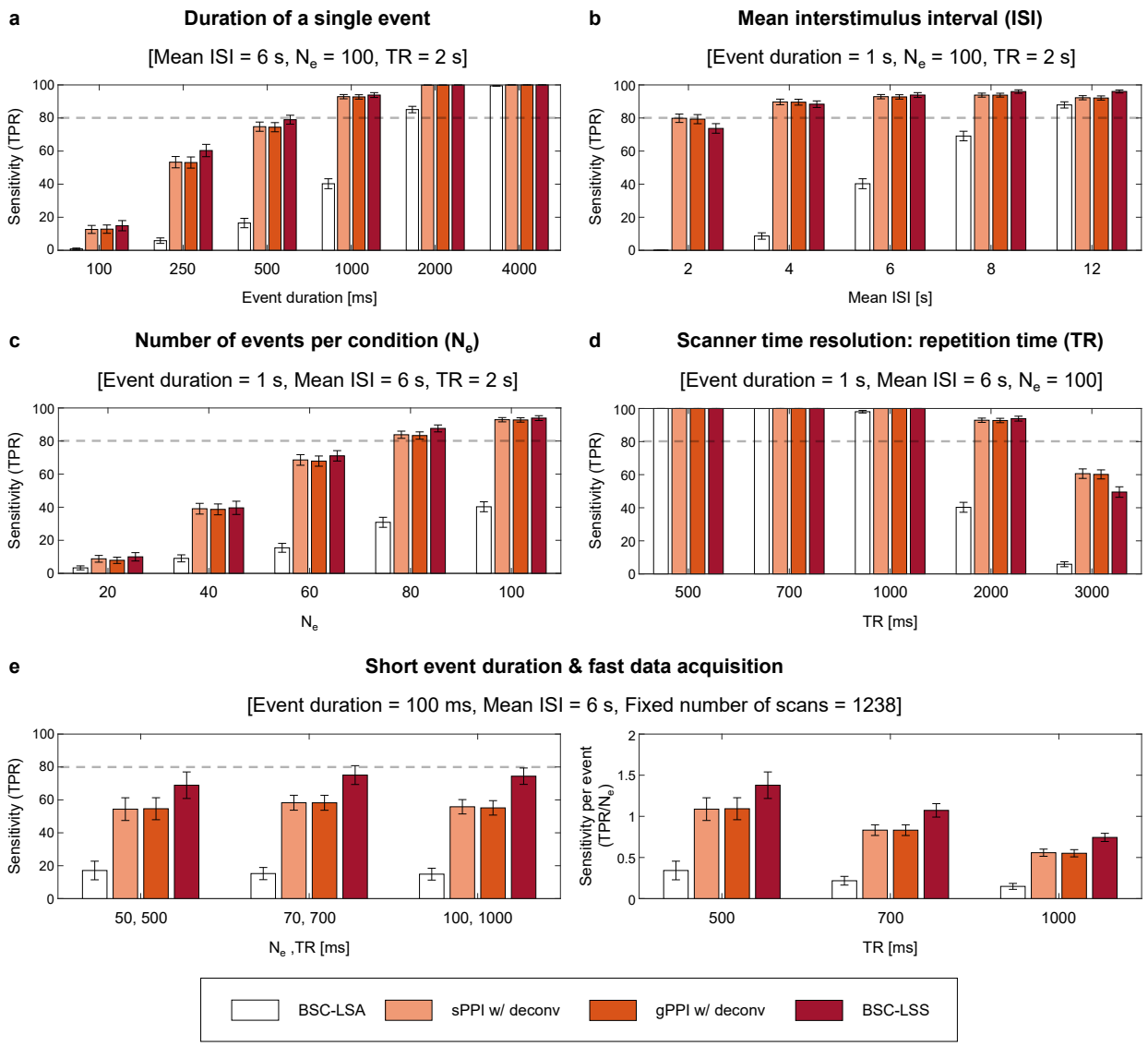


Fig. 7. Sensitivity of TMFC methods for the event-related design with different temporal parameters. Simulation results for sample size N = 100 and SNR = 0.4. The default event-related design consisted of one hundred 1 s events per condition and a mean ISI = 6 s. **(a)** Sensitivity depending on the duration of a single event. **(b)** Sensitivity depending on the mean interstimulus (ISI) interval. **(c)** Sensitivity depending on the number of events per condition (N_e). **(d)** Sensitivity depending on the repetition time (TR) for a fixed scan time (23.6 minutes). **(e)** Sensitivity for the event-related design with short event duration (100 ms), fast data acquisition (≤ 1000 ms) and fixed number of scans (1238). Fixing the number of scans resulted in variable scan times and numbers of events. The panel on the left side represents raw sensitivity (true positive rate, TPR). The panel on the right side represents normalised sensitivity per single event (TPR/N_e). sPPI and gPPI matrices were symmetrised. Error bars represent SD and were computed from 1000 random resamplings with replacement.

Rapid synchronisation can be revealed even with slow fMRI data acquisition

One intriguing question is the principal ability of BOLD fMRI to assess rapid modulations of gamma-band neuronal synchronisation evoked by short events given typically slow data acquisition (TR = 2 s). In the previous section, we showed that the most sensitive TMFC method (BSC-LSS) revealed a fairly small number of true positives (TPR = 15%) for the event-related design with one hundred 100 ms events per condition (Fig. 7a). When we doubled the number of events, the sensitivity was increased to TPR = 57%. Therefore, typically slow fMRI sequences can, in principle, detect task-related neuronal

synchronisation on the order of 100 ms. However, the sensitivity to such rapid synchronisation with slow data acquisition is relatively weak, and many events are required to detect them.

Importance of fast fMRI data acquisition for TMFC analysis

Another possible way to increase the sensitivity of TMFC methods to rapid neuronal synchronisation is to employ fast fMRI data acquisition techniques. Increasing the fMRI temporal resolution from TR = 2 s to TR = 500 ms resulted in 99% sensitivity for the event-related design with one hundred 100 ms events per condition when the BSC-LSS method was applied. At a fixed task duration (scan time = 23.6 minutes), decreasing TR from 2 s to 500 ms increased the number of fMRI data points fourfold (from 616 to 2464 scans). Thus, with the same scanning time, fast data sampling enables one to increase the temporal degree of freedom and thereby significantly increase the statistical power of TMFC methods.

Furthermore, fast fMRI data sampling may improve sensitivity *per se*, that is, by more precise insights into neuronal temporal dynamics rather than simply by providing more data points for a fixed scan time[41]. To test this assumption, in contrast to our previous simulations with fixed scan time and number of events, we considered a fixed number of data points (scans). As a result, the sensitivity of TMFC methods remained at the same level at different temporal resolutions, TR = 500/700/1000 ms, despite a reduction in total scan time and number of events per condition, N_e = 50/70/100 (Fig. 7e, left panel). This meant an increase in normalised sensitivity per single event (TPR/N_e) at shorter TRs (Fig. 7e, right panel). The BSC-LSS method had the highest sensitivity per event for short event duration and fast fMRI data acquisition. Therefore, despite the sluggishness of haemodynamic processes, a more accurate characterisation of ultra-slow BOLD-signal fluctuations allows for more precise insights into rapid task-related modulation of gamma-band synchronisation.

Genuine and spurious asymmetry of the PPI matrices

In the previous sections, we considered only symmetrised PPI matrices because averaging the upper and lower diagonal elements has become a standard procedure in TMFC analysis since they are considered quite similar[8,11]. Here, we use empirical and simulated data to determine when PPI matrices become asymmetric, making the averaging procedure problematic. Below, we report results only for the gPPI method since the sPPI and gPPI methods produce nearly identical results.

For the block design (working memory task), the correlations between the upper and lower diagonal elements of the group-mean gPPI matrices without and with deconvolution were 0.85 and 0.77, respectively (Extended Data Fig. 2a). Without deconvolution, the correlation coefficients for individual subjects ranged from 0.77 to 0.89 with a mean of 0.83 (Extended Data Fig. 2c). With deconvolution, individual correlation coefficients ranged from 0.56 to 0.75 (mean 0.68). For the event-related design (stop-signal task), the correlations between the upper and lower diagonal elements of the group-mean gPPI matrices without and with deconvolution were 0.90 and 0.40, respectively (Extended Data Fig. 2b). The correlation coefficients for individual subjects ranged from 0.78 to 0.93 with a mean of 0.90 without deconvolution (Extended Data Fig. 2d). In contrast, when deconvolution was applied, individual correlation coefficients ranged from 0.24 to 0.41 (mean 0.34). Similar results were obtained for other tasks with block and event-related designs taken from the HCP and CNP datasets (Supplementary Fig. S3). Therefore, empirical data showed that deconvolution slightly increases the asymmetry of gPPI matrices for block designs and substantially increases the asymmetry for event-related designs.

Next, we used simulations with symmetric ground-truth matrices to determine which parameters of the fMRI experiment can artificially increase the asymmetry of the PPI matrices. We found that the main factors for artificial matrix asymmetry are a low SNR (Extended Data Fig. 3a), small sample size (Extended Data Fig. 3b), short event duration (Extended Data Fig. 3c) and small number of events per condition (Extended Data Fig. 3e). In addition, the asymmetry of gPPI matrices increases slightly with larger TRs (Extended Data Fig. 3f) and is practically independent of the mean ISI duration (Extended

Data Fig. 3d). Therefore, the large asymmetry of the gPPI matrices for the event-related designs from the CNP dataset is most likely related to the short scan time and low SNR.

Finally, we considered simulations with asymmetric ground-truth matrices (Supplementary Fig. S4) to test whether the gPPI method could be used to obtain at least some information about the direction of connectivity. The causal influence that one neuronal system exerts over another at a synaptic or neuronal population level is referred to as effective connectivity (EC). Here, we correlated the asymmetric ground-truth matrix with the asymmetric gPPI matrix. For the block designs, we also calculated task-modulated EC (TMEC) matrices using the regression dynamic causal modelling (rDCM) method[42,43] (see Methods). The rDCM method, which is a conventional EC method, was used as a reference. As rDCM requires a relatively high SNR[42], we used SNR = 5 and twice the scan duration. If the gPPI method is unable to correctly evaluate the direction of information flow at a high SNR, then it would also fail at lower SNRs. A systematic comparison of TMEC methods is beyond the scope of the present work.

As a result, the gPPI method with deconvolution was able to reflect the actual direction of the information flow for the block design with twenty blocks per condition (Fig. 8). The correlation between the group-mean asymmetric gPPI matrix and ground-truth matrix for the "Cond A – Cond B" difference was 0.93, and correlation between the group-mean rDCM matrix and ground-truth matrix was 0.89. When the scan duration was halved, the correlation between the gPPI and ground-truth matrices dropped to 0.83, and further dropped to 0.15 when SNR decreased to 0.5. Similar results were obtained for the event-related designs: the correlation between the gPPI and ground-truth matrices was 0.85 for two hundred events per condition, reduced to 0.56 when the scan duration was halved, and further reduced to 0.03 when SNR decreased to 0.5. Without deconvolution, the correlation between the ground-truth and gPPI matrices was always close to zero ($|r| < 0.05$). Thus, only deconvolution allows gPPI to estimate the effective strength of task-modulated synaptic connections, but only at high SNRs and long scan times.

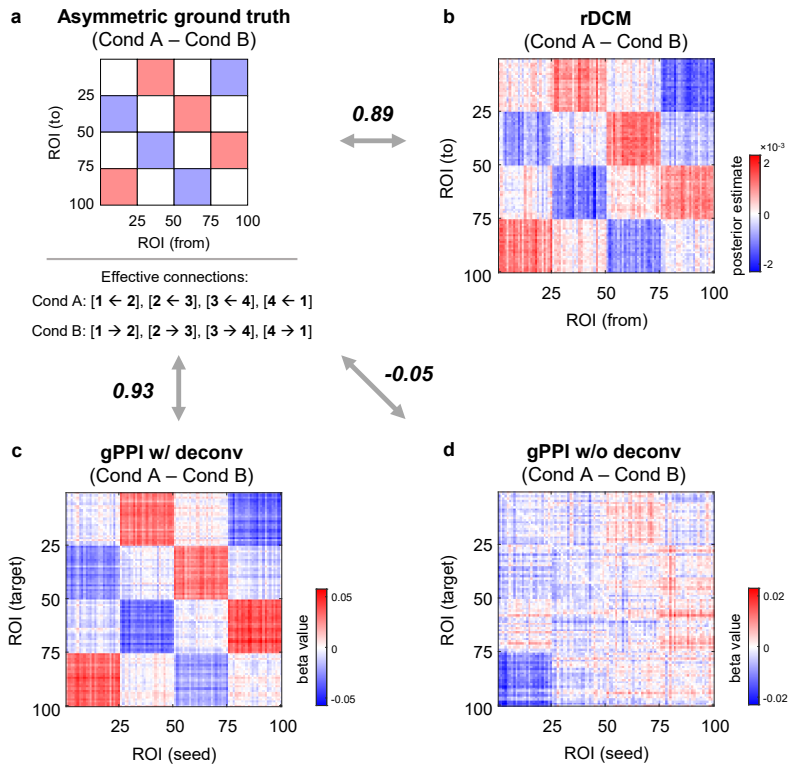


Fig. 8. gPPI with deconvolution can reveal the direction of information flow under high SNR and long scan time. Simulation results for the block design with twenty 20 s blocks per condition, sample size N = 100, TR = 2 s and very high SNR = 5. **(a)** Asymmetric ground-truth matrix of task-modulated effective connections. In "Cond A", synaptic weights were increased from module №1 to №4, from №4 to №3, from №3 to №2, and from №2 to №1. In "Cond B", synaptic weights were increased in the opposite direction. **(b)** Group-mean difference between rDCM matrices calculated for the "Cond A"

and "Cond B" blocks. **(c)** Group-mean asymmetric matrix generated by gPPI with (w/) deconvolution. Deconvolution allows modelling of psychophysiological interactions at the neuronal level. **(d)** Group-mean asymmetric matrix generated by gPPI without (w/o) deconvolution. Without deconvolution psychophysiological interactions, are modelled at the haemodynamic level. To evaluate the similarity between matrices, we used Pearson's r correlations.

Discussion

This is the first evaluation of existing whole-brain TMFC techniques using biophysically realistic large-scale neural mass simulations for a wide range of fMRI experimental settings. We identified the most effective TMFC methods for block and event-related designs, determined which data analysis procedures and parameters of the fMRI experiment increase their sensitivity and specificity, and demonstrated the principal capability of fMRI to detect rapid task-related neuronal synchronisation from sluggish BOLD signals at various temporal resolutions.

The simplest and most intuitive approach for constructing a whole-brain TMFC matrix is to directly calculate the correlation difference between task conditions, accounting for transient haemodynamic effects[11]. However, this approach is suitable only for task designs with long blocks since there must be enough time points after the removal of the transition periods (first six seconds of each block). We determined that the CorrDiff method has substantially lower sensitivity than other methods since cutting out the transition periods significantly reduces the temporal degrees of freedom.

A more sophisticated TMFC method available for both block and event-related designs is the PPI approach[12]. The psychophysiological interaction can be modelled at the haemodynamic level or at the neuronal level using the deconvolution procedure[15]. Standard PPI (sPPI) approach was originally proposed for seed-to-voxel analysis and optimised for designs with two task conditions[12]. Later, it was extended to a generalised form (gPPI) that enables the assessment of more than two task conditions per statistical model[13] and adapted for whole-brain ROI-to-ROI analysis[31,32,33]. In this study, we established that the most sensitive TMFC techniques for block designs and rapid event-related designs (mean ISI ≤ 4 s) are the sPPI and gPPI methods incorporating the deconvolution procedure. Although previous simulation studies suggested that the gPPI method is more sensitive than the sPPI method[13,18], we showed that the sensitivity of both methods is practically equivalent for block and event-related designs. This discrepancy is most likely due to the difference between previously used biologically simplified simulations based on box-car or delta functions[13,18] and biologically realistic simulations of TMFC based on the large-scale neuronal mass model used in the present work. Nevertheless, the gPPI method is preferable to the sPPI method since it provides more flexibility regarding task design and higher specificity if co-activations have not been removed before TMFC analysis using FIR task regression.

Previous empirical[44] and simulation studies[15,20] showed that the PPI approach with and without deconvolution produces similar results for block designs. However, disparities between PPI terms calculated at the haemodynamic and neuronal levels were more prominent for event-related designs since there are more high-frequency components[15]. Some authors suggest that the deconvolution step is mandatory for event-related designs[8], while others warn against its use, arguing that there is no deterministic way to deconvolve the haemodynamic response if its shapes are not known exactly[28]. While the deconvolution step is a default setting in SPM, it is not implemented in other popular neuroimaging software packages, such as FSL and the CONN toolbox. Based on our results, we argue that deconvolution can increase sensitivity by up to a factor of two for both event-related and block designs. Moreover, deconvolution can increase specificity several fold, effectively eliminating co-activations even without FIR regression. Without deconvolution, the gPPI method is susceptible to spurious inflation of TMFC estimates due to simple co-activations.

Since PPI is a regression method, it produces asymmetric (directed) TMFC matrices. However, to date there is no consensus whether PPI can be used to reveal the direction of information flow. Some authors used PPI without deconvolution as an effective connectivity method[35]. Others have used PPI with deconvolution to determine the direction of connectivity without making strong claims about the

underlying causal structure[33]. A popular view is that by using PPI we are making arbitrary directional assumptions based on a priori assignment of ROIs as sources or targets[14,29]. Here, we showed that PPI can reveal the direction of information flow only if the psychophysiological interaction is modelled at the neuronal level by using the deconvolution procedure. Without deconvolution, the PPI method provide no information about the underlying causal structure. In addition, we found that PPI matrices asymmetry can spuriously arise due to low SNR, small sample size, short scan time (small number of events) and short event duration. To make correct directional inferences, the PPI method, as well as the more sophisticated effective connectivity method, rDCM, requires high SNR and long scan times.

To create symmetrical PPI matrices, it was proposed to average upper and lower diagonal elements based on the empirical evidence that these elements are strongly correlated in block design tasks[11]. However, for task designs with low SNR and short scan time, the PPI matrices can become largely asymmetric, making the averaging procedure problematic. To completely avoid the asymmetry problem, it has been proposed to calculate symmetric TMFC matrices using correlational form of PPI (cPPI) based on partial correlations[14]. The cPPI method has not been previously validated on empirical or simulated data or compared with other TMFC methods. Here, we demonstrated that the cPPI method cannot estimate TMFC: it is unable to eliminate spontaneous task-independent activity and produces matrices similar to the RSFC, TSFC and BGFC matrices.

An alternative method for obtaining symmetrical whole-brain TMFC matrices for event-related designs is the BSC approach[16,17], which is based on correlations between beta estimates of the BOLD signal change for each individual trial. A simulation study by Cisler et al.[18] suggested that the BSC approach is more powerful than gPPI for event-related designs with many trial repetitions, while an empirical study by Di & Biswal[8] with a large sample size did not support this notion. The inconsistency between the results of these studies may be due to the discrepancy between real neurophysiological processes that cause TMFC and the biophysically unrealistic simulations used to mimic TMFC[18]. We determined that the BSC-LSS method is superior for most types of event-related designs (mean ISI > 4 s). The increase in sensitivity compared to that of the gPPI method is especially noticeable for short event durations (< 1 s). At the same time, the BSC-LSA method had the lowest sensitivity due to the multicollinearity problem and achieved reasonable sensitivity only for long ISI (≥ 12 s) and/or long event duration (≥ 2 s). Note that one of the popular BSC software packages (BASCO toolbox) utilises the LSA method.

Finally, we showed that short-term (100 ms) modulations of gamma-band neuronal synchronisation can, in principle, be recovered from ultra-slow BOLD signal oscillations, even at typically low temporal resolution (TR = 2 s). However, this requires many trials (> 200 per condition). Furthermore, fast fMRI data acquisition increases the sensitivity for short-term modulations not only by increasing the number of data points for a fixed scan time but also by providing more insights into neuronal temporal dynamics. Our results show the particular importance of fast fMRI sequences for TMFC analysis.

In conclusion, we provide practical recommendations for TMFC analysis:

1. Remove co-activations from time series using FIR regression, especially if you use gPPI without deconvolution or sPPI with/without deconvolution.
2. For primary TMFC analysis, use:
    a. gPPI for block designs and rapid event-related designs (mean ISI ≤ 4 s);
    b. BSC-LSS for all other designs.
3. Perform cross-method validations (secondary TMFC analysis):
    a. High correlation between raw TMFC matrices calculated by gPPI and BSC methods suggests that the results are reliable;
    b. Low correlation indicates insufficient or low-quality data for TMFC analysis.
4. Use the deconvolution procedure to increase sensitivity and specificity of PPI methods.
5. When averaging the upper and lower elements of the PPI matrix, report its symmetry:

a. High correlation between the upper and lower elements of the PPI matrix suggests that the results are reliable;
b. Low correlation indicates insufficiency or low quality of data for TMFC analysis.

6. Draw conclusions about the direction of information flow from PPI results with caution:
   a. Use the deconvolution procedure. Without deconvolution, PPI methods do not provide information about directional information flow.
   b. Make sure you have a high enough SNR and scan duration.
   c. Support preliminary whole-brain PPI results with confirmatory analysis for several selected ROIs using an advanced effective connectivity method such as DCM.
7. For BSC analysis, always use the LSS method.
8. When TMFC analysis is the main goal:
   a. Avoid using rapid event-related designs (mean ISI < 4 s);
   b. Increase the number of events as much as possible (TMFC analysis requires more events per condition than does the activation analysis);
   c. Use a longer event duration (if possible);
   d. Fast fMRI data acquisition is preferable (TR < 1 s).

Last, we want to note how our simulation approach can be improved in future work devoted to TMFC analysis. First, it is possible to introduce variability in neural mass and haemodynamic model parameters to account for the heterogeneity of brain regions. Second, oscillations other than those in the gamma-band range can be considered for modelling (alpha-band, beta-band, or multiple-frequency-band oscillations). Third, for fast fMRI acquisition, a penalty of reduced SNR per scan can be introduced due to reduced longitudinal magnetisation recovery.

Main references

1. Sporns, O., Tononi, G. & Kötter, R. The Human Connectome: A Structural Description of the Human Brain. *PLOS Computational Biology* **1**, e42 (2005).
2. Biswal, B. B., Yetkin, F. Z., Haughton, V. M. & Hyde, J. S. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine* **34**, 537–541 (1995).
3. Biswal, B. B. et al. Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences* **107**, 4734–4739 (2010).
4. Fox, M. D. et al. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences* **102**, 9673–9678 (2005).
5. Yeo, B. T. T. et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology* **106**, 1125–1165 (2011).
6. Seeley, W. W., Crawford, R. H., Zhou, J., Miller, B. L. & Greicius, M. D. Neurodegenerative Diseases Target Large-Scale Human Brain Networks. *Neuron* **62**, 42–52 (2009).
7. Zhang, D.-Y. & Raichle, M. E. Disease and the brain's dark energy. *Nature Reviews Neurology* **6**, 15–28 (2010).
8. Di, X. & Biswal, B. B. Toward Task Connectomics: Examining Whole-Brain Task Modulated Connectivity in Different Task Domains. *Cerebral Cortex* **29**, 1572–1583 (2018).
9. Norman-Haignere, S. V., McCarthy, G., Chun, M. M. & Turk-Browne, N. B. Category-Selective Background Connectivity in Ventral Visual Cortex. *Cerebral Cortex* **22**, 391–402 (2011).
10. Cole, M., Bassett, D. S., Power, J. D., Braver, T. S. & Petersen, S. E. Intrinsic and Task-Evoked Network Architectures of the Human Brain. *Neuron* **83**, 238–251 (2014).
11. Di, X., Reynolds, R. & Biswal, B. B. Imperfect (de)convolution may introduce spurious psychophysiological interactions and how to avoid it. *Human Brain Mapping* **38**, 1723–1740 (2017).
12. Friston, K. J. et al. Psychophysiological and Modulatory Interactions in Neuroimaging. *NeuroImage* **6**, 218–229 (1997).
13. McLaren, D. G., Ries, M. L., Xu, G. & Johnson, S. C. A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *NeuroImage* **61**, 1277–1286 (2012).
14. Fornito, A., Harrison, B. J., Zalesky, A. & Simons, J. S. Competitive and cooperative dynamics of large-scale brain functional networks supporting recollection. *Proceedings of the National Academy of Sciences* **109**, 12788–12793 (2012).
15. Gitelman, D. R., Penny, W. D., Ashburner, J. & Friston, K. J. Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. *NeuroImage* **19**, 200–207 (2003).
16. Rissman, J., Gazzaley, A. & D'Esposito, M. Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage* **23**, 752–763 (2004).
17. Mumford, J. A., Turner, B. L., Ashby, F. G. & Poldrack, R. A. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage* **59**, 2636–2643 (2012).
18. Cisler, J. M., Bush, K. & Steele, J. S. A comparison of statistical methods for detecting context-modulated functional connectivity in fMRI. *NeuroImage* **84**, 1042–1052 (2014).
19. Abdulrahman, H. & Henson, R. N. Effect of trial-to-trial variability on optimal event-related fMRI design: Implications for Beta-series correlation and multi-voxel pattern analysis. *NeuroImage* **125**, 756–766 (2016).

20. Kim, J.-E. & Horwitz, B. Investigating the neural basis for fMRI-based functional connectivity in a blocked design: application to interregional correlations and psycho-physiological interactions. *Magnetic Resonance Imaging* **26**, 583–593 (2008).
21. Cole, M. et al. Task activations produce spurious but systematic inflation of task functional connectivity estimates. *NeuroImage* **189**, 1–18 (2019).
22. Wilson, H. R. & Cowan, J. D. Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. *Biophysical Journal* **12**, 1–24 (1972).
23. Deco, G., Jirsa, V. K., McIntosh, A. M., Sporns, O. & Kötter, R. Key role of coupling, delay, and noise in resting brain fluctuations. *Proceedings of the National Academy of Sciences* **106**, 10302–10307 (2009).
24. Friston, K. J., Harrison, L. H. & Penny, W. D. Dynamic causal modelling. *NeuroImage* **19**, 1273–1302 (2003).
25. Zucker, R. A. & Regehr, W. G. Short-Term Synaptic Plasticity. *Annual Review of Physiology* **64**, 355–405 (2002).
26. Stephan, K. et al. Ten simple rules for dynamic causal modeling. *NeuroImage* **49**, 3099–3109 (2010).
27. Regehr, W. G. Short-Term Presynaptic Plasticity. *Cold Spring Harbor Perspectives in Biology* **4**, a005702 (2012).
28. O'Reilly, J., Woolrich, M. W., Behrens, T. E. J., Smith, S. M. & Johansen-Berg, H. Tools of the trade: psychophysiological interactions and functional connectivity. *Social Cognitive and Affective Neuroscience* **7**, 604–609 (2012).
29. Fornito, A., Zalesky, A. & Breakspear, M. Graph analysis of the human connectome: Promise, progress, and pitfalls. *NeuroImage* **80**, 426–444 (2013).
30. Stephan, K. E. & Friston, K. J. Analyzing effective connectivity with functional magnetic resonance imaging. *Wiley Interdisciplinary Reviews: Cognitive Science* **1**, 446–459 (2010).
31. Minati, L., Grisoli, M., Seth, A. K. & Critchley, H. D. Decision-making under risk: A graph-based network analysis using functional MRI. *NeuroImage* **60**, 2191–2205 (2012).
32. Cocchi, L. et al. Complexity in Relational Processing Predicts Changes in Functional Brain Network Dynamics. *Cerebral Cortex* **24**, 2283–2296 (2013).
33. Gerchen, M. F., Bernal-Casas, D. & Kirsch, P. Analyzing task-dependent brain network changes by whole-brain psychophysiological interactions: A comparison to conventional analysis. *Human Brain Mapping* **35**, 5071–5082 (2014).
34. Smith, D., Gseir, M., Speer, M. E. & Delgado, M. R. Toward a cumulative science of functional integration: A meta-analysis of psychophysiological interactions. *Human Brain Mapping* **37**, 2904–2917 (2016).
35. Wang, Y. et al. Multimodal mapping of the face connectome. *Nature Human Behaviour* **4**, 397–411 (2020).
36. Barch, D. M. et al. Function in the human connectome: Task-fMRI and individual differences in behavior. *NeuroImage* **80**, 169–189 (2013).
37. Poldrack, R. A. et al. A phenome-wide examination of neural and cognitive function. *Scientific Data* **3**, (2016).
38. Mateo, C., Knutsen, P. M., Tsai, P. S., Shih, A. Y. & Kleinfeld, D. Entrainment of arteriole vasomotor fluctuations by neural activity is a basis of Blood-Oxygenation-Level-Dependent "Resting-State" connectivity. *Neuron* **96**, 936-948.e3 (2017).
39. Tzourio-Mazoyer, N. et al. Automated Anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI Single-Subject Brain. *NeuroImage* **15**, 273–289 (2002).
40. Cabral, J., Hugues, E. & Kringelbach, M. L. Modeling the outcome of structural disconnection on resting-state functional connectivity. *NeuroImage* **62**, 1342–1353 (2012).
41. Dowdle, L. T. et al. Statistical power or more precise insights into neuro-temporal dynamics? Assessing the benefits of rapid temporal sampling in fMRI. *Progress in Neurobiology* **207**, 102171 (2021).
42. Frässle, S. et al. Regression DCM for FMRI. *NeuroImage* **155**, 406–421 (2017).
43. Frässle, S. et al. Whole-brain estimates of directed connectivity for human connectomics. *NeuroImage* **225**, 117491 (2021).
44. Di, X. & Biswal, B. B. Psychophysiological interactions in a visual checkerboard task: reproducibility, reliability, and the effects of deconvolution. *Frontiers in Neuroscience* **11**, (2017).

Methods

Empirical data

We analysed preprocessed fMRI data for two block design tasks (working memory and social cognition tasks, N = 100) from the HCP dataset[36,45], two event-related tasks (stop-signal and task-switching tasks, N = 115) from the CNP dataset[37,46] and resting-state data from both datasets. For details about the scanning parameters and task designs, see Supplementary Information 2. The HCP and CNP preprocessing pipelines included realignment, spatial artefact/distortion correction, co-registration between functional and structural images, and normalisation to Montreal Neurological Institute (MNI) space. The CNP pipeline also included a slice-timing correction. Additionally, we smoothed the functional data with a 4 mm Gaussian kernel using SPM12 (https://www.fil.ion.ucl.ac.uk/spm/software/spm12).

To construct empirical FC matrices, we used a set of functionally defined ROIs covering the whole brain[47]. The ROIs were defined as spheres with a radius of 4 or 5 mm. We discarded ROIs that did not have data for at least one subject. As a result, we utilised 239 ROIs for the HCP dataset and 246 ROIs for the CNP dataset.

Large-scale neural mass model

The simulation procedure included five steps. First, we simulated gamma-band oscillatory neuronal activity for 100 interconnected brain regions using Wilson-Cowan equations[22,23] and manipulated the synaptic weights depending on the task conditions to control the ground-truth TMFC. Transient activity-dependent modulation of synaptic strength, lasting from tens of milliseconds to several minutes, is referred to as short-term synaptic plasticity[25,26,27]. Second, we independently simulated co-activations using box-car functions to evaluate their impact on spurious inflation of TMFC estimates. Third, we applied the Balloon-Windkessel haemodynamic model[24] to convert oscillatory neuronal activity and co-activations into BOLD signals. Fourth, we downsampled the BOLD signal to different time resolutions to assess the potential benefits of fast data acquisition for TMFC estimation. Fifth, we added white Gaussian noise to model scanner measurement error.

Our simulation approach expands previous TMFC simulation studies in several ways. First, we used a large-scale neural mass model instead of delta or boxcar functions[13,15,18,19]. Second, we applied biophysically realistic simulations not only to block designs[20,21] but also to different event-related designs. Third, previous biophysically realistic simulations have indirectly modulated FC by injecting task stimulation into brain regions[20,21]. In the current work, we directly manipulated ground-truth FC by changing synaptic weights between neuronal units depending on the task context, which corresponds to short-term synaptic plasticity[25,26,27]. Finally, we ensured that the neural mass model generated spontaneous oscillations in the gamma band. In contrast, the most recent TMFC simulation study by Cole et al.[21] used a large-scale neural mass model without inhibitory subpopulations, where limit-cycle oscillations cannot emerge under any model parameters (Supplementary Fig. S5). For a more detailed overview of the results and limitations of previous TMFC simulation studies, see Supplementary Table S1.

The Wilson–Cowan neural mass model achieves a good balance between biophysical realism and mathematical abstraction[23,48,49]. According to this model, a single neuronal population can be represented as synaptically coupled excitatory and inhibitory subpopulations described by two ordinary differential equations with a non-linear saturation function[22]. Each population can produce self-sustained limit-cycle oscillations as a result of feedback between coupled excitatory and inhibitory subpopulations. The Wilson-Cowan units were connected through the excitatory subpopulations[23,49,50] with a signal transmission delay of 25 ms[51]. In line with previous large-scale neural mass simulations[23,49], we chose to set the model parameters such that each coupled Wilson-Cowan unit produced gamma-band oscillations (≈ 40 Hz) due to the following considerations.

A large body of literature demonstrates that gamma-band oscillations and synchronisation are linked to sensory processing[52,53], motor acts[54], and cognitive processes[55,56] and are thought to underlie information processing and transmission[57,58]. In addition, the spectral power (or envelope) of gamma-band oscillations fluctuates very slowly with time, and brain regions with shared function demonstrate co-fluctuation of gamma-band envelopes[59,60]. Moreover, a multitude of animal and human studies have shown that local field potential power in the gamma band is the closest electrophysiological correlate of spontaneous and evoked BOLD signals[61,62,63,64,65]. The mechanisms for the relationship between neuronal activity and BOLD signals have not yet been fully determined. However, a recent study by Mateo et al.[38] elucidated these mechanisms using optogenetic manipulations and concurrently measuring local field potential, arteriole diameter and blood oxygenation in awake mice. They provided direct evidence that an increase in gamma-band power leads to an increase in arteriole diameter, and an increase in arteriole diameter leads to an increase in blood oxygenation. This chain of processes can be described by a coupled oscillator model (Fig. 3a). For details on the applied model and its parameters, see Supplementary Information 3.1 and Supplementary Table S2. Synaptic activity was calculated as the sum of all inputs to excitatory and inhibitory subpopulations of the Wilson-Cowan unit and was considered a proxy for the local field potential[20,66].

Ground-truth synaptic weight matrices

The construction of the synaptic weight matrices involved three steps. First, we drew synaptic weights for each subject from a Gaussian distribution with a mean of one and a standard deviation of 0.1. Then, we multiplied the synaptic weights within and between functional modules by weighting factors that determined the network structure (Supplementary Table S3). Finally, we normalised the synaptic weights so that all inputs to each region summed to one, following previous simulation studies[21,49].

Analysis of simulated synaptic activity

We additionally analysed simulated synaptic activity before converting it into the BOLD signal. Initially, we downsampled the raw time series from a temporal resolution of 0.1 ms to 5 ms and bandpass filtered it in a narrow carrier frequency range [$f_{carrier}$ – 2 Hz, $f_{carrier}$ + 2 Hz], where $f_{carrier}$ = 40 Hz[67]. Next, we employed the Hilbert transform to obtain the instantaneous amplitudes and phases of the narrowband signal. The instantaneous amplitudes, or amplitude envelopes, were further cross-correlated with the simulated BOLD signals, while the instantaneous phases were utilised to estimate gamma-band neuronal synchronisation based on the phase-locking value method[68].

BOLD signal generation

The synaptic activity was converted to the BOLD signal using the Balloon-Windkessel haemodynamic model[24]. For more information about the model, see Supplementary Information 3.2. We used the standard parameters taken from Friston et al.[24] (Supplementary Table S4), which have previously been used in whole-brain RSFC simulation studies[23,40,50].

The BOLD signal generated from fast oscillatory activity corresponded to spontaneous oscillations and task-modulated oscillations during the rest and task conditions. To simulate the BOLD signal related to simple co-activations, we used boxcar activation functions as input to the Balloon-Windkessel model. Boxcar activation functions were set equal to one during task conditions and equal to zero at all other times. We normalised the BOLD signal related to co-activations such that the ratio between the standard deviations of the oscillatory-related signal and co-activation-related signal was determined by the scaling factor: SF = $\sigma_{oscill}/\sigma_{coact}$. In all simulations with co-activations, we set SF = 1.

Next, we downsampled the raw BOLD signals from 0.1 ms to distinct time resolutions corresponding to different TRs and added white Gaussian noise as measurement error. The SNR was defined as the ratio between the standard deviations of the signal and noise, SNR = $\sigma_{signal}/\sigma_{noise}$, which is commonly used in DCM studies[42,43]. To compare the different TMFC methods, we varied the SNR between 0.3 and

0.5, and the default SNR was 0.4. To compare the gPPI and rDCM methods, we increased the SNR to 5, as the rDCM method requires a higher SNR.

Simulation experiments and task designs

By default, we considered symmetric synaptic weight matrices (Fig. 3c). The block design included 10 blocks each for the "A" and "B" conditions, alternating with "Rest" blocks. Each block lasted for 20 s. The total duration of the block design was ≈ 13 minutes. The default event-related design included 100 events for the "A" and "B" conditions interleaved by "Rest" periods. Each event lasted for 1 s. The ISI was randomly jittered between 4–8 s (mean ISI = 6 s). The ISI was defined as the interval between the end of one event and the start of the next event. The total duration of the default event-related design was ≈ 24 minutes. In other event-related designs, we separately varied event duration = [100 ms, 250 ms, 500 ms, 1 s, 2 s, 4 s], number of events per condition = [20, 40, 60, 80, 100], and mean ISI = [2 s, 4 s, 6 s, 8 s, 12 s]. Stimulus onset timings were determined by Chris Rorden's fMRI Design software (https://github.com/neurolabusc/fMRI-Simulator), which improves the statistical efficiency of the task design. By default, we used typical slow fMRI data acquisition, TR = 2 s. In a separate experiment, we varied TR = [500 ms, 700 ms, 1 s, 2 s, 3 s]. For the simulation with asymmetric synaptic weight matrices (Supplementary Fig. S4), we used the block design. All simulations were performed for N = 100 subjects.

Resting-state, background and task-state functional connectivity

To calculate the empirical RSFC, BGFC and TSFC matrices, we used the CONN toolbox release 21.a (www.nitrc.org/projects/conn)[69]. The mean time series extracted from functional ROIs were filtered using a bandpass filter of 0.01 Hz to 0.1 Hz and corrected for head motion and physiological noise using 24 motion regressors[70] and 6 anatomical component-based noise correction (aCompCorr) regressors[71]. Pearson's r correlation coefficients were calculated for each pair of regions and then converted to Fisher's Z. Empirical RSFC and TSFC matrices were calculated based on the whole time series of denoised resting-state and task-state BOLD signals, respectively. To calculate the BGFC matrix for the block design task, we regressed out the stimulus-evoked haemodynamic responses (co-activations) from the time series using an FIR basis set with 64 post-stimulus time bins of 0.72 s duration[9]. For the event-related task, we used an FIR basis set with 32 post-stimulus time bins of 1 s duration.

Task-modulated functional connectivity

The TMFC matrices for the empirical and simulated data were constructed in a similar way. For all TMFC analyses, we considered differences between two task conditions: (1) "2-back" and "0-back" for the working memory task, (2) "Social interaction" and "Random interaction" for the social cognition task, (3) "Go" and "Stop" for the stop-signal task, (4) "Switch" and "No switch" for the task-switching task, and (5) "Cond A" and "Cond B" for simulated data. All TMFC matrices were calculated without and with FIR task regression prior to TMFC analysis. The FIR basis sets were the same as for the BGFC analysis.

Direct correlation difference

To calculate the direct correlation difference between task blocks, we removed the first six seconds in each block to account for transient haemodynamic effects[8,11]. We then concatenated the time series for each task condition, calculated Pearson's r correlation coefficients, and converted them to Fisher's Z. Finally, we calculated the Fisher's Z difference between the two conditions.

Psychophysiological interaction

The PPI approach applies a general linear model to reveal the relationships between the BOLD signal in the "target" region and several explanatory variables: (1) expected task-evoked haemodynamic responses (psychological regressor, reflecting co-activations), (2) BOLD signal in the "seed" region

(physiological regressor, reflecting task-independent connectivity, similar to BGFC), (3) element-by-element product of psychological and physiological variables (psychophysiological interaction regressor, reflecting TMFC), and (4) non-neuronal nuisance variables (head motion and physiological noise). The standard PPI (sPPI) method includes a single PPI regressor for the difference between task conditions[12]. The generalised PPI (gPPI) approach includes multiple PPI terms for each task condition, which enables us to consider more than two task conditions in a single general linear model[13]. To account for the fact that psycho-physiological interactions occur at the neuronal level, it was proposed to deconvolve the BOLD signal in the seed ROI and calculate the element-by-element product of the deconvolved physiological regressor and unconvolved psychological regressor[15].

The PPI matrices consist of the beta estimates for the PPI regressors. The beta estimates change when the "target" and "seed" regions are swapped. However, the directionality of the PPI approach is arbitrary, as it is based on the a priori designation of regions as either sources or targets rather than a dynamic model of causal effects posed at the neuronal level combined with a forward biophysical model linking neuronal dynamics with haemodynamic responses[29]. The correlational PPI (cPPI) approach is a modification of the sPPI approach that avoids arbitrary directionality assumptions[14]. It is based on the partial correlation between PPI terms, removing the variance associated with the psychological and physiological regressors. To calculate the group-mean cPPI matrices, we converted partial correlations to Fisher's Z.

For all PPI analyses, we used SPM's volume of interest function to extract time series from the ROIs. Representative ROI time series were extracted as first eigenvariates after signal pre-whitening (using an autoregressive model of order one, AR(1) model), high-pass filtering (cut-off of 128 s), and adjustment for effects of no interest (session-specific constant terms, head motion and aCompCorr regressors). For more details about different PPI approaches, see Supplementary Information 4.

Beta-series correlation

The BSC approach is based on the correlations between beta estimates of stimulus-evoked haemodynamic responses for individual task trials. According to the originally proposed BSC approach, all events are modelled by separate regressors in a single general linear model[16]. This BSC approach, also known as the Least Squares All approach (LSA), was initially developed for slow event-related designs. However, the LSA approach is unsuitable for rapid event-related designs for two reasons. First, beta estimates can become unreliable due to correlations between regressors for individual trials. Shorter ISIs lead to higher correlations between regressors. This multicollinearity results in highly unstable estimates due to a limited amount of variability unique to an individual event[17]. Second, rapid event-related designs consist of many trial repetitions due to the short ISI. As the number of individual trial regressors approaches the number of time points, the design matrix becomes close to singular (e.g., the stop-signal task consists of 126 trials and 182 time points) and cannot be reliably inverted using ordinary least squares[72].

The alternative approach, called the Least Squares Separate approach (LSS), solves this problem by modelling each individual event against all other events with separate general linear models[17]. We considered both BSC-LSA and BSC-LSS approaches. To obtain the BSC matrices, we calculated Pearson's r correlations between the mean beta values extracted from ROIs separately for each task condition, converted them to Fisher's Z and calculated the difference between Fisher's Z values. For more details about different BSC approaches, see Supplementary Information 5.

Task-modulated effective connectivity: regression dynamic causal modelling

The most popular EC method is DCM, which employs an explicit forward or generative model of how observed haemodynamic signals emerge from hidden neuronal and biophysical states[24]. Originally, DCM was introduced for fMRI data to describe changes in hidden neuronal states via a bilinear differential equation. It enables us to estimate the effective strength of task-independent (intrinsic)

synaptic connections (*A* matrix), task-modulated (extrinsic) synaptic connections (*B* matrix), and the direct influence of driving inputs that cause activations (*C* matrix). The problem is that the original DCM is limited to parameter estimation of small brain networks (approximately ten nodes), as the model inversion becomes ill-posed and computationally demanding for large-scale (whole-brain) networks. To solve this problem, the regression DCM (rDCM) approach, which applies several simplifications to the original DCM, was proposed[42]. For more information about the rDCM approach, see Supplementary Information 6.

The task-modulated effective connectivity (TMEC) matrix cannot be directly obtained using the rDCM approach because it is based on the linear neural state equation without the *B* matrix. If we feed the entire time series of the resting-state or task-state BOLD signal into rDCM, the *A* matrix will reflect the resting-state effective connectivity (RSEC) and task-state effective connectivity (TSEC) matrices, respectively (cf. Fig. 1). In the latter case, the *A* matrix will depend on both spontaneous (intrinsic) and task-modulated (extrinsic) oscillations. To calculate the TMEC matrix, we propose calculating two *A* matrices for concatenated "Cond A" and "Cond B" block time series after removing the first six seconds of each block. The difference between these matrices will subtract spontaneous (intrinsic) EC and result in TMEC[40]. The driving input *C* matrix was set to zero, as we removed the transition periods.

Comparison of functional connectivity matrices

For the empirical data, we considered unthresholded weighted FC matrices. To evaluate similarity between the group-mean RSFC, TSFC, and TMFC matrices produced by different methods, we calculated the Pearson's r correlations between the lower diagonal elements. To evaluate the asymmetry of the raw (i.e., before symmetrisation procedure) group-mean and individual gPPI matrices, we calculated the Pearson's r correlations between the upper and lower diagonal elements.

For the simulated data, we compared the ground-truth TMFC matrix and thresholded binary TMFC matrices to evaluate the sensitivity and specificity of distinct TMFC methods (Supplementary Information 1). To obtain thresholded binary matrices, we determined significant FC differences between "Cond A" and "Cond B" using a two-sided one-sample t test. The Benjamini–Hochberg procedure[73] was used to control for the false discovery rate (FDR) at the 0.001 level.

Additionally, we used Bayesian analysis to provide evidence for no difference between the sPPI and gPPI methods. Bayesian comparisons were based on the default t test with a Cauchy (0, $r = 1/2^{0.5}$) prior implemented in JASP software (https://jasp-stats.org).

Ethics declaration

This study involves an analysis of simulated fMRI data and open-access fMRI datasets. The HCP data were acquired using protocols approved by the Washington University institutional review board. The CNP procedures were approved by the Institutional Review Boards at UCLA and the Los Angeles County Department of Mental Health.

Data availability

The HCP data are available at https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release. The CNP data are available at https://openfmri.org/dataset/ds000030. Simulated data are available at https://github.com/Masharipov/TMFC_simulations.

Code availability

Python and MATLAB code for task-modulated functional connectivity simulations, along with user-friendly Jupyter notebooks, are available at https://github.com/Masharipov/TMFC_simulations. The

SPM12-based TMFC toolbox with graphical user interface (GUI) for task-modulated functional connectivity analysis is available at https://github.com/Masharipov/TMFC_toolbox.

Methods references

45. Glasser, M. F. et al. The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80**, 105–124 (2013).
46. Gorgolewski, K. J., Durnez, J. & Poldrack, R. A. Preprocessed Consortium for Neuropsychiatric Phenomics dataset. *F1000Research* **6**, 1262 (2017).
47. Seitzman, B. A. et al. A set of functionally-defined brain regions with improved representation of the subcortex and cerebellum. *NeuroImage* **206**, 116290 (2020).
48. Cabral, J., Kringelbach, M. L. & Deco, G. Functional connectivity dynamically evolves on multiple time-scales over a static structural connectome: Models and mechanisms. *NeuroImage* **160**, 84–96 (2017).
49. Papadopoulos, L., Lynn, C. D., Battaglia, D. & Bassett, D. S. Relations between large-scale brain connectivity and effects of regional stimulation depend on collective dynamical state. *PLOS Computational Biology* **16**, e1008144 (2020).
50. Messé, A., Rudrauf, D., Giron, A. & Marrelec, G. Predicting functional connectivity from structural connectivity via computational models using MRI: An extensive comparison study. *NeuroImage* **111**, 65–75 (2015).
51. Ringo, J. L., Doty, R. W., Demeter, S. & Simard, P. Time Is of the Essence: A Conjecture that Hemispheric Specialization Arises from Interhemispheric Conduction Delay. *Cerebral Cortex* **4**, 331–343 (1994).
52. Brosch, M., Budinger, E. & Scheich, H. Stimulus-Related gamma oscillations in primate auditory cortex. *Journal of Neurophysiology* **87**, 2715–2725 (2002).
53. Liu, J. & Newsome, W. T. Local Field Potential in Cortical Area MT: stimulus tuning and behavioral correlations. *The Journal of Neuroscience* **26**, 7779–7790 (2006).
54. Schoffelen, J.-M., Oostenveld, R. & Fries, P. Neuronal coherence as a mechanism of effective corticospinal interaction. *Science* **308**, 111–113 (2005).
55. Engel, A., Fries, P. & Singer, W. Dynamic predictions: Oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience* **2**, 704–716 (2001).
56. Fries, P. Neuronal Gamma-Band synchronization as a fundamental process in cortical computation. *Annual Review of Neuroscience* **32**, 209–224 (2009).
57. Fries, P. Rhythms for Cognition: Communication through Coherence. *Neuron* **88**, 220–235 (2015).
58. Palmigiano, A., Geisel, T., Wolf, F. & Battaglia, D. Flexible information routing by transient synchrony. *Nature Neuroscience* **20**, 1014–1022 (2017).
59. Leopold, D. A., Murayama, Y. & Logothetis, N. K. Very slow activity fluctuations in monkey visual cortex: Implications for functional brain imaging. *Cerebral Cortex* **13**, 422–433 (2003).
60. Nir, Y. et al. Interhemispheric correlations of slow spontaneous neuronal fluctuations revealed in human sensory cortex. *Nature Neuroscience* **11**, 1100–1108 (2008).
61. Logothetis, N. K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–157 (2001).
62. Niessing, J. et al. Hemodynamic Signals Correlate Tightly with Synchronized Gamma Oscillations. *Science* **309**, 948–951 (2005).
63. Nir, Y. et al. Coupling between Neuronal Firing Rate, Gamma LFP, and BOLD fMRI Is Related to Interneuronal Correlations. *Current Biology* **17**, 1275–1285 (2007).
64. Shmuel, A. & Leopold, D. A. Neuronal correlates of spontaneous fluctuations in fMRI signals in monkey visual cortex: Implications for functional connectivity at rest. *Human Brain Mapping* **29**, 751–761 (2008).
65. Schölvinck, M. L., Maier, A., Ye, F. Q., Duyn, J. H. & Leopold, D. A. Neural basis of global resting-state fMRI activity. *Proceedings of the National Academy of Sciences* **107**, 10238–10243 (2010).
66. Ulloa, A. & Horwitz, B. Embedding Task-Based Neural Models into a Connectome-Based Model of the Cerebral Cortex. Frontiers in *Neuroinformatics* **10**, (2016).
67. Deco, G. et al. Single or multiple frequency generators in on-going brain activity: A mechanistic whole-brain model of empirical MEG data. *NeuroImage* **152**, 538–550 (2017).
68. Lachaux, J.-P., Rodríguez, E., Martinerie, J. & Varela, F. J. Measuring phase synchrony in brain signals. *Human Brain Mapping* **8**, 194–208 (1999).
69. Whitfield-Gabrieli, S. & Nieto-Castañón, A. Conn: A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks. *Brain Connectivity* **2**, 125–141 (2012).
70. Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S. J. & Turner, R. Movement-Related effects in fMRI time-series. *Magnetic Resonance in Medicine* **35**, 346–355 (1996).
71. Behzadi, Y., Restom, K., Liau, J. & Liu, T. T. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage* **37**, 90–101 (2007).
72. Di, X., Zhang, Z. & Biswal, B. B. Understanding psychophysiological interaction and its relations to beta series correlation. *Brain Imaging and Behavior* **15**, 958–973 (2020).
73. Benjamini, Y. & Hochberg, Y. Controlling the false discovery Rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-methodological* **57**, 289–300 (1995).

Acknowledgements

The current study was supported by the program of fundamental studies of IHB RAS (theme number 122041500046-5). Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research and by the McDonnell Center for Systems Neuroscience at Washington University. Another part of the data was provided by the Consortium for Neuropsychiatric Phenomics (NIH Roadmap for Medical Research grants UL1-

DE019580, RL1MH083268, RL1MH083269, RL1DA024853, RL1MH083270, RL1LM009833, PL1MH083271, and PL1NS062410).

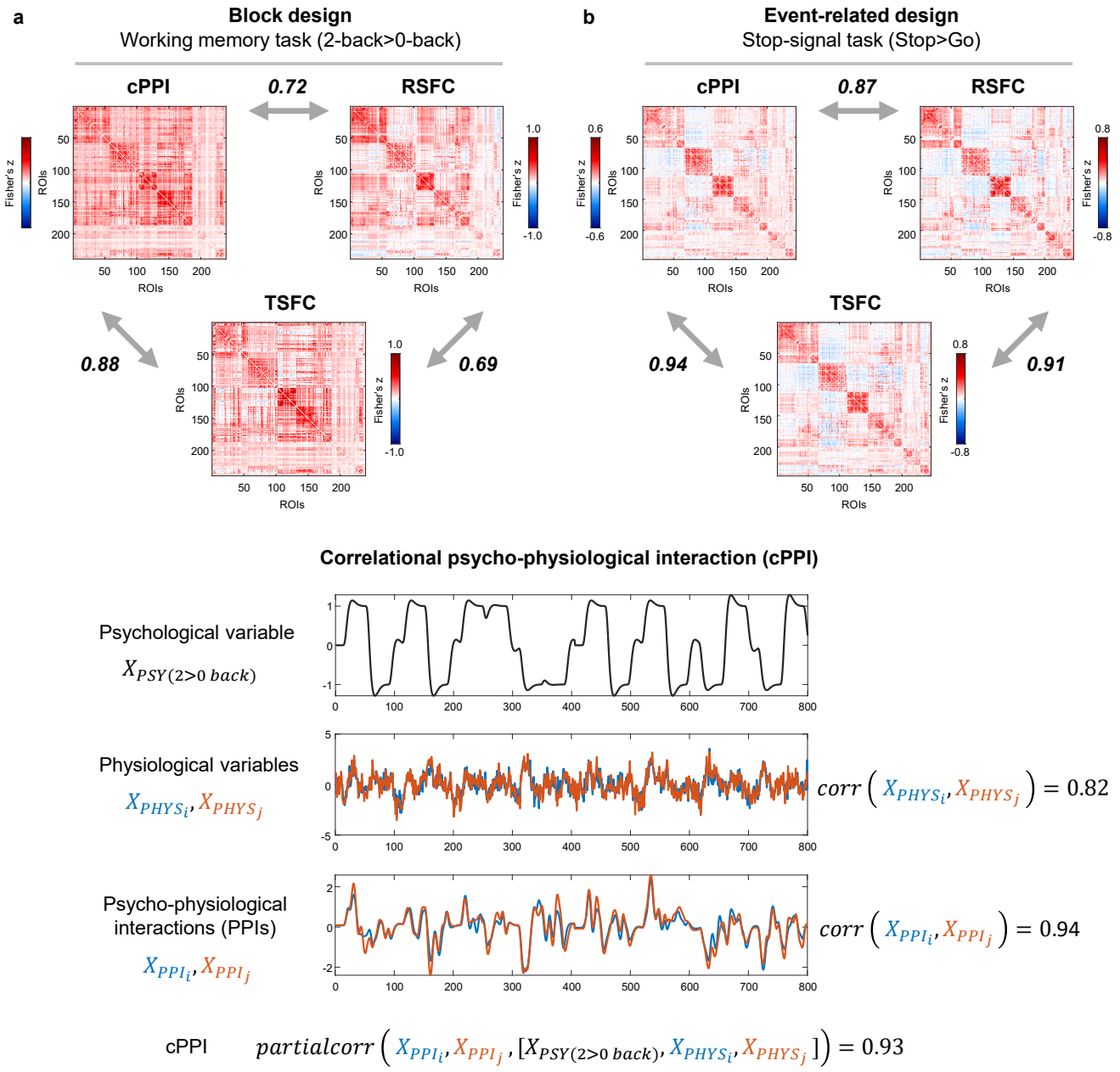
Author contributions

RM: Conceptualisation, Methodology, Software, Formal analysis, Writing – original draft, review & editing, Visualisation. IK: Methodology, Software, Formal analysis, Writing – review & editing, Visualisation. AK: Writing – review & editing. DC: Writing – review & editing, Funding acquisition. MK: Conceptualisation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

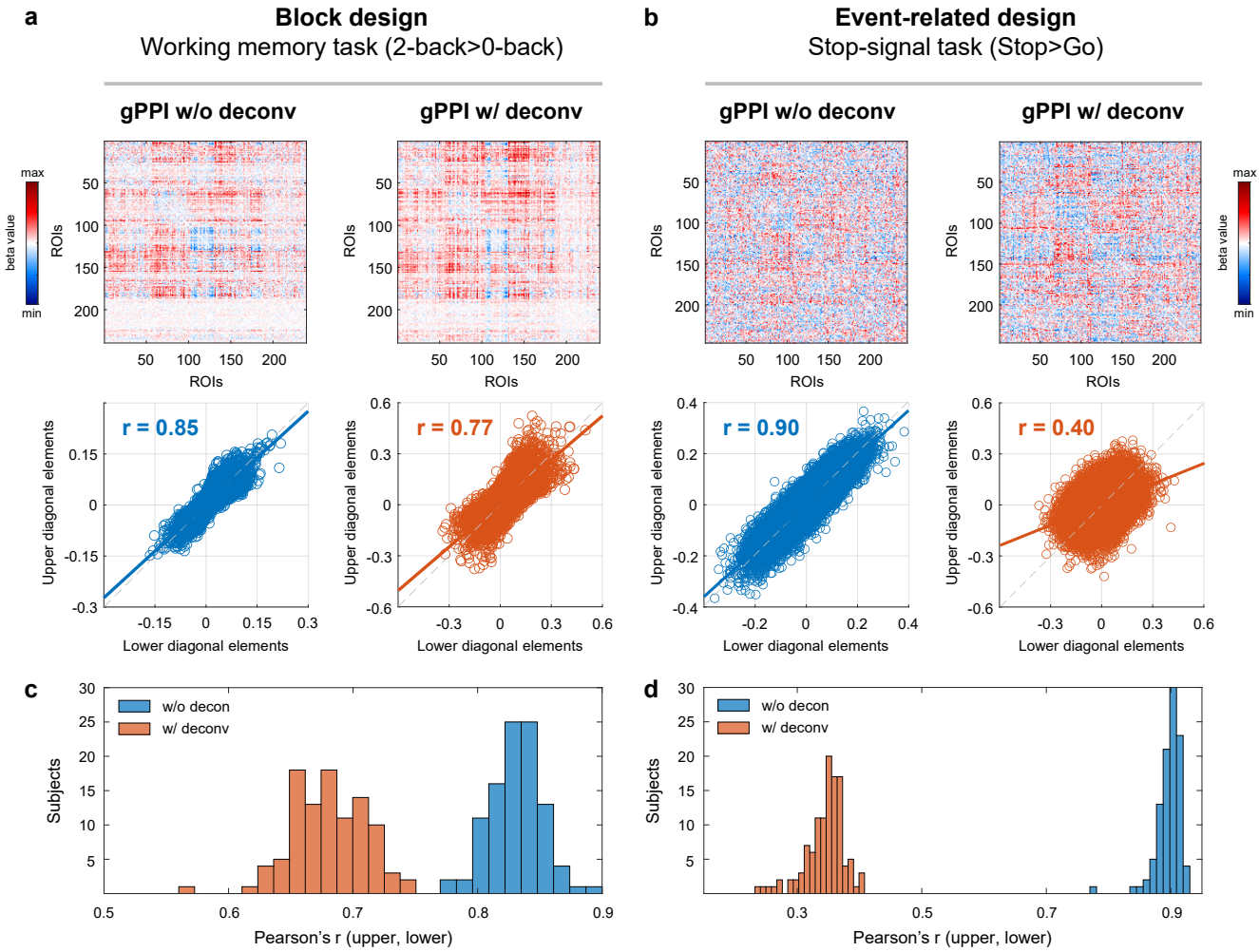
Competing interests

The authors declare no competing interests.

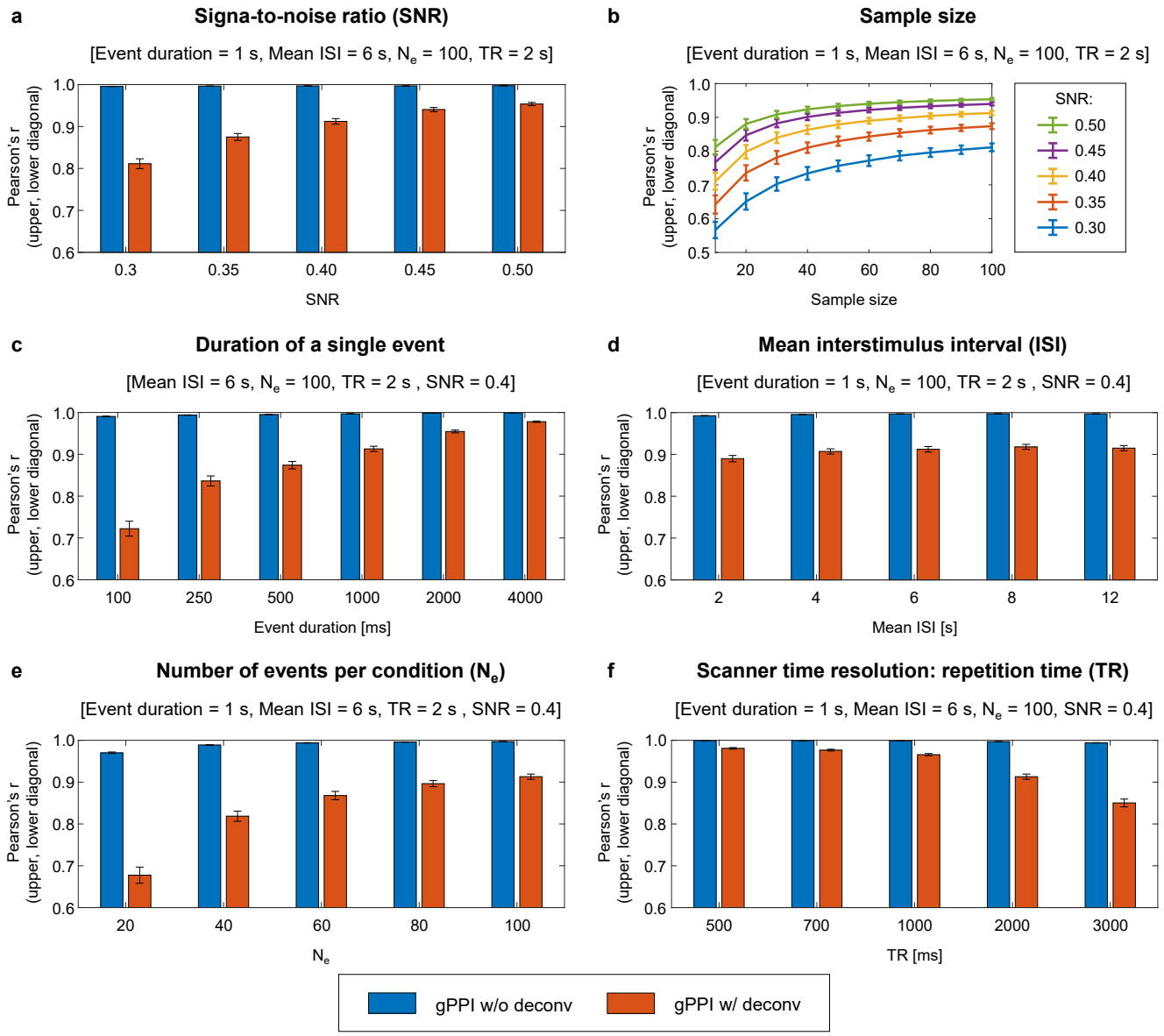
Extended data



Extended Data Fig. 1. The cPPI approach fails to eliminate spontaneous task-independent activity.
The cPPI approach produces FC matrices similar to the RSFC and TSFC matrices. **(a)** Correlations for the block design task (working memory task) and resting-state data from the HCP dataset. **(b)** Correlations for the event-related design task (stop-signal task) and resting-state data from the CNP dataset. **(c)** Illustration of the scPPI approach. The working memory task time series were taken from the two ROIs as an example. All PPI terms were calculated with the deconvolution step.



Extended Data Fig. 2. Asymmetry of gPPI matrices in block and event-related designs. **(a, b)** Group-mean gPPI matrices without symmetrisation procedure for the block and event-related design tasks (upper panels). Scatterplots for the upper and lower diagonal elements of the corresponding gPPI matrices (lower panels). **(c, d)** Histograms of correlations between the upper and lower diagonal elements of individual subjects' matrices. Blue and red colours represent gPPI results without (w/o) and with (w/) the deconvolution step, respectively.



Extended Data Fig. 3. Asymmetry of gPPI matrices in block and event-related designs. We calculated correlations between the upper and lower diagonal elements of the group-mean gPPI matrices to assess matrix asymmetry. By default, we considered the event-related design to consist of one hundred 1 s events per condition, mean ISI = 6 s, TR = 2 s, and sample size N = 100. **(a)** Dependence of gPPI matrix asymmetry on SNR. **(b)** Asymmetry depending on sample size. **(c)** Asymmetry depending on the duration of a single event. **(d)** Asymmetry depending on the mean interstimulus (ISI) interval. **(e)** Asymmetry depending on the number of events per condition (N_e). **(f)** Asymmetry depending on the repetition time (TR) for a fixed scan time (23.6 minutes). Error bars represent SD and were computed from 1000 random resamplings with replacement. Blue and red bars represent gPPI matrix symmetry without (w/o) and with (w/) the deconvolution step, respectively.

Extended Data Table 1. Marginal mean comparison of the sensitivity of TMFC methods for the block design.

Comparison		Mean	SE	t	p_{bonf}	Prior Odds	Posterior Odds	$Log(BF_{10, U})$
CorrDiff	gPPI (w/ dec)	-29.757	0.053	-559.255	< .001	0.320	∞	1215.642
	gPPI (w/o dec)	-2.659	0.053	-49.975	< .001	0.320	772.877	7.791
	sPPI (w/ dec)	-30.126	0.053	-566.197	< .001	0.320	∞	1248.660
	sPPI (w/o dec)	-2.901	0.053	-54.513	< .001	0.320	6437.832	9.911
gPPI (w/ dec)	gPPI (w/o dec)	27.098	0.053	509.281	< .001	0.320	∞	994.462
	sPPI (w/ dec)	**-0.369**	**0.053**	**-6.942**	**< .001**	**0.320**	**0.009**	**-3.595**
	sPPI (w/o dec)	26.856	0.053	504.742	< .001	0.320	∞	973.814
gPPI (w/o dec)	sPPI (w/ dec)	-27.467	0.053	-516.222	< .001	0.320	∞	1024.142
	sPPI (w/o dec)	**-0.241**	**0.053**	**-4.539**	**< .001**	**0.320**	**0.008**	**-3.700**
sPPI (w/ dec)	sPPI (w/o dec)	27.226	0.053	511.684	< .001	0.320	∞	1003.137

The marginal means are averaged over the levels of the signal-to-noise ratio. p values were adjusted for 10 t tests using Bonferroni correction. Bayesian comparisons are based on the default t test with a Cauchy (0, $r = 1/2^{0.5}$) prior. The posterior odds have been corrected for multiple testing by fixing the prior probability that the null hypothesis holds across all comparisons to 0.5. The "U" in the Bayes factor denotes that it is uncorrected. PPI terms were calculated with (w/) and without (w/o) the deconvolution step (dec). sPPI and gPPI matrices were symmetrised. Comparisons that show at least "substantial" evidence in favour of the null hypothesis ($LogBF_{10} < 1.1$) are highlighted in bold.

Extended Data Table 2. Marginal mean comparison of the sensitivity of TMFC methods for the event-related design.

Comparison		Mean	SE	t	p_{bonf}	Prior Odds	Posterior Odds	$Log(BF_{10, U})$
BSC-LSA	BSC-LSS	-40.601	0.042	-963.775	< .001	0.260	∞	2622.547
	gPPI (w/ dec)	-39.015	0.042	-926.125	< .001	0.260	∞	2372.147
	gPPI (w/o dec)	**-0.856**	**0.042**	**-20.317**	**< .001**	**0.260**	**0.020**	**-2.556**
	sPPI (w/ dec)	-39.188	0.042	-930.251	< .001	0.260	∞	2397.846
	sPPI (w/o dec)	**-0.892**	**0.042**	**-21.183**	**< .001**	**0.260**	**0.023**	**-2.447**
BSC-LSS	gPPI (w/ dec)	1.586	0.042	37.650	< .001	0.260	3.884	2.704
	gPPI (w/o dec)	39.745	0.042	943.458	< .001	0.260	∞	2530.145
	sPPI (w/ dec)	1.412	0.042	33.524	< .001	0.260	1.044	1.390
	sPPI (w/o dec)	39.708	0.042	942.592	< .001	0.260	∞	2530.481
gPPI (w/ dec)	gPPI (w/o dec)	38.159	0.042	905.808	< .001	0.260	∞	2283.728
	sPPI (w/ dec)	**-0.174**	**0.042**	**-4.126**	**< .001**	**0.260**	**0.006**	**-3.718**
	sPPI (w/o dec)	38.122	0.042	904.942	< .001	0.260	∞	2283.728
gPPI (w/o dec)	sPPI (w/ dec)	-38.332	0.042	-909.934	< .001	0.260	∞	2309.017
	sPPI (w/o dec)	**-0.037**	**0.042**	**-0.867**	**1.000**	**0.260**	**0.006**	**-3.790**
sPPI (w/ dec)	sPPI (w/o dec)	38.296	0.042	909.068	< .001	0.260	∞	2309.050

The marginal means are averaged over the levels of the signal-to-noise ratio. p values were adjusted for 15 t tests using Bonferroni correction. Bayesian comparisons are based on the default t test with a Cauchy (0, $r = 1/2^{0.5}$) prior. The posterior odds have been corrected for multiple testing by fixing the prior probability that the null hypothesis holds across all comparisons to 0.5. The "U" in the Bayes factor denotes that it is uncorrected. PPI terms were calculated with (w/) and without (w/o) the deconvolution step (dec). sPPI and gPPI matrices were symmetrised. Comparisons that show at least "substantial" evidence in favour of the null hypothesis ($LogBF_{10} < 1.1$) are highlighted in bold.

Extended Data Table 3. Pairwise comparison of the sensitivity of TMFC methods for event-related designs with different event durations.

Comparison		Mean	SE	t	p_{bonf}	Prior Odds	Posterior Odds	Log($BF_{10, U}$)
				Event duration = 100 ms				
BSC-LSA	BSC-LSS	-13.995	0.096	-145.989	< .001	0.059	∞	2358.555
	gPPI (w/ dec)	-11.968	0.096	-124.837	< .001	0.059	∞	2489.800
	sPPI (w/ dec)	-11.944	0.096	-124.589	< .001	0.059	∞	2471.492
BSC-LSS	gPPI (w/ dec)	2.028	0.096	21.153	< .001	0.059	$6.822\times10^{+49}$	117.569
	sPPI (w/ dec)	2.052	0.096	21.400	< .001	0.059	$4.661\times10^{+50}$	119.491
gPPI (w/ dec)	sPPI (w/ dec)	0.024	0.096	0.247	1.000	0.059	0.003	-2.968
				Event duration = 250 ms				
BSC-LSA	BSC-LSS	-54.516	0.096	-568.674	< .001	0.059	∞	4464.545
	gPPI (w/ dec)	-47.253	0.096	-492.905	< .001	0.059	∞	4389.013
	sPPI (w/ dec)	-47.354	0.096	-493.961	< .001	0.059	∞	4364.078
BSC-LSS	gPPI (w/ dec)	7.264	0.096	75.769	< .001	0.059	∞	724.413
	sPPI (w/ dec)	7.162	0.096	74.713	< .001	0.059	$2.608\times10^{+303}$	701.464
gPPI (w/ dec)	sPPI (w/ dec)	-0.101	0.096	-1.056	1.000	0.059	0.004	-2.757
				Event duration = 500 ms				
BSC-LSA	BSC-LSS	-62.376	0.096	-650.665	< .001	0.059	∞	4834.299
	gPPI (w/ dec)	-57.776	0.096	-602.673	< .001	0.059	∞	4648.812
	sPPI (w/ dec)	-58.079	0.096	-605.840	< .001	0.059	∞	4700.323
BSC-LSS	gPPI (w/ dec)	4.601	0.096	47.991	< .001	0.059	$4.219\times10^{+216}$	501.620
	sPPI (w/ dec)	4.297	0.096	44.825	< .001	0.059	$5.011\times10^{+200}$	464.951
gPPI (w/ dec)	sPPI (w/ dec)	-0.304	0.096	-3.167	0.426	0.059	0.052	-0.126
				Event duration = 1000 ms				
BSC-LSA	BSC-LSS	-53.705	0.096	-560.211	< .001	0.059	∞	4917.987
	gPPI (w/ dec)	-52.539	0.096	-548.052	< .001	0.059	∞	4883.486
	sPPI (w/ dec)	-52.721	0.096	-549.943	< .001	0.059	∞	4907.266
BSC-LSS	gPPI (w/ dec)	1.166	0.096	12.159	< .001	0.059	$2.242\times10^{+68}$	160.206
	sPPI (w/ dec)	0.984	0.096	10.269	< .001	0.059	$3.346\times10^{+51}$	121.462
gPPI (w/ dec)	sPPI (w/ dec)	-0.181	0.096	-1.891	1.000	0.059	0.296	1.604
				Event duration = 2000 ms				
BSC-LSA	BSC-LSS	-14.754	0.096	-153.907	< .001	0.059	∞	3472.149
	gPPI (w/ dec)	-14.717	0.096	-153.512	< .001	0.059	∞	3465.769
	sPPI (w/ dec)	-14.721	0.096	-153.555	< .001	0.059	∞	3466.808
BSC-LSS	gPPI (w/ dec)	0.038	0.096	0.395	1.000	0.059	$1.392\times10^{+14}$	35.389
	sPPI (w/ dec)	0.034	0.096	0.352	1.000	0.059	$1.347\times10^{+12}$	30.752
gPPI (w/ dec)	sPPI (w/ dec)	-0.004	0.096	-0.043	1.000	0.059	0.004	-2.591
				Event duration = 4000 ms				
BSC-LSA	BSC-LSS	-0.509	0.096	-5.309	< .001	0.059	∞	1090.983
	gPPI (w/ dec)	-0.509	0.096	-5.310	< .001	0.059	∞	1091.225
	sPPI (w/ dec)	-0.509	0.096	-5.310	< .001	0.059	∞	1091.225
BSC-LSS	gPPI (w/ dec)	-8.000×10^{-5}	0.096	-8.345×10^{-4}	1.000	0.059	0.008	-1.997
	sPPI (w/ dec)	-8.000×10^{-5}	0.096	-8.345×10^{-4}	1.000	0.059	0.008	-1.997
gPPI (w/ dec)	sPPI (w/ dec)	-1.107×10^{-11}	0.096	-1.155×10^{-10}	1.000	0.059	NaN	NaN

p values were adjusted for 36 t tests using Bonferroni correction. Bayesian comparisons are based on the default t test with a Cauchy (0, r = $1/2^{0.5}$) prior. The posterior odds have been corrected for multiple testing by fixing the prior probability that the null hypothesis holds across all comparisons to 0.5. The "U" in the Bayes factor denotes that it is uncorrected. PPI terms were calculated with the deconvolution step (w/dec). sPPI and gPPI matrices were symmetrised. NaN – data are essentially constant (sPPI and gPPI methods reached 100% sensitivity).

Extended Data Table 4. Pairwise comparison of the sensitivity of TMFC methods for event-related designs with different interstimulus intervals.

Comparison		Mean	SE	t	p_{bonf}	Prior Odds	Posterior Odds	Log($BF_{10, U}$)
				Mean ISI = 2 s				
BSC-LSA	BSC-LSS	-73.642	0.084	-878.110	< .001	0.072	∞	5659.180
	gPPI (w/ dec)	-79.141	0.084	-943.686	< .001	0.072	∞	6077.271
	sPPI (w/ dec)	-79.639	0.084	-949.620	< .001	0.072	∞	6171.758
BSC-LSS	gPPI (w/ dec)	-5.500	0.084	-65.577	< .001	0.072	$4.574\times10^{+279}$	646.576
	sPPI (w/ dec)	-5.997	0.084	-71.511	< .001	0.072	∞	751.354
gPPI (w/ dec)	sPPI (w/ dec)	-0.498	0.084	-5.934	< .001	0.072	28.379	5.980
				Mean ISI = 4 s				
BSC-LSA	BSC-LSS	-79.651	0.084	-949.765	< .001	0.072	∞	5991.980
	gPPI (w/ dec)	-81.047	0.084	-966.411	< .001	0.072	∞	6229.187
	sPPI (w/ dec)	-81.248	0.084	-968.802	< .001	0.072	∞	6265.782
BSC-LSS	gPPI (w/ dec)	-1.396	0.084	-16.645	< .001	0.072	$1.804\times10^{+53}$	125.261
	sPPI (w/ dec)	-1.596	0.084	-19.037	< .001	0.072	$9.991\times10^{+70}$	166.117
gPPI (w/ dec)	sPPI (w/ dec)	-0.201	0.084	-2.391	1.000	0.072	0.141	0.675
				Mean ISI = 6 s				
BSC-LSA	BSC-LSS	-53.705	0.084	-640.382	< .001	0.072	∞	4917.987
	gPPI (w/ dec)	-52.539	0.084	-626.483	< .001	0.072	∞	4883.486
	sPPI (w/ dec)	-52.721	0.084	-628.644	< .001	0.072	∞	4907.266
BSC-LSS	gPPI (w/ dec)	1.166	0.084	13.899	< .001	0.072	$2.707\times10^{+68}$	160.206
	sPPI (w/ dec)	0.984	0.084	11.738	< .001	0.072	$4.039\times10^{+51}$	121.462
gPPI (w/ dec)	sPPI (w/ dec)	-0.181	0.084	-2.161	1.000	0.072	0.357	1.604
				Mean ISI = 8 s				
BSC-LSA	BSC-LSS	-26.735	0.084	-318.789	< .001	0.072	∞	3637.117
	gPPI (w/ dec)	-24.487	0.084	-291.980	< .001	0.072	∞	3437.127
	sPPI (w/ dec)	-24.536	0.084	-292.569	< .001	0.072	∞	3447.562
BSC-LSS	gPPI (w/ dec)	2.248	0.084	26.809	< .001	0.072	$3.678\times10^{+296}$	685.502
	sPPI (w/ dec)	2.199	0.084	26.220	< .001	0.072	$5.130\times10^{+292}$	676.624
gPPI (w/ dec)	sPPI (w/ dec)	-0.049	0.084	-0.589	1.000	0.072	0.006	-2.553
				Mean ISI = 12 s				
BSC-LSA	BSC-LSS	-8.093	0.084	-96.505	< .001	0.072	∞	2144.008
	gPPI (w/ dec)	-4.093	0.084	-48.810	< .001	0.072	∞	945.145
	sPPI (w/ dec)	-4.281	0.084	-51.046	< .001	0.072	∞	1020.727
BSC-LSS	gPPI (w/ dec)	4.000	0.084	47.695	< .001	0.072	∞	1416.543
	sPPI (w/ dec)	3.812	0.084	45.459	< .001	0.072	∞	1393.549
gPPI (w/ dec)	sPPI (w/ dec)	-0.188	0.084	-2.236	1.000	0.072	0.656	2.212

p values were adjusted for 30 t tests using Bonferroni correction. Bayesian comparisons are based on the default t test with a Cauchy (0, r = $1/2^{0.5}$) prior. The posterior odds have been corrected for multiple testing by fixing the prior probability that the null hypothesis holds across all comparisons to 0.5. The "U" in the Bayes factor denotes that it is uncorrected. PPI terms were calculated with the deconvolution step (w/dec). sPPI and gPPI matrices were symmetrised.

Extended Data Table 5. Pairwise comparison of the sensitivity of TMFC methods for event related designs with different numbers of events per condition.

Comparison		Mean	SE	t	p_{bonf}	Prior Odds	Posterior Odds	Log($BF_{10, U}$)
				Number of events = 20				
BSC-LSA	BSC-LSS	-6.437	0.115	-56.026	< .001	0.072	∞	1276.956
	gPPI (w/ dec)	-4.597	0.115	-40.008	< .001	0.072	∞	1063.888
	sPPI (w/ dec)	-5.213	0.115	-45.371	< .001	0.072	∞	1225.250
BSC-LSS	gPPI (w/ dec)	1.840	0.115	16.018	< .001	0.072	$2.454\times10^{+61}$	143.990
	sPPI (w/ dec)	1.224	0.115	10.656	< .001	0.072	$5.334\times10^{+26}$	64.176
gPPI (w/ dec)	sPPI (w/ dec)	-0.616	0.115	-5.362	< .001	0.072	$2.074\times10^{+7}$	19.482
				Number of events = 40				
BSC-LSA	BSC-LSS	-30.593	0.115	-266.263	< .001	0.072	∞	3204.757
	gPPI (w/ dec)	-29.627	0.115	-257.853	< .001	0.072	∞	3370.030
	sPPI (w/ dec)	-30.063	0.115	-261.647	< .001	0.072	∞	3362.369
BSC-LSS	gPPI (w/ dec)	0.966	0.115	8.410	< .001	0.072	345085.451	15.386
	sPPI (w/ dec)	0.530	0.115	4.616	< .001	0.072	0.834	2.453
gPPI (w/ dec)	sPPI (w/ dec)	-0.436	0.115	-3.794	0.028	0.072	0.290	1.395
				Number of events = 60				
BSC-LSA	BSC-LSS	-55.639	0.115	-484.251	< .001	0.072	∞	4454.717
	gPPI (w/ dec)	-52.771	0.115	-459.289	< .001	0.072	∞	4468.663
	sPPI (w/ dec)	-53.374	0.115	-464.532	< .001	0.072	∞	4434.657
BSC-LSS	gPPI (w/ dec)	2.868	0.115	24.962	< .001	0.072	$1.233\times10^{+77}$	180.143
	sPPI (w/ dec)	2.266	0.115	19.719	< .001	0.072	$5.065\times10^{+46}$	110.175
gPPI (w/ dec)	sPPI (w/ dec)	-0.602	0.115	-5.243	< .001	0.072	39.868	6.320
				Number of events = 80				
BSC-LSA	BSC-LSS	-56.530	0.115	-492.004	< .001	0.072	∞	4772.541
	gPPI (w/ dec)	-52.463	0.115	-456.609	< .001	0.072	∞	4599.074
	sPPI (w/ dec)	-52.738	0.115	-459.000	< .001	0.072	∞	4611.084
BSC-LSS	gPPI (w/ dec)	4.067	0.115	35.395	< .001	0.072	$5.693\times10^{+266}$	616.861
	sPPI (w/ dec)	3.792	0.115	33.004	< .001	0.072	$1.555\times10^{+240}$	555.696
gPPI (w/ dec)	sPPI (w/ dec)	-0.275	0.115	-2.391	1.000	0.072	0.156	0.777
				Number of events = 100				
BSC-LSA	BSC-LSS	-53.705	0.115	-467.415	< .001	0.072	∞	4917.987
	gPPI (w/ dec)	-52.539	0.115	-457.270	< .001	0.072	∞	4883.486
	sPPI (w/ dec)	-52.721	0.115	-458.847	< .001	0.072	∞	4907.266
BSC-LSS	gPPI (w/ dec)	1.166	0.115	10.145	< .001	0.072	$2.707\times10^{+68}$	160.206
	sPPI (w/ dec)	0.984	0.115	8.568	< .001	0.072	$4.039\times10^{+51}$	121.462
gPPI (w/ dec)	sPPI (w/ dec)	-0.181	0.115	-1.577	1.000	0.072	0.357	1.604

p values were adjusted for 30 t tests using Bonferroni correction. Bayesian comparisons are based on the default t test with a Cauchy (0, r = $1/2^{0.5}$) prior. The posterior odds have been corrected for multiple testing by fixing the prior probability that the null hypothesis holds across all comparisons to 0.5. The "U" in the Bayes factor denotes that it is uncorrected. PPI terms were calculated with the deconvolution step (w/dec). sPPI and gPPI matrices were symmetrised.

Extended Data Table 6. Pairwise comparison of the sensitivity of TMFC methods for event-related designs with different repetition times.

Comparison		Mean	SE	t	p_{bonf}	Prior Odds	Posterior Odds	Log($BF_{10, U}$)
				TR = 500 ms				
BSC-LSA	BSC-LSS	-73.642	0.084	-878.110	< .001	0.072	0.016	-1.500
	gPPI (w/ dec)	-79.141	0.084	-943.686	< .001	0.072	0.016	-1.500
	sPPI (w/ dec)	-79.639	0.084	-949.620	< .001	0.072	0.016	-1.500
BSC-LSS	gPPI (w/ dec)	-5.500	0.084	-65.577	< .001	0.072	NaN	NaN
	sPPI (w/ dec)	-5.997	0.084	-71.511	< .001	0.072	NaN	NaN
gPPI (w/ dec)	sPPI (w/ dec)	-0.498	0.084	-5.934	< .001	0.072	NaN	NaN
				TR = 700 ms				
BSC-LSA	BSC-LSS	-79.651	0.084	-949.765	< .001	0.072	$7.303\times10^{+57}$	135.870
	gPPI (w/ dec)	-81.047	0.084	-966.411	< .001	0.072	$7.303\times10^{+57}$	135.870
	sPPI (w/ dec)	-81.248	0.084	-968.802	< .001	0.072	$7.303\times10^{+57}$	135.870
BSC-LSS	gPPI (w/ dec)	-1.396	0.084	-16.645	< .001	0.072	NaN	NaN
	sPPI (w/ dec)	-1.596	0.084	-19.037	< .001	0.072	NaN	NaN
gPPI (w/ dec)	sPPI (w/ dec)	-0.201	0.084	-2.391	1.000	0.072	NaN	NaN
				TR = 1000 ms				
BSC-LSA	BSC-LSS	-53.705	0.084	-640.382	< .001	0.072	∞	1534.286
	gPPI (w/ dec)	-52.539	0.084	-626.483	< .001	0.072	∞	1527.832
	sPPI (w/ dec)	-52.721	0.084	-628.644	< .001	0.072	∞	1526.842
BSC-LSS	gPPI (w/ dec)	1.166	0.084	13.899	< .001	0.072	$1.189\times10^{+23}$	55.767
	sPPI (w/ dec)	0.984	0.084	11.738	< .001	0.072	$1.352\times10^{+25}$	60.501
gPPI (w/ dec)	sPPI (w/ dec)	-0.181	0.084	-2.161	1.000	0.072	0.007	-2.391
				TR = 2000 ms				
BSC-LSA	BSC-LSS	-26.735	0.084	-318.789	< .001	0.072	∞	4917.987
	gPPI (w/ dec)	-24.487	0.084	-291.980	< .001	0.072	∞	4883.486
	sPPI (w/ dec)	-24.536	0.084	-292.569	< .001	0.072	∞	4907.266
BSC-LSS	gPPI (w/ dec)	2.248	0.084	26.809	< .001	0.072	$2.707\times10^{+68}$	160.206
	sPPI (w/ dec)	2.199	0.084	26.220	< .001	0.072	$4.039\times10^{+51}$	121.462
gPPI (w/ dec)	sPPI (w/ dec)	-0.049	0.084	-0.589	1.000	0.072	0.357	1.604
				TR = 3000 ms				
BSC-LSA	BSC-LSS	-8.093	0.084	-96.505	< .001	0.072	∞	4370.359
	gPPI (w/ dec)	-4.093	0.084	-48.810	< .001	0.072	∞	4916.715
	sPPI (w/ dec)	-4.281	0.084	-51.046	< .001	0.072	∞	5003.281
BSC-LSS	gPPI (w/ dec)	4.000	0.084	47.695	< .001	0.072	∞	1414.757
	sPPI (w/ dec)	3.812	0.084	45.459	< .001	0.072	∞	1498.276
gPPI (w/ dec)	sPPI (w/ dec)	-0.188	0.084	-2.236	1.000	0.072	0.210	1.073

p values were adjusted for 30 t tests using Bonferroni correction. Bayesian comparisons are based on the default t test with a Cauchy (0, r = 1/sqrt(2)) prior. The posterior odds have been corrected for multiple testing by fixing the prior probability that the null hypothesis holds across all comparisons to 0.5. The "U" in the Bayes factor denotes that it is uncorrected. PPI terms were calculated with the deconvolution step (w/dec). sPPI and gPPI matrices were symmetrised. NaN – data are essentially constant (sPPI, gPPI and BSC-LSS methods reached 100% sensitivity).