applied sciences

MDPI

Article

Nonparametric Correlation Estimator via Solving Cubic Equations and Its Application to Brain Functional Connectivity Analysis

Shenyuan Yang [1], Gary Green [2], Jian Zhang [3], André Gouws [2] and Jie Li [1,3,*]

[1] Information College, Lijiang Culture and Tourism College, Lijiang 674199, China; shenyuan.yang@ljwhlyxy.edu.cn
[2] York Neuroimaging Centre, University of York, Innovation Way, York YO10 5NY, UK; gary.green@york.ac.uk (G.G.); dr.andre.gouws@gmail.com (A.G.)
[3] School of Engineering, Mathematics and Physics, University of Kent, Canterbury CT2 7NF, UK; j.zhang-79@kent.ac.uk
* Correspondence: jie.li@ljwhlyxy.edu.cn

Abstract

In this paper, we propose a novel nonparametric estimator of the correlation coefficient, obtained as the root of the cubic score equation of a locally weighted Gaussian likelihood. The estimator is nonparametric in the functional form of the correlation but assumes a conditional bivariate Gaussian working model with standardised margins: pointwise consistency requires only correctly specified conditional moments, whereas Gaussianity determines the asymptotic variance. Furthermore, we adopt a local linear smoothing technique to correct the boundary effects, which are common in nonparametric estimation. We establish the consistency and asymptotic normality of the proposed estimator and demonstrate its performance through simulation studies with non-Gaussian data. Additionally, we apply our method to a single-participant resting-state MEG recording, estimating time-varying power-envelope correlations in six frequency bands and comparing them with static and sliding-window correlations.

Keywords: nonparametric estimation; cubic equation; Yeo-7 network; dynamic functional connectivity; time-varying correlation

1. Introduction

Human brain activity exhibits temporal dynamics, with the coupling between brain regions changing across multiple timescales in response to cognitive tasks, brain states, and disease progression [1–3]. Functional connectivity (FC), which characterises the statistical dependence between regional brain signals, has become an important framework for investigating the organisation of brain function, disease-related network abnormalities, and potential biomarkers [4–6]. Unlike fMRI, which relies on haemodynamic signals, MEG directly measures the magnetic fields generated by neuronal currents and provides millisecond-level temporal resolution for tracking transient changes in oscillatory phase and amplitude [7–9]. It therefore offers a particularly fine-grained temporal basis for investigating dynamic functional connectivity (dFC), beyond what can be captured by static measures based on time-averaged connectivity [8,9]. Consequently, recent research has increasingly shifted from static functional connectivity towards dFC, with growing interest in characterising time-varying patterns of brain networks, transitions between connectivity states, and their associations with cognitive processes and clinical phenotypes [1,6,10].

Check for updates

Academic Editor: Alexander N. Pisarchik

Received: 11 September 2026
Revised: 27 October 2026
Accepted: 5 November 2026
Published: 9 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

In neuropsychiatric and neurodegenerative disorders, alterations in brain network organisation may precede or accompany the emergence of clinical symptoms. In Alzheimer's disease (AD), for example, MEG studies have reported detectable alterations in the complexity of dynamic brain networks, transitions between connectivity states, and frequency-specific coupling, with some of these features showing potential for distinguishing patients from healthy controls [3,11]. These findings suggest that dFC may provide complementary information for early detection, disease stratification, and the assessment of disease progression. Consequently, current research into dynamic functional connectivity is broadly categorised into two complementary strands: fMRI-based approaches and MEG-based approaches, which are discussed in the sections that follow.

Dynamic functional connectivity based on fMRI. Owing to its relatively high spatial resolution and the availability of extensive publicly accessible datasets, fMRI has been the predominant data source for dFC research [12,13]. Over the past five years, fMRI-based dFC studies have expanded beyond static, time-averaged connectivity towards the characterisation of time-varying organisation and disease-related network abnormalities [14–16]. Consequently, attention in this field is increasingly shifting away from classification accuracy alone towards the stability, interpretability, and cross-dataset reproducibility of dynamic abnormality patterns.

Dynamic functional connectivity based on MEG. With millisecond-level temporal resolution, MEG is particularly well suited to investigating rapid neuronal oscillations and interactions between spatially distributed brain regions [6,8,17]. In recent years, MEG-based functional connectivity research has progressed from analyses of single-frequency bands or static group differences towards the joint characterisation of multiple frequency bands, time-varying connectivity states, and the temporal complexity of network dynamics [10]. Within the Alzheimer's disease spectrum, previous studies have reported an overall reduction or reorganisation of MEG-derived dFC, with differences particularly involving the α and β bands, frontotemporal regions, and the default mode network, as well as associations with cognitive function and structural alterations [11]. Measures of dynamic network complexity, state-transition characteristics, and multi-band deep-learning features have also shown potential for patient classification and early prediction of clinical progression [3].

Families of electrophysiological connectivity estimators. Connectivity measures for source-reconstructed MEG and EEG can be broadly divided into two families [18]. Phase-based measures estimate the consistency of phase differences between band-limited signals, such as the phase-locking value (PLV) [19], the imaginary part of coherency (IC) [20], the phase-lag index (PLI) and its weighted version (wPLI) [21,22], and the phase linearity measurement (PLM) [23]; PLI, wPLI, and PLM discard zero-lag coupling and are therefore insensitive to instantaneous source leakage. Amplitude-based measures, in contrast, quantify the co-fluctuation of band-limited amplitude or power envelopes. The amplitude- or power-envelope correlation (AEC/PEC), computed after leakage correction by symmetric orthogonalisation [4,6,24], has been reported as one of the most reproducible MEG connectivity metrics [25] and is partly dissociated from phase coupling [26]. Time-resolved versions of these measures are usually obtained with sliding windows, with hidden Markov models of the envelopes or of the phase coupling [27,28], or with instantaneous Hilbert-based measures [29]. Sliding-window estimates, however, depend strongly on the window length and can fluctuate even when the underlying process is stationary, which has motivated surrogate-based tests of dynamic connectivity [30–33]. Table 1 summarises these families.

Table 1. Families of MEG/EEG functional connectivity estimators.

Measure	Family	Zero-Lag Leakage	Typical Time-Resolved Variant
PLV [19]	phase	sensitive	sliding window; instantaneous phase difference
IC [20]	phase	insensitive	sliding window
PLI/wPLI [21,22]	phase	insensitive	sliding window; HMM phase states [28]
PLM [23]	phase	insensitive	sliding window
AEC/PEC [4,24]	amplitude	corrected	sliding window; HMM on envelopes [27]
Proposed CL estimator	amplitude	corrected	kernel smoothing, pair-specific bandwidth

Nevertheless, existing MEG-dFC analyses predominantly rely on sliding-window approaches, epoch-based aggregation, amplitude-envelope correlations, phase synchronisation, or adaptive segmentation methods, and the resulting estimates remain sensitive to choices of window length and time-frequency smoothing [3,10]. To improve the estimation of direct connectivity and the characterisation of information flow, recent work has incorporated structural connectivity priors, partial coherence analysis, and time-lagged multidimensional pattern connectivity approaches [34]. Overall, MEG-dFC research is moving beyond demonstrating the presence of connectivity abnormalities towards the robust extraction, statistical interpretation, and independent validation of dynamic abnormality patterns. However, statistical inference frameworks that fully exploit the high temporal resolution of MEG data remain underdeveloped.

The estimator proposed here belongs to the amplitude-based family. It can be viewed as a likelihood-based kernel generalisation of sliding-window PEC, in which the rectangular window is replaced by a smooth kernel whose width is chosen by blocked cross-validation separately for every pair of regions. The estimate always lies in $(-1, 1)$, and it is corrected for boundary bias at the start and end of the recording, where sliding windows must be truncated; its bias, variance, and convergence rate are known. Existing nonparametric approaches to time-varying correlation, including sliding-window Pearson correlation, Nadaraya–Watson smoothing, and local polynomial likelihood [35–39], either lack one of these properties or become computationally demanding when applied to all region pairs of a whole-brain parcellation. We therefore aim to provide a pairwise, data-adaptive, and statistically characterised time-varying PEC estimator that is fast enough for tens of thousands of region pairs. The estimator does not measure phase coupling, and we do not propose it as a replacement for phase-based measures.

Concretely, we combine kernel-weighted local maximum likelihood with a cubic estimating equation for the correlation, determine a bandwidth separately for each pair, and assemble the pairwise estimates into a time-varying correlation matrix. Because independently estimated pairwise correlations need not form a valid correlation matrix, the assembled matrix is corrected to be positive definite before network-level summaries are computed (Section 4).

The proposed estimator retains desirable asymptotic properties, including consistency and asymptotic normality, while allowing the bandwidth to adapt to the local characteristics of each variable pair. This pair-specific bandwidth selection provides greater flexibility than a single globally optimised bandwidth and is particularly advantageous for high-dimensional applications in which the underlying dependence structure may vary substantially across variable pairs.

The rest of this paper is organised as follows. In Section 2, we introduce the proposed nonparametric estimation method for time-varying correlation coefficients. Section 3 presents simulation studies to evaluate the performance of the proposed method. In Section 4, we apply the method to MEG data for dynamic functional connectivity analysis. Finally, Section 5 concludes the paper and discusses future research directions.

2. Methodology

2.1. Overview

We begin with a high-level description of the estimation procedure. Section 2.2 introduces the local Gaussian working likelihood, whose score equation is a cubic in the correlation ρ with kernel-weighted coefficients A_n and B_n. The assumptions are stated in Section 2.3, and consistency and asymptotic normality are established in Section 2.4. A local-linear version of B_n removes the first-order boundary bias (Definition 1, Theorem 5). For multichannel data, the estimator is applied separately to every pair of channels, each with its own bandwidth, and the pairwise estimates are assembled into a correlation matrix that is corrected to be positive definite. Algorithm 1 outlines the complete procedure as used for the MEG data in Section 4.

Algorithm 1 Pairwise time-varying correlation for multichannel series.

Input: log-envelopes $y_r(t_k)$, $r = 1, \dots, p$, $k = 1, \dots, T$; Gaussian kernel K truncated at $\pm 4\sigma$; bandwidth grid $\mathcal{H}$ and a constant-correlation option; cross-validation gap g; pilot scale σ_p; truncation constant ε.

(1) Standardise: $x_r = (y_r - m_r)/s_r$, with kernel-weighted local mean m_r and standard deviation s_r at scale σ_p.
(2) Select the bandwidth: for each pair (r, q) and each $h \in \mathcal{H}$, compute $A_n(t)$ (local constant) and $B_n^{\mathrm{CL}}(t)$ (local linear) at the validation times with the kernel set to zero on $|t - t_k| \le g$, and evaluate the Gaussian predictive log-likelihood; do the same for the constant-correlation model; set h_{rq} to the best candidate (coarse grid, then local refinement).
(3) Estimate: with the full kernel at h_{rq}, compute $A_n(t)$ and $B_n^{\mathrm{CL}}(t)$ for all t; if the cubic (13) has no root in $(-1, 1)$, truncate B_n^{CL} to $\pm(1 - \varepsilon)A_n/2$; solve for $\hat{\rho}_{rq}(t)$.
(4) Assemble: for each t, form $\hat{R}(t)$; if it is not positive definite, clip its eigenvalues at 10^{-6} and rescale to unit diagonal.
(5) Summarise: mean squared correlation $\overline{r^2}(t)$; Yeo-7 network averages of Fisher-z values (no thresholding).

Output: $\{\hat{\rho}_{rq}(t)\}$, $\{h_{rq}\}$, $\{\hat{R}(t)\}$.

2.2. Model and Estimator

Let $(U_i, \mathbf{X}_i)$, $\mathbf{X}_i = (X_{i1}, X_{i2})^\top$, $i = 1, \dots, n$, be i.i.d. copies of $(U, \mathbf{X})$, where $U \in [-1, 1]$ has density f_U and, conditional on $U = u$, $\mathbf{X} = (X_1, X_2)^\top$ follows a bivariate normal distribution with mean zero and covariance matrix $\Sigma(\rho(u))$:

$$\mathbf{X} = (X_1, X_2)^\top \mid (U = u) \sim N(0,\ \Sigma(\rho(u))), \qquad \Sigma(\rho(u)) = \begin{pmatrix} 1 & \rho(u) \\ \rho(u) & 1 \end{pmatrix},\ \rho(u) \in (-1, 1).$$

The model assumes that, conditionally on $U = u$, both variables have mean zero and unit variance, so that $\Sigma(\rho(u))$ is a correlation matrix. These are localconditions: they hold after global standardisation only if the marginal means and variances do not vary with u. In applications, we therefore centre and scale each series locally before estimation (Section 4.2). The Gaussian form is a working model used to derive the estimating equation; Remark 2 states which results depend on it. The goal is to estimate the correlation function $\rho(u)$ nonparametrically. Define the negative Gaussian log-likelihood contribution (up to some normalising constants)

$$\ell(\mathbf{x}; \rho) = \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{1 - \rho^2} + \log(1 - \rho^2), \qquad \mathbf{x} = (x_1, x_2)^\top \in \mathbb{R}^2,\ \rho \in (-1, 1).$$

The kernel-weighted local criterion at u_0 is

$$Q_n(\rho; u_0) = \frac{1}{n}\sum_{i=1}^{n} K_h(U_i - u_0)\,\ell(\mathbf{X}_i; \rho), \qquad K_h(t) = \frac{1}{h}K(t/h),$$

where $K(\cdot)$ is the kernel function. Define the local weighted MLE as

$$\hat{\rho}(u_0) \in \underset{\rho\in(-1,1)}{\arg\min}\, Q_n(\rho; u_0).$$

One can differentiate Q_n with respect to ρ, set it to zero, and simplify to get the cubic score equation as follows:

$$\rho^3 - B_n(u_0)\rho^2 + \big(A_n(u_0) - 1\big)\rho - B_n(u_0) = 0, \tag{1}$$

where A_n, B_n are the kernel-weighted quantities below. We write $g(\rho; A, B) = \rho^3 - B\rho^2 + (A-1)\rho - B$, so that (1) reads $g(\rho; A_n(u_0), B_n(u_0)) = 0$:

$$A_n(u_0) = \frac{\sum_{i=1}^{n}(X_{i1}^2 + X_{i2}^2)K_h(U_i - u_0)}{\sum_{i=1}^{n} K_h(U_i - u_0)}, \quad B_n(u_0) = \frac{\sum_{i=1}^{n} X_{i1}X_{i2}K_h(U_i - u_0)}{\sum_{i=1}^{n} K_h(U_i - u_0)}.$$

Coincidentally, the Nadaraya–Watson estimator of the correlation coefficient is exactly the ratio of these two quantities:

$$\hat{\rho}^{\mathrm{NW}}(u_0) = \frac{2\sum_{i=1}^{n} X_{i1}X_{i2}K_h(U_i - u_0)}{\sum_{i=1}^{n}(X_{i1}^2 + X_{i2}^2)K_h(U_i - u_0)} = \frac{2B_n(u_0)}{A_n(u_0)}. \tag{2}$$

The corresponding population quantities are $A(u_0) = \mathbb{E}\big[X_1^2 + X_2^2 \mid U = u_0\big]$, $B(u_0) = \mathbb{E}[X_1X_2 \mid U = u_0]$. Because $\Sigma(\rho(u))$ is a correlation matrix for every u, we have

$$A(u) \equiv 2 \quad \text{for all } u, \qquad B(u_0) = \rho(u_0). \tag{3}$$

Here, u_0 can be regarded as a fixed point in the design interval while u is a generic point in the design interval; we write $\rho_0 = \rho(u_0)$. The fact that $A(\cdot)$ is exactly constant drives almost everything in the theory of consistency and limit distribution. Especially, if $A_n(u_0) = 2$, one real root of (1) is equal to $B_n(u_0)$. For an estimator of the correlation coefficient, Kendall et al. [40] pointed out that there is at least one real root of Equation (1) lying in the interval $[-1, 1]$. Let $\hat{\rho}(u_0)$ be the real root of cubic Equation (1) that minimises $Q_n(\rho; u_0)$; we call it the cubic-equation (CE) estimator. Next, we will develop the theory of uniform consistency of the estimator $\hat{\rho}(u_0)$ over compact subsets of the design interval $(-1, 1)$, limit distribution, optimal bandwidth selection, and the convergence rate of the estimator $\hat{\rho}(u_0)$. Moreover, we also discuss the boundary behaviour of the estimator $\hat{\rho}(u_0)$ and propose a method to correct the boundary effects.

2.3. Assumptions

(A1) **Kernel**. K is a symmetric, bounded, and Lipschitz continuous probability density function with compact support $[-1, 1]$, $K \geqq 0$, $\mu_1(K) = \int tK(t)\,\mathrm{d}t = 0$, $\mu_2(K) = \int t^2K(t)\,\mathrm{d}t < \infty$, $R(K) = \int K^2(t)\,\mathrm{d}t < \infty$.

(A2) **Design Density**. The density f_U of U is bounded away from zero and infinity on $[-1, 1]$, i.e., $0 < f_{\min} \leqq f_U(u) \leqq f_{\max} < \infty$. In the interior of $[-1, 1]$, f_U is twice continuously differentiable with bounded second derivative; at ± 1, f_U has finite one-sided derivatives.

(A3) **Smoothness and Range of the Correlation Function**. $\rho \in C^2[-1, 1]$, with $\sup_u |\rho'(u)| < \infty$, $\sup_u |\rho''(u)| < \infty$, and there is $\eta \in (0, 1)$ such that $\sup_{u\in[-1,1]} |\rho(u)| \leqq 1 - \eta$. The last condition keeps the true curve a fixed distance away from the

boundary ± 1 of the parameter space, so that $\Sigma(\rho(u))$ is uniformly positive definite and $\log\{1-\rho(u)^2\} \geqq \log(2\eta-\eta^2) > -\infty$.

(A4) **Kernel Bandwidth**. The kernel bandwidth h satisfies $h \to 0$ and $nh \to \infty$ as $n \to \infty$; for the uniform statements, we require $nh/\log(1/h) \to \infty$.

(A5) **Moment Condition**. For $k = 1, 2$ and all $u \in [-1, 1]$, $\mathbb{E}[X_k \mid U = u] = 0$ and $\mathbb{E}[X_k^2 \mid U = u] = 1$; moreover, $\sup_{u\in[-1,1]} \mathbb{E}\left[(X_1^2 + X_2^2)^4 \mid U = u\right] < \infty$ and the conditional moments of (X_1, X_2) up to order four are continuous in u.

(A6) $\{(U_i, \mathbf{X}_i)\}_{i=1}^n$ are i.i.d. observations.

For moment condition (A5), under the Gaussian working model, the first two conditions hold by construction, because $\Sigma(\rho(u))$ is a correlation matrix, and the moment bound holds because Gaussian moments are polynomials in $\rho(u)$. Pointwise consistency of $\hat{\rho}(u_0)$ requires only the first two conditions; the uniform statements in Theorems 1 and 2 also use the sub-exponential conditional tails of X_k^2 and X_1X_2 implied by the Gaussian model, and the variance formulas in Theorems 3–5 use its fourth moments (Remark 2). Theorems 3–5 assume the conditional Gaussian model of Section 2 in addition to (A1)–(A6). The theory is developed for a random design. For an equally spaced fixed design with independent errors, the same results hold with $f_U \equiv 1/2$ [41]; for serially dependent data, as in Section 4, the variance formulas do not apply. These assumptions are standard in the literature of nonparametric regression and kernel smoothing, ensuring the well-definedness of the estimation problem and the validity of asymptotic results; see Fan et al. [35], Fan and Gijbels [36], Fan et al. [37], Fan and Yao [38], Staniswalis and Lee [39].

2.4. Theory

Theorem 1. *Under the assumptions (A1)–(A6), for any $\delta > 0$, we have*

$$\sup_{u_0\in[-1+\delta,1-\delta]} |A_n(u_0) - 2| \xrightarrow{p} 0, \qquad \sup_{u_0\in[-1+\delta,1-\delta]} |B_n(u_0) - \rho(u_0)| \xrightarrow{p} 0.$$

Proof. Immediate from Lemma 2 in the online Supplementary Materials together with Equation (3) ($A(u) \equiv 2$, $B(u_0) = \rho(u_0)$), and the assumption (A4) which guarantees $\sqrt{\log(1/h)/(nh)} + h^2 \to 0$. □

Theorem 1 upgrades pointwise consistency of the coefficients of the cubic equation to uniform convergence over compact subsets of the design interval. This stronger form of convergence ensures that the root-continuity argument holds uniformly rather than just pointwise, in u_0.

For finite-sample, the cubic equation may have three real roots, or one real root and two complex roots. The following proposition shows that the cubic equation has at least one real root in $(-1, 1)$ for all u_0 in compact subsets of the design interval.

Proposition 1. *For every n and every u_0 with $\sum_i w_i(u_0) > 0$, where $w_i(u_0) = K_h(U_i - u_0)$, the empirical cubic $g_n(\rho; u_0) := g(\rho; A_n(u_0), B_n(u_0)) = 0$ has at least one real root in $[-1, 1]$.*

The proof, given in the online Supplementary Materials, uses only the non-negativity of the weights: averaging $(X_{i1} \pm X_{i2})^2 \geq 0$ gives $|B_n(u_0)| \leq A_n(u_0)/2$, so that $g_n(1; u_0) \geq 0 \geq g_n(-1; u_0)$, and the intermediate value theorem yields a root in $[-1, 1]$. This establishes the empirical, finite-sample analogue of the classical population-level result by Kendall et al. [40]. Crucially, our proof demonstrates that this relationship holds exactly for any non-negative kernel weights, without relying on asymptotic approximations.

Theorem 2. *Uniform Consistency of $\hat{\rho}(u_0)$. Under Assumptions (A1)–(A6), for any $\delta > 0$ and compact set $\mathcal{U}_\delta = [-1+\delta, 1-\delta]$, we have*

$$\sup_{u_0 \in \mathcal{U}_\delta} \left|\hat{\rho}(u_0) - \rho(u_0)\right| \xrightarrow{p} 0, \qquad n \to \infty.$$

Theorem 2 establishes that the estimator $\hat{\rho}(u_0)$ converges uniformly to the true correlation function $\rho(u_0)$ over compact subsets of the design interval. The main idea of proof is to first show that the population cubic equation has a unique simple root ρ_0, then use the uniform convergence of the empirical coefficients to the population coefficients (Theorem 1) and the continuity of roots of polynomials, i.e., Rouché's theorem [42], to conclude that the empirical root $\hat{\rho}(u_0)$ converges uniformly to $\rho(u_0)$; the proof details of Theorem 2 are in the online Supplementary Materials. Having shown that $\hat{\rho}(u_0)$ is uniformly consistent, we next examine its limiting distribution. The derivation is carried out separately for interior points and for boundary points of the design interval, reflecting the fact that the asymptotic behaviour differs between these two regions.

Theorem 3. *Interior Asymptotic Normality. Under Assumptions (A1)–(A6), for any fixed interior point $u_0 \in (-1, 1)$ with $h \to 0$ and $nh \to \infty$, we have*

$$\sqrt{nh}\Big(\hat{\rho}(u_0) - \rho(u_0) - \mathrm{Bias}_\rho(u_0)\Big) \xrightarrow{d} N\left(0,\ \frac{R(K)}{f_U(u_0)} \cdot \frac{\left(1-\rho(u_0)^2\right)^2}{1+\rho(u_0)^2}\right), \tag{4}$$

where the bias term is

$$\mathrm{Bias}_\rho(u_0) = \frac{h^2 \mu_2(K)}{2}\Big[\rho''(u_0) + 2\frac{f_U'(u_0)}{f_U(u_0)}\rho'(u_0)\Big]. \tag{5}$$

Remark 1. *The asymptotic variance in (4) has the form $R(K)/\big(f_U(u_0) I(\rho_0)\big)$, where $I(\rho) = (1+\rho^2)/(1-\rho^2)^2$ is the Fisher information for the correlation parameter ρ in a bivariate normal distribution with known unit marginal variances. This coincides with the general local-likelihood theory for nonparametric varying-coefficient models [35,39], providing an independent verification of the calculations.*

Remark 2 (Role of the Gaussian working model)**.** *The consistency argument does not actually require Gaussianity. Under the moment conditions $\mathbb{E}[X_k \mid U = u] = 0$, $\mathbb{E}[X_k^2 \mid U = u] = 1$ and $\mathbb{E}[X_1 X_2 \mid U = u] = \rho(u)$, the population cubic factorises as $(\rho - \rho_0)(\rho^2 + 1)$ for any conditional distribution, so $\hat{\rho}(u_0)$ is pointwise consistent. The uniform consistency result over compact sets additionally relies on sub-exponential conditional tails of X_k^2 and $X_1 X_2$, unless a stronger bandwidth condition is imposed. Gaussianity only affects the asymptotic variance; if the conditional distribution is not Gaussian, the linearisation used in the proof of Theorem 3 gives the asymptotic variance*

$$\frac{R(K)}{f_U(u_0)} \mathrm{Var}\Big\{X_1 X_2 - \frac{\rho_0}{1+\rho_0^2}(X_1^2 + X_2^2)\ \Big|\ U = u_0\Big\},$$

which equals the variance in (4) under normality. For an elliptical distribution with kurtosis parameter κ ($\mathbb{E}X_1^4 = 3(1+\kappa)$), the variance factor $(1-\rho_0^2)^2/(1+\rho_0^2)$ becomes $(1+\kappa)\frac{(1-\rho_0^2)^2}{1+\rho_0^2} + \kappa\frac{\rho_0^2(1-\rho_0^2)^2}{(1+\rho_0^2)^2}$. The unit-variance condition, by contrast, is essential. If both conditional variances equal $s(u) \neq 1$, the root is biased, to first order in $s-1$, by $-(s-1)\rho_0(1-\rho_0^2)/(1+\rho_0^2)$, whereas the Nadaraya–Watson ratio $2B_n/A_n$ is invariant to a common scale.

Theorem 4. *Boundary Asymptotic Normality. Under Assumptions (A1)–(A6), let $u_0 = -1 + ch$ for fixed $c \in [0, \infty)$ (left boundary), with $h \to 0$ and $nh \to \infty$. Define the truncated kernel moments*

$$\mu_j(K, c) = \int_{-c}^{1} t^j K(t)\, \mathrm{d}t, \qquad R(K, c) = \int_{-c}^{1} K(t)^2\, \mathrm{d}t. \tag{6}$$

Then,

$$\sqrt{nh}\left(\hat{\rho}(u_0) - \rho(u_0) - h\rho'(u_0)\frac{\mu_1(K,c)}{\mu_0(K,c)}\right) \xrightarrow{d} N\left(0,\ \frac{R(K,c)}{\mu_0(K,c)^2 f_U(-1^+)} \cdot \frac{\left(1-\rho(u_0)^2\right)^2}{1+\rho(u_0)^2}\right). \tag{7}$$

The corresponding result holds symmetrically for the right boundary $u_0 = 1 - ch$.

Remark 3. *The qualitative difference between Theorems 3 and 4 is the order of bias: $O(h^2)$ in the interior versus $O(h)$ at the boundary. This is a first-order boundary effect characteristic of local-constant (Nadaraya–Watson) type estimators. Higher-order kernels or local-linear modifications can reduce the boundary bias to $O(h^2)$.*

The asymptotic bias-variance trade-off in Theorems 3 and 4 determines the optimal bandwidth selection and the resulting convergence rates. The following corollary characterises these optimal rates.

Corollary 1. *Optimal Bandwidth and Convergence Rates. Under Assumptions (A1)–(A6), the optimal bandwidth and convergence rates differ between interior and boundary points as follows:*

(i) For interior points $u_0 \in (-1, 1)$, the asymptotic mean squared error $\mathrm{AMSE}(h) = \mathbb{E}[(\hat{\rho}(u_0) - \rho(u_0))^2]$ is minimised at

$$h_{\mathrm{opt}}^{\mathrm{int}} = \left(\frac{C_2(u_0)}{4C_1(u_0)\, n}\right)^{1/5} \asymp n^{-1/5}, \tag{8}$$

where

$$C_1(u_0) = \left(\frac{\mu_2(K)}{2}\right)^2 \left[\rho''(u_0) + 2\frac{f_U'(u_0)}{f_U(u_0)}\rho'(u_0)\right]^2,$$
$$C_2(u_0) = \frac{R(K)}{f_U(u_0)} \cdot \frac{\left(1-\rho(u_0)^2\right)^2}{1+\rho(u_0)^2}.$$

At this bandwidth, $\mathrm{AMSE}(h_{\mathrm{opt}}^{\mathrm{int}}) \asymp n^{-4/5}$ and $\hat{\rho}(u_0) - \rho(u_0) = O_p(n^{-2/5})$.

(ii) For boundary points $u_0 = -1 + ch$ with fixed $c \in [0, \infty)$, the optimal bandwidth is

$$h_{\mathrm{opt}}^{\mathrm{bd}} = \left(\frac{C_4}{2C_3\, n}\right)^{1/3} \asymp n^{-1/3}, \tag{9}$$

where

$$C_3 = \left[\rho'(u_0)\frac{\mu_1(K,c)}{\mu_0(K,c)}\right]^2,$$
$$C_4 = \frac{R(K,c)}{\mu_0(K,c)^2 f_U(-1^+)} \cdot \frac{\left(1-\rho(u_0)^2\right)^2}{1+\rho(u_0)^2}.$$

At this bandwidth, $\mathrm{AMSE}(h_{\mathrm{opt}}^{\mathrm{bd}}) \asymp n^{-2/3}$ and $\hat{\rho}(u_0) - \rho(u_0) = O_p(n^{-1/3})$.

Remark 4. *The boundary convergence rate $n^{-1/3}$ is strictly slower than the interior rate $n^{-2/5}$. This is the standard penalty for local-constant smoothing at design boundaries, where the bias is $O(h)$ rather than $O(h^2)$. This penalty can be eliminated by using local-linear likelihood methods, which achieve $O(h^2)$ bias uniformly across the domain, including at boundaries.*

Next, we address the boundary rate discrepancy. The suboptimal $n^{-1/3}$ convergence rate at boundaries is not inherent to the cubic-equation framework itself, but rather a consequence of using the Nadaraya–Watson estimator for $B_n(u_0)$. We develop a boundary-corrected variant using a local linear method that restores the faster $n^{-2/5}$ rate uniformly across the entire design interval, including boundary points.

The Nadaraya–Watson estimator $B_n(u_0)$ estimates $\rho(u_0) = \mathbb{E}[X_1X_2 \mid U = u_0]$ using a truncated kernel whose first moment $\mu_1(K, c)$ in (6) is nonzero once the window $[u_0 - h, u_0 + h]$ is truncated by the boundary $U \geq -1$. In contrast, $A_n(u_0)$ suffers no bias at any order, because the population function $A(u) \equiv 2$ is constant; thus, $A_n(u_0)$ estimates a zero derivative, and kernel asymmetry is irrelevant. Consequently, the boundary correction need only modify $B_n(u_0)$.

This is precisely the situation addressed by local-linear kernel regression [37,38,41,43,44]. While the Nadaraya–Watson smoother exhibits $O(h)$ design bias at boundaries, the local-linear smoother achieves $O(h^2)$ bias uniformly across the support, including boundaries. This automatic boundary correction is a well-known advantage of local polynomial methods. Because $B(u_0) = \rho(u_0)$ is the regression function of X_1X_2 on U, we can replace $B_n(u_0)$ by its local-linear approximation in the cubic equation without altering the consistency and asymptotic normality arguments previously established.

We define the boundary-corrected estimator by replacing the Nadaraya–Watson component $B_n(u_0)$ with a local-linear fit, while retaining the local-constant estimator for $A_n(u_0)$.

Definition 1 (Boundary-Corrected Estimator)**.** *For $u_0 \in [-1, 1]$, define $(\hat{a}, \hat{b})$ as the minimiser of the local weighted least squares criterion*

$$(\hat{a}, \hat{b}) = \arg\min_{a,b} \sum_{i=1}^{n} K_h(U_i - u_0) \Big[X_{i1}X_{i2} - a - b(U_i - u_0) \Big]^2, \tag{10}$$

and set $B_n^{\mathrm{CL}}(u_0) := \hat{a}$.

Explicitly, let

$$\hat{S}_j(u_0) = \frac{1}{n} \sum_{i=1}^{n} K_h(U_i - u_0)(U_i - u_0)^j,$$
$$\hat{T}_j(u_0) = \frac{1}{n} \sum_{i=1}^{n} K_h(U_i - u_0)(U_i - u_0)^j X_{i1}X_{i2}, \quad j = 0, 1, 2.$$

Then,

$$B_n^{\mathrm{CL}}(u_0) = \frac{\hat{S}_2(u_0)\hat{T}_0(u_0) - \hat{S}_1(u_0)\hat{T}_1(u_0)}{\hat{S}_2(u_0)\hat{S}_0(u_0) - \hat{S}_1(u_0)^2} = \frac{1}{n} \sum_{i=1}^{n} \omega_i(u_0)\, X_{i1}X_{i2}, \tag{11}$$

where

$$\omega_i(u_0) = K_h(U_i - u_0)\, \frac{\hat{S}_2(u_0) - \hat{S}_1(u_0)(U_i - u_0)}{\hat{S}_2(u_0)\hat{S}_0(u_0) - \hat{S}_1(u_0)^2} \tag{12}$$

are the equivalent kernel weights. Unlike the Nadaraya–Watson weights $K_h(U_i - u_0) \geq 0$, these weights can be negative near the boundary, which is the mechanism that eliminates the first-order bias.

Define the boundary-corrected estimator $\hat{\rho}^{\mathrm{CL}}(u_0)$ as the unique real root in $(-1, 1)$ of the modified cubic equation

$$\rho^3 - B_n^{\mathrm{CL}}(u_0)\rho^2 + \big(A_n(u_0) - 1\big)\rho - B_n^{\mathrm{CL}}(u_0) = 0, \tag{13}$$

where $A_n(u_0)$ remains the local-constant estimator defined previously.

The local-linear estimator $B_n^{\mathrm{CL}}(u_0)$ can equivalently be viewed as a local-constant estimator using a boundary-corrected kernel. At the population level, the weights in (11) correspond to the boundary kernel

$$K^*(t;c) := \frac{S_2(c) - S_1(c)\,t}{S_0(c)S_2(c) - S_1(c)^2}\,K(t), \qquad t \in [-c, 1], \tag{14}$$

where, as a shorthand for the truncated moments of (6),

$$S_j(c) := \mu_j(K, c), \quad j = 0, 1, 2, 3, \tag{15}$$

and c parameterises the distance from the left boundary via $u_0 = -1 + ch$. This construction is analogous to the Müller and Gasser–Müller boundary kernels [45,46]. The key property of $K^*(t;c)$ is that it satisfies the moment conditions

$$\int_{-c}^{1} K^*(t;c)\,\mathrm{d}t = 1, \qquad \int_{-c}^{1} t\,K^*(t;c)\,\mathrm{d}t = 0, \tag{16}$$

which hold exactly for every $c \geq 0$. The zero first moment is what eliminates the $O(h)$ bias term, regardless of the degree of truncation.

Remark 5. *As u_0 moves into the interior, $c \to \infty$ and the kernel support becomes effectively symmetric. By Assumption (A1), $S_1(c) \to \mu_1(K) = 0$, so $K^*(t;c) \to K(t)$. The boundary correction vanishes precisely where it is not needed, ensuring continuity between boundary and interior behaviour.*

We now establish that the local-linear modification eliminates the $O(h)$ boundary bias. Because $A_n(u_0)$ and $B_n^{\mathrm{CL}}(u_0)$ are built with different effective kernels near the boundary, namely, the normalised truncated kernel $\bar{K}(t;c) := K(t)/\mu_0(K, c)$ and the boundary kernel $K^*(t;c)$, the asymptotic variance must account for their joint fluctuation. Combining Lemmas 4 and 5 in the online Supplementary Materials with a linearisation of the cubic equation, we obtain the following result.

Theorem 5. *Under the Gaussian model of Section 2 and Assumptions (A1)–(A6), let $u_0 = -1 + ch$ with $c \geq 0$ fixed, $h \to 0$, $nh \to \infty$ and $nh^5 = O(1)$. Then,*

$$\sqrt{nh}\Big(\hat{\rho}^{\mathrm{CL}}(u_0) - \rho_0 - \mathrm{Bias}_\rho^{\mathrm{CL}}(u_0)\Big) \xrightarrow{d} N\big(0,\ V_{\mathrm{CL}}(u_0, c)\big), \tag{17}$$

where

$$\mathrm{Bias}_\rho^{\mathrm{CL}}(u_0) = \frac{h^2}{2}\rho''(u_0)\,\mu_2(K^*, c), \tag{18}$$

$$V_{\mathrm{CL}}(u_0, c) = \frac{1}{f_U(-1^+)}\left[\frac{(1-\rho_0^2)^2}{1+\rho_0^2}\,R(K^*, c) + \frac{4\rho_0^2}{1+\rho_0^2}\,\Delta(c)\right], \tag{19}$$

with $\mu_2(K^*, c) = \int_{-c}^{1} t^2 K^*(t;c)\,\mathrm{d}t$, $R(K^*, c) = \int_{-c}^{1} K^*(t;c)^2\,\mathrm{d}t$ and $\Delta(c) = \int_{-c}^{1}\{K^*(t;c) - \bar{K}(t;c)\}^2\,\mathrm{d}t$.

Sketch of Proof. Write $\alpha_n = A_n(u_0) - 2$ and $\beta_n = B_n^{\mathrm{CL}}(u_0) - \rho_0$. Substituting $A_n = 2 + \alpha_n$ and $B_n^{\mathrm{CL}} = \rho_0 + \beta_n$ into (13) gives, for every ρ, $g(\rho; A_n, B_n^{\mathrm{CL}}) = (\rho - \rho_0)(\rho^2 + 1) + \alpha_n \rho - \beta_n(\rho^2 + 1)$. Hence, the root satisfies the exact identity

$$\hat{\rho}^{\mathrm{CL}}(u_0) - \rho_0 = \beta_n - \frac{\hat{\rho}^{\mathrm{CL}}(u_0)}{1 + \hat{\rho}^{\mathrm{CL}}(u_0)^2}\,\alpha_n = \beta_n - \frac{\rho_0}{1 + \rho_0^2}\,\alpha_n + o_p\big((nh)^{-1/2}\big),$$

where the second step uses consistency, $\alpha_n = O_p((nh)^{-1/2})$ and the fact that $\rho \mapsto \rho/(1+\rho^2)$ is 1-Lipschitz. Since $A(u) \equiv 2$, $\mathbb{E}[\alpha_n \mid U_1, \ldots, U_n] = 0$ exactly, so the bias comes from β_n alone (Lemma 4 in the online Supplementary Materials). The stochastic parts of α_n and β_n are linear smoothers with equivalent kernels $\bar{K}$ and K^*, and they are jointly asymptotically normal with the covariance given in Lemma 5 in the online Supplementary Materials, which involves the cross term $\int_{-c}^{1} \bar{K} K^*$. The variance of $\beta_n - \rho_0 \alpha_n/(1+\rho_0^2)$ then simplifies to (19), using $(1+\rho_0^2)^2 = (1-\rho_0^2)^2 + 4\rho_0^2$. Existence and uniqueness of the real root near ρ_0 follow from the same Rouché-theorem argument as in Theorem 2, which requires only $(A_n, B_n^{\mathrm{CL}}) \xrightarrow{p} (2, \rho_0)$. The full proof is given in the online Supplementary Materials. □

Remark 6. *The second term in* (19) *reflects the fact that* A_n *and* B_n^{CL} *use different sets of weights. Writing* $K^* - \bar{K} = S_1(c)\{S_1(c) - S_0(c)t\}K(t)/[S_0(c)\{S_0(c)S_2(c) - S_1(c)^2\}]$, *the difference simplifies to*

$$\Delta(c) = \frac{S_1(c)^2}{S_0(c)^2\{S_0(c)S_2(c) - S_1(c)^2\}^2} \int_{-c}^{1} K(t)^2\{S_1(c) - S_0(c)t\}^2\,\mathrm{d}t \;\geq 0,$$

which vanishes in the interior, where $c \geq 1$, $S_0 = 1$, $S_1 = 0$ *and* $K^* = \bar{K} = K$, *so that Theorem 5 reduces to Theorem 3;* $\Delta(c)$ *also vanishes when* $\rho_0 = 0$. *In the interior, the fluctuations of* A_n *therefore partly cancel those of* B_n, *reducing the variance from* $(1+\rho_0^2)R(K)/f_U$ *to* $(1-\rho_0^2)^2 R(K)/\{(1+\rho_0^2)f_U\}$, *whereas at the boundary, this cancellation is only partial. The correction is not negligible when* $|\rho_0|$ *is large. For the Epanechnikov kernel at* $c = 0$, *for example,* V_{CL} *exceeds its first term by factors of 1.23, 1.93, 2.84 and 47.9 for* $\rho_0 = 0.3, 0.5, 0.6$ *and* 0.9. *A Monte Carlo check of* (19) *is reported in the online Supplementary Materials.*

Remark 7. *The critical feature of Theorem 5 is that the bias is* $O(h^2)$ *uniformly for all* $c \geq 0$, *from the extreme boundary* ($c = 0$) *through the deep interior* ($c \to \infty$). *This stands in contrast to Theorem 4, where boundary bias was* $O(h)$, *necessitating a slower optimal bandwidth.*

The uniform $O(h^2)$ bias in Theorem 5 restores the faster convergence rate at boundary points. The asymptotic mean squared error is

$$\mathrm{AMSE}_{\mathrm{CL}}(h; u_0) = h^4\, C_1^*(u_0, c) + \frac{1}{nh}\, C_2^*(u_0, c), \tag{20}$$

where

$$C_1^*(u_0, c) = \frac{1}{4}\Big[\rho''(u_0)\mu_2(K^*, c)\Big]^2, \tag{21}$$

$$C_2^*(u_0, c) = V_{\mathrm{CL}}(u_0, c) \quad \text{as defined in (19).} \tag{22}$$

This has the same $h^4 + 1/(nh)$ structure as the interior case in Corollary 1. Minimizing (20) yields

$$h_{\mathrm{opt}}^{\mathrm{CL}}(u_0, c) = \left(\frac{C_2^*(u_0, c)}{4C_1^*(u_0, c)\, n}\right)^{1/5} \asymp n^{-1/5}, \tag{23}$$

with corresponding convergence rate

$$\hat{\rho}^{\mathrm{CL}}(u_0) - \rho(u_0) = O_p\left(n^{-2/5}\right). \tag{24}$$

Corollary 2. *Under Assumptions (A1)–(A6), the boundary-corrected estimator $\hat{\rho}^{\mathrm{CL}}(u_0)$ achieves the optimal rate $n^{-2/5}$ uniformly across the entire design interval $[-1,1]$, including boundary points. The optimal bandwidth is $h_{\mathrm{opt}}^{\mathrm{CL}} \asymp n^{-1/5}$ uniformly, eliminating the boundary penalty present in the local-constant estimator.*

To summarise the improvement, the local-linear correction raises the boundary convergence rate from $n^{-1/3}$ to $n^{-2/5}$, matching the interior rate. Only the finite-sample constants $C_1^*(u_0, c)$ and $C_2^*(u_0, c)$ depend on boundary proximity through $\mu_2(K^*, c)$ and $V_{\mathrm{CL}}(u_0, c)$; the polynomial orders in n and h are uniform across the support.

Finite-sample existence guarantee. The deterministic existence result in Proposition 1 relied on the nonnegativity of the Nadaraya–Watson weights, which ensured that $|B_n(u_0)| \leq A_n(u_0)/2$ because $2|X_{i1}X_{i2}| \leq X_{i1}^2 + X_{i2}^2$. The local-linear weights (12) may be negative, however, so this bound does not necessarily hold in finite samples, and (13) can then fail to have a root in $[-1,1]$. This does not affect the asymptotic theory: by Theorem 5, a unique real root near $\rho(u_0) \in (-1,1)$ exists with probability approaching one. In practice, when (13) has no root in $(-1,1)$, we truncate $B_n^{\mathrm{CL}}(u_0)$ to $\pm(1-\varepsilon)A_n(u_0)/2$ with $\varepsilon = 10^{-3}$, which guarantees a root in $(-1,1)$ and pins the estimate near the truncation bound. This truncation does not affect the asymptotic theory, since the probability that it is needed tends to zero. In the rerun simulations, the truncation was used in 7.64% of evaluations at $n = 100$ and 2.89% at $n = 3200$, decreasing monotonically with n. It was markedly more frequent under the boundary sampling strategy S2 (12.18% at $n = 100$, 5.28% at $n = 3200$) than under the uniform strategy S1 (3.10% at $n = 100$, 0.50% at $n = 3200$); the full breakdown is given in the online Supplementary Materials.

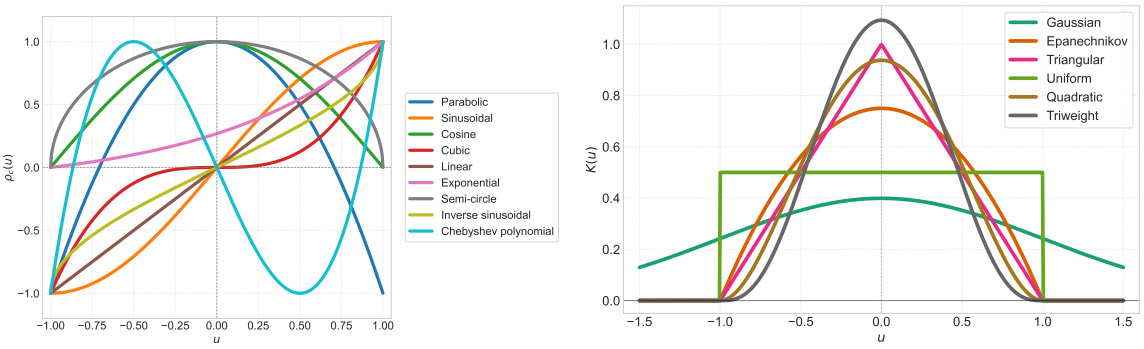
3. Simulation Study

In this section, we conduct a series of simulation studies to compare the performance of the estimation method based on solving the cubic equation with the other methods. The estimators we consider in this section include the Nadaraya–Watson estimator in (2), the cubic-equation-based estimator established in Section 2, the cubic-equation-based estimator with boundary correction using local linear adjustments, and the direct local linear estimator described in the online Supplementary Materials. We call these four estimators NW, CE, CL, and LL, respectively. LL denotes the local-linear likelihood estimator in which $\tanh^{-1}\rho(u)$ is approximated linearly around u_0 and the kernel-weighted Gaussian likelihood is maximised (details in the online Supplementary Materials).

Correlation Function. In this simulation study, we consider using the underlying correlation functions as follows: parabolic function $\rho(u) = 1 - 2u^2$, sinusoidal function $\rho(u) = \sin(\pi u/2)$, Cosine function $\rho(u) = \cos(\pi u/2)$, cubic function $\rho(u) = u^3$, linear function $\rho(u) = u$, exponential function $\rho(u) = (\exp(u) - \exp(-1))/(\exp(1) - \exp(-1))$, upper half semi-circle $\rho(u) = \sqrt{1-u^2}$, normalised inverse sinusoidal function $2/\pi * \arcsin(u)$, and the third Chebyshev polynomial $\rho(u) = 4u^3 - 3u$. For simplicity, we denote them in order as $F1, F2, F3, F4, F5, F6, F7, F8, F9$. These correlation functions represent a wide range of correlation structures, including monotonic, non-monotonic, and oscillatory behaviours. The choice of these functions allows us to evaluate the performance of the four estimators under various scenarios that may arise in practical applications, see Figure 1a. To satisfy Assumption (A3), every function is used in the rescaled form $\rho_c(u) = 0.999999\,\rho(u)$, so that $\sup_u |\rho_c(u)| = 0.999999$ ($\eta = 10^{-6}$); the rescaling keeps $\Sigma(\rho_c(u))$ non-singular at every u and leaves shape, monotonicity, and turning points unchanged, but because η is

small, functions that approach ± 1 probe the estimators close to the limit of (A3); throughout this section, the true correlation function $\rho(\cdot)$ refers to this rescaled version. F7 and F8 have unbounded first derivatives at $u = \pm 1$ and therefore violate the smoothness part of (A3) even after rescaling; we retain them as stress tests of the estimators outside the conditions of the theory.

Kernel Function. We use Gaussian, Epanechnikov, Triangular, Uniform, Quadratic, and Triweight kernel functions in the simulation study, see Figure 1b. For simplicity, we denote them in order as $K1, K2, K3, K4, K5, and K6$, respectively.



(**a**) (**b**)

Figure 1. Correlation and kernel functions: (**a**) The underlying correlation functions (rescaled by 0.999999). (**b**) The kernel functions.

Data generation and parameter settings. The sample size n is set to 100, 200, 400, 800, 1600, and 3200. Given the underlying correlation function and kernel function, we first generate the $U_i, i = 1, 2, \ldots, n$ from the Uniform distribution $\mathcal{U}(-1, 1)$, then for each U_i, randomly generate the X_{i1}, X_{i2} from the bivariate normal distribution with zero mean and covariance matrix $\begin{bmatrix} 1 & \rho(U_i) \\ \rho(U_i) & 1 \end{bmatrix}$. Once we obtain the observations (U_i, X_{i1}, X_{i2}), we apply the four estimators to estimate the correlation function. The optimal bandwidth is selected via grid search over 100 points spaced logarithmically from 10^{-3} to 10^{0}. To evaluate the performance of the four estimators at boundaries, we adopt two sampling strategies: (i) randomly sample $u_j^{\text{test}}, j = 1, 2, \ldots, m$ from the Uniform distribution $\mathcal{U}(-1, 1)$, and (ii) randomly sample 25%, 50%, 25% of $u_j^{\text{test}}, j = 1, 2, \ldots, m$ from the Uniform distributions $\mathcal{U}(-1, -0.9], \mathcal{U}(-0.9, 0.9)$ and $\mathcal{U}[0.9, 1)$, respectively. The first strategy allows us to evaluate the performance of the four estimators across the entire support, while the second strategy focuses on the boundary regions. We use the mean squared error (MSE) to evaluate the performance of the four estimators, which is defined as $\text{MSE} = \frac{1}{m}\sum_{j=1}^{m}(\hat{\rho}(u_j^{\text{test}}) - \rho(u_j^{\text{test}}))^2$, where m is the number of test points. We set $m = 100$ in this simulation. The total number of combinations of the sample size, the correlation functions, kernel functions, and sampling strategies is $6 \times 9 \times 6 \times 2 = 648$. For each combination, we repeat the simulation 200 times to obtain the MSE for each estimator. Furthermore, we also record the computation time for each estimator to evaluate their computational efficiency.

The effect of sample size. Figure 2 shows the log-transformed MSEs of four methods under the two sampling strategies against the different sample size. The "uniform" and "boundary" sampling strategies are denoted as "S1" and "S2", respectively, in the rest of this section. Given the sample size n, for every combination of methods and sampling strategies, we estimate the kernel density of the log-transformed MSEs by pooling them from across 9 correlation functions, 6 kernel functions, and 200 replications. The left dot on the kernel density estimation (KDE) is the corresponding median of the log-transformed MSEs. The KDE is estimated as follows: we adopt a univariate Gaussian KDE;

the bandwidth is selected according to the normal-reference rule $h = 1.06\hat{\sigma}N^{-1/5}$, where $N = 9 \times 6 \times 200 = 10{,}800$ and $\hat{\sigma}$ is the standard deviation of log-transformed MSEs [47]. The resulting KDEs are used solely to provide a smoothed visual representation of the empirical distributions and are not used for formal statistical inference.

Figure 2 highlights several important facts.

- For every combination of estimation method and sampling strategy, the MSE decreases as the sample size increases, in agreement with the theoretical results established in Section 2. For example, under the NW+S1 configuration, the MSE decreases from approximately $10^{-1.75}$ at $n = 100$ to nearly $10^{-3.0}$ at $n = 3200$ in Figure 2.
- When comparing the two sampling strategies, we find that, for every combination of estimation method and sample size, the uniform sampling strategy (S1) consistently yields a more favourable distribution of the MSE than the boundary sampling strategy (S2). This pattern is consistent across all combinations of methods and sample sizes. This result is largely attributable to the differences in the design of sampling strategies. S2 draws 50% of the test points from the narrow boundary regions $(-1, -0.9] \cup [0.9, 1)$, which are more challenging for estimation due to the boundary effects. In contrast, S1 samples across the entire support, leading to a more balanced representation of the correlation function and thus lower MSEs. The scheme of S2 is introduced specifically to assess the effectiveness of the boundary correction in LL and CL methods. Figure 2 provide clear evidence that the boundary correction indeed reduces the estimation error. For instance, when $n = 200$, the vertical distance between the median MSEs of LL under S1 and S2 (represented by two green dots) is substantially smaller than that of NW under S1 and S2 (represented by two blue dots). Similarly, the vertical distance between the median MSEs of CL under S1 and S2 (represented by two orange dots) is smaller than that of CE under S1 and S2 (represented by two red dots). This indicates that the boundary correction effectively mitigates the additional estimation error induced by boundary sampling. A more detailed assessment of the boundary effect is provided later, based on the MSE calculated specifically from test points in the boundary regions.
- From a method comparison perspective, NW and CE do not incorporate boundary correction, while LL and CL do. Consequently, the latter two methods consistently outperform the former two in terms of MSE, regardless of the sampling strategy adopted. A more detailed comparison of the four methods is provided later in this section.

Furthermore, we generate the heatmap of the median logarithmic MSEs across different correlation functions and kernel functions for each combination of estimation method, sampling strategy, and sample size in Figure S3 of the online Supplementary Materials. This heatmap provides a comprehensive overview of the performance of the four methods under various experimental settings, allowing for a more detailed comparison beyond the sample size analysis presented earlier.

The boundary correction. In order to specifically evaluate the performance of the four estimators in the boundary regions, we calculate the MSEs based on the test points sampled from the boundary regions $(-1, -0.9] \cup [0.9, 1)$ under the boundary sampling strategy. The boundary MSEs for each estimator in the boundary regions are defined as follows:

$$\mathrm{MSE}_{\mathrm{B}} = \frac{1}{m_{\mathrm{B}}} \sum_{j \in \mathcal{J}} (\hat{\rho}(u_j^{\mathrm{test}}) - \rho(u_j^{\mathrm{test}}))^2,$$

where $\mathcal{J} := \{j | u_j^{\mathrm{test}} \in (-1, -0.9] \cup [0.9, 1)\}$ and m_B is the cardinality of set $\mathcal{J}$. The boundary MSEs of the four methods against the sample size n are shown in Figure 3. Similar to Figure 2, for each sample size n, the violin boxplot of boundary MSEs is based on

the pooled boundary $\mathrm{MSE}_{\mathrm{B}}$ values across nine correlation functions, six kernel functions, and 200 replications. From Figure 3, we can observe that the boundary MSEs of LL and CL are consistently lower than those of NW and CE across all sample sizes. This indicates that the boundary correction implemented in LL and CL effectively reduces the estimation error in the boundary regions, confirming the theoretical results established in Section 2. LL attains the smallest boundary MSE for $n \leq 1600$, which is consistent with Theorem 5: the boundary variance of CL contains the additional term in (19), whereas LL does not suffer from this mismatch of weights. CL nevertheless improves on CE and NW at every sample size, because at their optimal bandwidths, the $O(h)$ boundary bias of the local-constant estimators dominates.

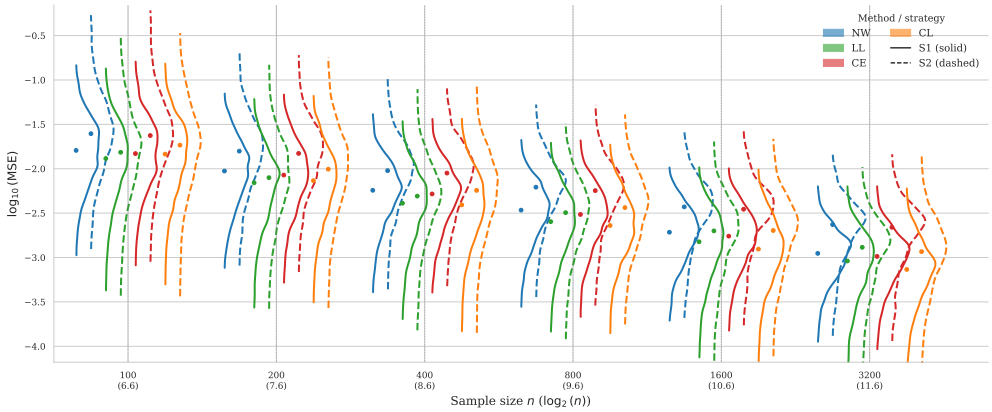


Figure 2. The MSEs of four methods (NW, LL, CE, and CL) against the sample size n. Both the x-axis and y-axis are plotted on logarithmic scales to accommodate the wide range of the data and improve visual clarity. "S1" and "S2" represent the uniform sampling strategy and boundary strategy, respectively. For each sample size n, the specific method and sampling strategy, the kernel density of the log-transformed MSEs is estimated from pooled MSE values across 9 correlation functions, 6 kernel functions, and 200 replications. The dot on the left of each kernel density function represents the corresponding median value of the log-transformed MSE.

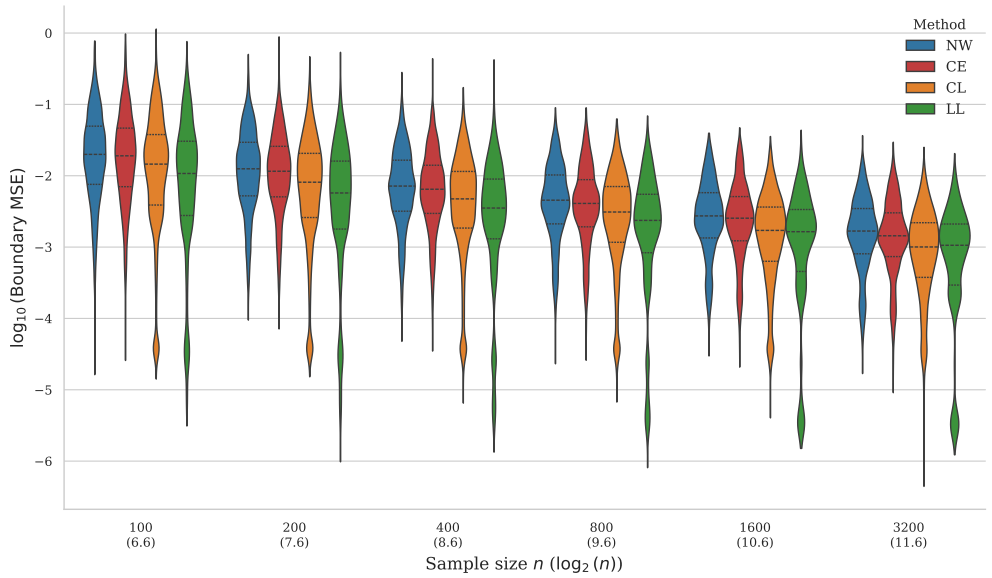


Figure 3. The boundary MSEs of four methods (NW, LL, CE, and CL) against the sample size n. Both the x-axis and y-axis are plotted on logarithmic scales to accommodate the wide range of the data and improve visual clarity. For each sample size n, the specific method and sampling strategy, the violin boxplot is based on the pooled $\mathrm{MSE}_{\mathrm{B}}$ values across 9 correlation functions, 6 kernel functions, 200 replications.

The time consumption. Figure 4 presents heatmaps of the time consumption across different methods, sample sizes, sampling strategies, correlation functions, and kernel functions. The horizontal axis consists of the 54 combinations of correlation functions and kernel functions, while the vertical axis represents the 48 combinations of estimation methods, sample sizes, and sampling strategies. Each cell reports the median time consumption over the 200 simulation replications for a given combination of correlation function, kernel function, method, sample size, and sampling strategy. Darker shading indicates a larger median time consumption. Among the four methods, LL consistently incurs the highest computational cost. For all methods, time consumption increases with sample size, reflecting the expected decline in computational efficiency as the sample size increases. In contrast, the choice of correlation function and kernel function has relatively little effect on the time consumption.

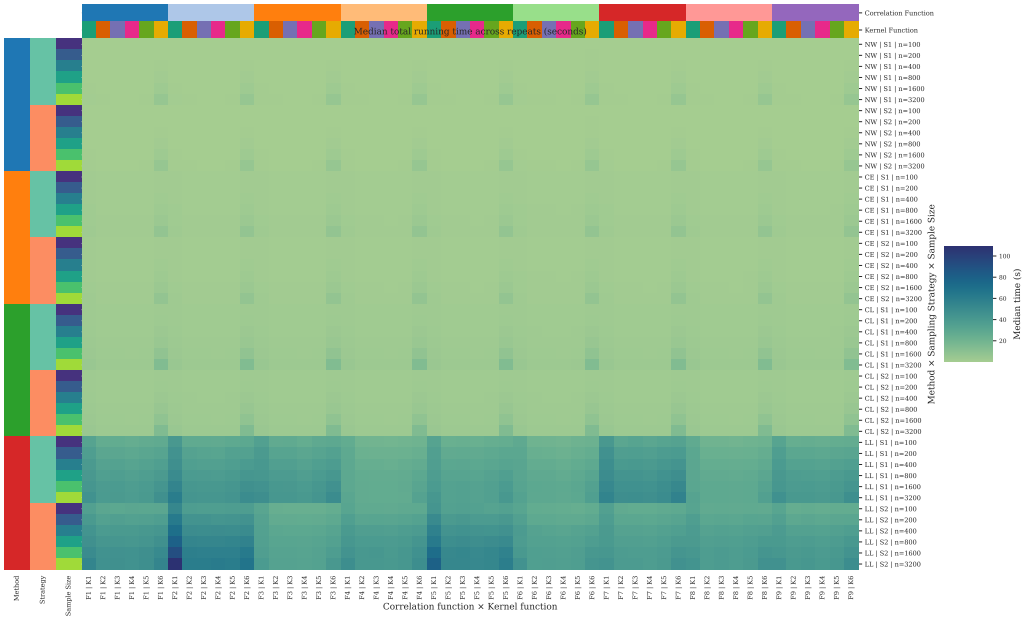
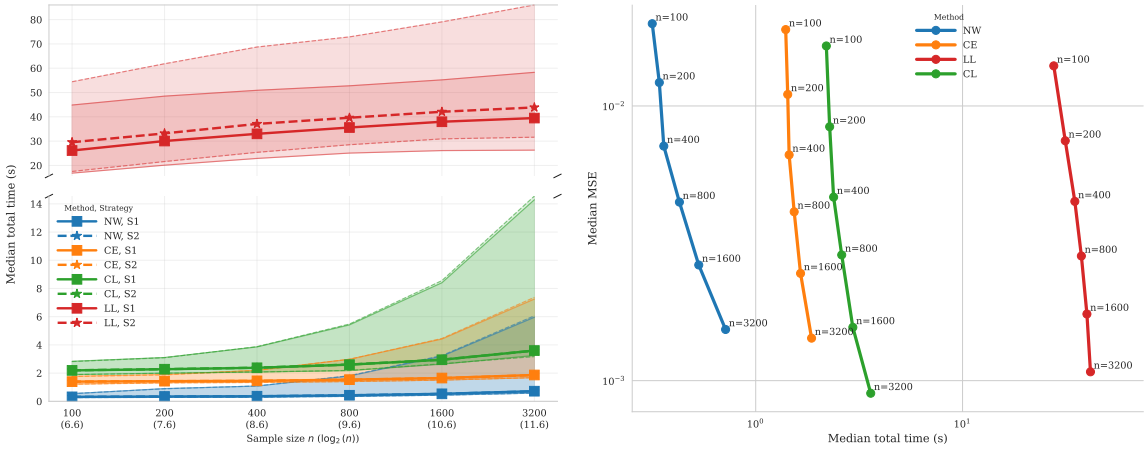


Figure 4. Heatmap of the median time consumption (in seconds) for each combination of estimation method, sample size, sampling strategy, correlation function, and kernel function, with darker shading indicating larger time consumption. The top and left axes indicate the specific combinations of correlation functions, kernel functions, methods, sample sizes, and sampling strategies which are represented in the heatmap.

To further examine how time consumption varies with sample size, we take the median time consumption across the nine correlation functions, six kernel functions, and 200 simulation replications, and plot time consumption against sample size for the eight combinations of method and sampling strategy. The resulting plots are shown in Figure 5a. The solid line with square markers represents the median time consumption under sampling strategy S1 as a function of sample size, whereas the dashed line with star markers represents the corresponding median under sampling strategy S2. The 95% empirical confidence intervals are also shown using different colours, with their boundaries indicated by solid and dashed lines for S1 and S2, respectively. As expected, the time consumption increases with sample size for all methods and under both sampling strategies. Moreover, S2 consistently requires more time consumption than S1 across all sample sizes and methods. This difference is attributable to the greater number of test points located near the boundaries under S2, which increases the computational burden associated with estimation in the boundary regions.

To provide a more direct illustration of the trade-off between estimation accuracy and computational efficiency, we plot the accuracy–speed trade-off for the four methods, as shown in Figure 5b. Each point represents the median time consumption plotted against the median MSE. From this perspective, a particularly noteworthy result occurs when $n = 3200$: CL achieves a lower MSE than LL while requiring substantially less time consumption. Overall, LL provides the highest estimation accuracy but at the greatest computational cost. In contrast, CL offers a more favourable balance between accuracy and computational efficiency and, in some settings, even outperforms LL in both respects. NW and CE are computationally more efficient, but this advantage comes at the cost of lower estimation accuracy.



(**a**) (**b**)

Figure 5. Computational cost of the four estimation methods and their trade-off with estimation accuracy. (**a**) Median time consumption as a function of sample size for the four methods under the sampling strategies S1 and S2, showing that time consumption increases with sample size and is larger under S2; (**b**) the accuracy–speed trade-off, where each point plots the median time consumption against the median MSE, illustrating that LL achieves the highest accuracy at the greatest computational cost while CL offers the most favourable balance.

How to choose the method? In practice, estimation accuracy alone is insufficient; the associated computational burden imposes a critical constraint, particularly in online or high-throughput settings. For instance, in typical MEG source reconstruction, the sampling rate is often 1000 Hz, yielding 1000 time points per second, amounting to 600,000 time points over a 10 min recording. Even when downsampled to 100 Hz, this still results in 60,000 observation points, and the combinatorial pairing across cortical regions further inflates the total computational load, making algorithmic efficiency a primary practical concern.

To balance statistical precision against computational cost, we summarise each method by rank-based indices computed over the 648 configurations $\nu = (n, F, K, S) \in \mathcal{C}$. For each configuration ν, we take the median MSE and the median runtime of each method $m \in \{\mathrm{NW}, \mathrm{LL}, \mathrm{CE}, \mathrm{CL}\}$ over its 200 replications, and rank the four methods within configuration ν on each criterion, with rank 1 assigned to the best method. Averaging these configuration-wise ranks over all configurations gives the mean MSE rank $\bar{R}_m^{\mathrm{MSE}}$ and the mean runtime rank $\bar{R}_m^{\mathrm{time}}$, and averaging the average of the two ranks within each configuration gives the dual rank $\bar{R}_m^{\mathrm{dual}}$, which weights accuracy and speed equally. For a weight $w \in [0, 1]$ on accuracy, the effective score $\mathrm{eff}_m(w)$ averages, over configurations, w times the min–max normalised $\log_{10}$ median MSE plus $1 - w$ times the min–max normalised $\log_{10}$ median runtime, where the normalisation is across the four methods within each configuration; the four methods are then ranked by this score for each $w \in \{0.1, \ldots, 0.9\}$

to obtain the effective rank $R_m^{\text{eff}}(w)$. For all three indices, a smaller value indicates a better method, and their exact definitions are given in the online Supplementary Materials. The resulting rankings in Table 2 provide a comprehensive picture of how the methods compare under varying preferences for accuracy versus computational efficiency. We can observe that CL exhibits the most robust overall performance: it attains the best effective rank $R_m^{\text{eff}}(w)$ for all weights $w = 0.4$ to 0.8 and ranks second at $w = 0.3$ and $w = 0.9$, and it also achieves the smallest dual rank $\bar{R}_m^{\text{dual}} = 2.402$, indicating a strong balance between estimation accuracy and computational cost. LL delivers the highest estimation accuracy, as reflected by the smallest mean MSE rank $\bar{R}_m^{\text{MSE}} = 1.465$, but its dual rank $\bar{R}_m^{\text{dual}} = 2.732$ is the least favourable among the four methods owing to its substantially higher computational cost; consequently, LL attains the top effective rank only when accuracy is assigned the largest weight ($w = 0.9$). NW is computationally the most efficient ($w = 0.1, 0.2, 0.3$), yielding the second-best dual rank $\bar{R}_m^{\text{dual}} = 2.430$, yet it suffers from the worst estimation accuracy ($\bar{R}_m^{\text{MSE}} = 3.860$), and its effective rank deteriorates as w increases, so that it is only attractive when computational speed is prioritised over precision. CE performs moderately on both dimensions, occupying an intermediate position in the effective ranking across all weights.

Table 2. The ranking of the four methods in terms of MSE, dual criterion, and effective score rank under different weights w.

Method	$\bar{R}_m^{\text{MSE}}$	$\bar{R}_m^{\text{dual}}$	R_m^{eff} with Different Weight								
			0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
CL	1.804	2.402	3	3	2	1	1	1	1	1	2
LL	1.465	2.732	4	4	4	4	4	2	2	2	1
CE	2.872	2.436	2	2	3	3	3	3	3	3	3
NW	3.860	2.430	1	1	1	2	2	4	4	4	4

Wilcoxon Signed-rank Analysis: CL vs. LL. In this section, we perform one-sided paired Wilcoxon signed-rank tests to compare the performance of the CL and LL methods across all simulation cells. The tests are conducted by pairing Monte Carlo replicates via the repetition index, with the alternative hypothesis that CL has smaller mean squared error (MSE) than LL. To be concrete, let $\text{MSE}_{r,\nu}^{\text{CL}}$ and $\text{MSE}_{r,\nu}^{\text{LL}}$ denote the MSEs of the CL and LL methods, respectively, for the r-th repetition in the ν-th simulation cell. For each cell $\nu \in \mathcal{C}$, we define the paired differences $d_{r,\nu} := \text{MSE}_{r,\nu}^{\text{CL}} - \text{MSE}_{r,\nu}^{\text{LL}}$. The one-sided paired Wilcoxon signed-rank test is then applied to test whether CL is better than LL in terms of MSE:

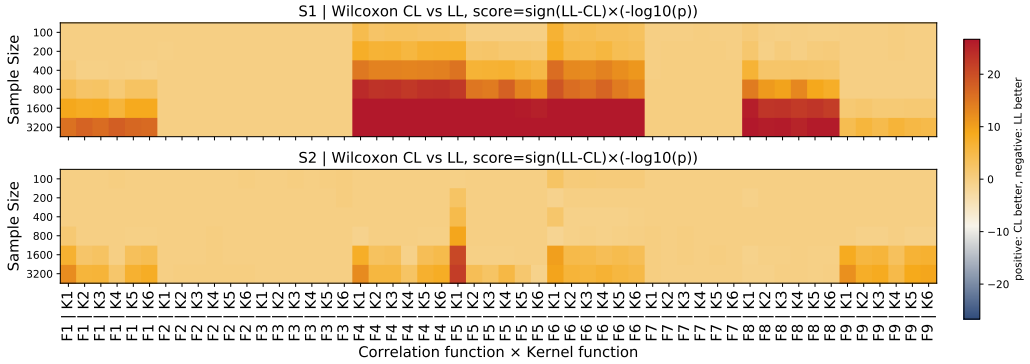
$$\mathbb{H}_0 : \text{median}(d_{r,\nu}) \geq 0 \quad \text{vs} \quad \mathbb{H}_1 : \text{median}(d_{r,\nu}) < 0.$$

For each simulation cell ν, we report the one-sided p-value (p_ν), and mean difference $\bar{d}_\nu := \frac{1}{200}\sum_{r=1}^{200} d_{r,\nu}$. A cell is marked as "CL better" if $p_\nu < 0.05$ and $\bar{d}_\nu < 0$. To summarise the overall performance across all simulation cells, we unify significance and direction by the transformed score $s_\nu := -\log_{10}(p_\nu) \cdot \text{sign}(\bar{d}_\nu)$. The resulting scores can then be visualised using heatmaps in Figure 6a. Furthermore, we compute the significance win rate, defined as the proportion of cells where CL is significantly better than LL, and demonstrate it in Figure 6b.

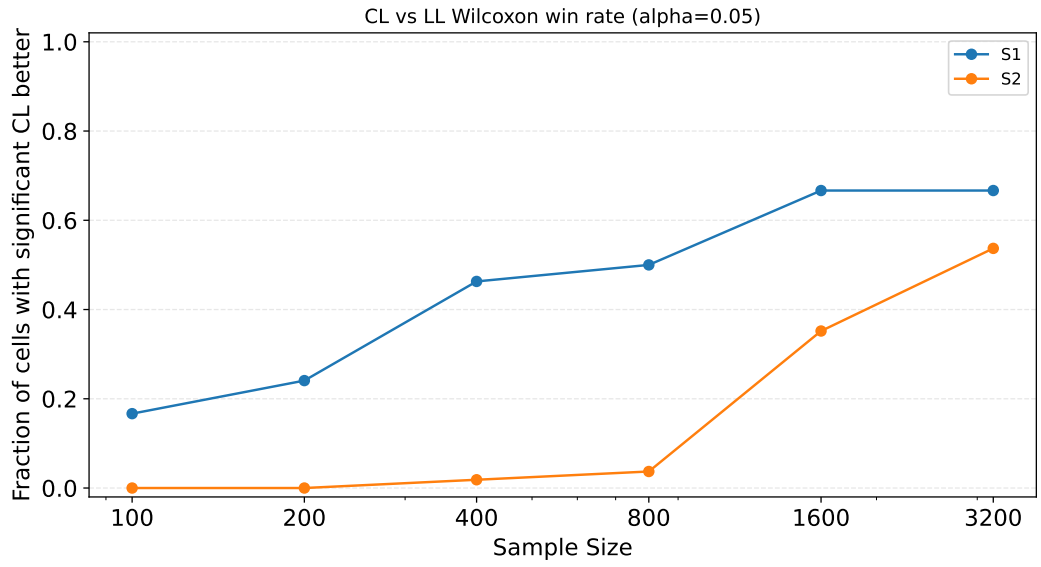
Across 648 simulation cells, 213 show nominal significance ($p_\nu < 0.05$), of which 212 favoured CL over LL and 1 favoured LL over CL. The sole case favouring LL over CL is $(F6, K1, S2, n = 800)$, with $p = 0.0385$ and $\bar{d} = 7.136 \times 10^{-5}$. Furthermore, we apply BH-FDR at $q = 0.05$ over 648 cells and yield 203 discoveries, all of which favour CL over LL. Results are robust to BH-FDR correction. The proportion of CL-favouring significant

cells increases with sample size (from 12.96% at $n = 100$ to 60.19% at $n = 3200$), and is markedly higher under S1 than S2 (47.53% vs. 17.90%).

Robustness to non-Gaussian data. To examine Remark 2 in finite samples, we repeated the simulation for F1, F2, and F5 with the Epanechnikov kernel, $n \in \{400, 1600\}$ and both test-point strategies. The data were generated as $(X_1, X_2) = \Sigma^{1/2}(\rho_c(U))\, W$, where W followed either a multivariate t_5 distribution scaled to unit variances, which has heavy tails and does not satisfy the eighth moment condition in (A5), or independent standardised Gamma(2) components (skewness $\sqrt{2}$, excess kurtosis 3); in both cases, $\Sigma^{1/2}$ is the symmetric square root, so the conditional correlation remains $\rho_c(U)$. Compared with the Gaussian design, the median MSE of every estimator increased under both departures. For the Gamma(2) design, the median ratios were 1.00–1.08 for NW, 1.09–1.18 for CE and 1.21–1.31 for CL, and for the t_5 design, they were 1.57–1.75 for NW, 1.49–1.91 for CE and 1.72–2.14 for CL; the heavy-tailed design thus cost all three estimators roughly a factor of two. The ranking of the estimators changed little: CL kept the smallest median MSE in 20 of the 24 combinations, and the only reorderings occurred under the skewed design, where NW and CE exchanged places (NW ranking first in one cell), whereas under the t_5 design, the ranking was unchanged in every cell.



(**a**)



(**b**)

Figure 6. Wilcoxon signed-rank test results for CL vs. LL across simulation cells. (**a**) The heatmap of Wilcoxon signed-rank test transformed scores for CL vs. LL across simulation cells. (**b**) The significance win rate of CL over LL across simulation cells.

4. Real Data Analysis

Magnetoencephalography (MEG) provides a non-invasive measure of the magnetic fields generated by neuronal currents in the cerebral cortex, with millisecond-scale temporal resolution, making it particularly well suited for investigating the dynamic organisation of functional brain network [48,49]. In resting-state MEG studies, power envelope correlation (PEC) has become a widely used approach for quantifying the synchrony of low-frequency (<1 Hz) fluctuations in the amplitudes of neural oscillations across cortical regions [4–6,50]. A key finding from this line of research is that the power envelopes of band-limited neural oscillations exhibit significant and reproducible correlations between distinct brain regions. These inter-regional correlations are thought to reflect the spatial organisation of resting-state functional networks, including the default mode and visual networks [4,6,50].

Traditional PEC estimates functional connectivity by computing the Pearson correlation coefficient between power-envelope time series over the entire recording period, thereby yielding a static, time-invariant measure of connectivity:

$$\rho_{rq} = \mathrm{corr}\big(e_r(t), e_q(t)\big) = \frac{\sum_t \big(e_r(t) - \bar{e}_r\big)\big(e_q(t) - \bar{e}_q\big)}{\sqrt{\sum_t \big(e_r(t) - \bar{e}_r\big)^2 \sum_t \big(e_q(t) - \bar{e}_q\big)^2}},$$

where $e_r(t)$ denotes the power envelope of the r-th ROI at time t. However, functional connectivity in the resting brain is inherently dynamic, with the strength of inter-regional coupling varying over time [4,49]. The nonparametric correlation estimator CL proposed in this study extends the conventional static correlation framework by incorporating time t into the observation model, thereby representing the correlation as a smooth function $\rho(t)$. This formulation allows the strength of functional connectivity to vary continuously over time, without imposing a specific parametric form on its temporal evolution. In the MEG application, $\hat{\rho}(t)$ provides a nonparametric estimate of time-varying functional connectivity between pairs of ROIs.

In this case study, we use MEG data from the Cam-CAN dataset [51] to demonstrate the practical utility of the proposed nonparametric correlation estimation method. The Cam-CAN dataset contains resting-state MEG recordings from healthy adults across a broad range of ages. The recordings were acquired at a sampling frequency of 1000 Hz and lasted 8 min 40 s, providing a rich temporal profile of ongoing neural activity. To provide a clear illustration of the proposed method, we focused on a single participant (sub-CC110033).

Our analysis has two distinct aims. The first is to show that the estimator can be applied to whole-brain MEG data and yields interpretable network summaries (Sections 4.1, 4.2 and 4.4). The second is to examine whether its statistical properties (time-varying estimation, pair-specific bandwidths, and boundary correction) make a measurable difference compared with conventional approaches. To this end, we benchmark the estimator against static and sliding-window correlation (Section 4.3), run a surrogate test of the estimated fluctuations (online Supplementary Materials), and report each network result in Section 4.4 together with its sliding-window and static counterparts. Since only one participant is analysed, the network results should be read as a proof of concept rather than as population-level findings.

To keep the computational burden manageable, we restricted the analysis to a moderate number of brain regions. Instead of the commonly used 68-region Desikan–Killiany (DK) atlas [52], we adopted the 200-region Schaefer Local-Global Parcellation rather than its 1000-region counterpart [13]. Because functional network organisation can vary substantially across frequency bands [16,53–55], we followed the conventional classification of neural oscillations and extracted power envelopes separately for six frequency bands: delta (1–4 Hz),

theta (4–8 Hz), alpha (8–12 Hz), beta (12–40 Hz), slow gamma (40–80 Hz), and fast gamma (80–120 Hz) [56–58]. The extracted power envelopes were subsequently log-transformed and used as input to the proposed nonparametric correlation estimation method.

We first describe the procedure for extracting power envelopes from the raw MEG recordings and then demonstrate how the proposed method can be applied to these real-world data.

4.1. Power Envelope Extraction

We first prepare the raw data required for the analysis. Specifically, we utilize the resting-state MEG recording, which has been preprocessed using MaxFilter to implement the Signal Space Separation (SSS), alongside the corresponding empty-room recording for environmental noise estimation. The spatial alignment between the MEG sensor space and the structural MRI space is achieved using a pre-calculated MEG–MRI coregistration transformation matrix. Because the experimental recordings have already undergone the SSS procedure [59,60], Maxwell filtering is not repeated in the subsequent analysis pipeline.

Data preprocessing. The preprocessing pipeline consists of the following steps. (1) Channel selection and bad-channel handling. For both the experimental and empty-room recordings, we retained only MEG channels and excluded channels marked as "bads". We then restricted both datasets to their common set of channels to ensure consistency in the subsequent noise covariance estimation. The empty-room data were reordered to match the channel ordering of the experimental recording. (2) Band-pass filtering. To remove very slow drifts while retaining the lowest frequency range of interest (1–4 Hz), both recordings were high-pass filtered at 0.5 Hz using a zero-phase FIR filter designed with 'firwin' [8,61]. This cutoff preserves the lowest frequencies of the delta band without introducing substantial attenuation. (3) Power line noise removal. Notch filters were applied at 50, 100, and 150 Hz to suppress power-line interference and its harmonics.

Source Reconstruction and Atlas Mapping. It consists of four main steps:

1. **Anatomical processing and source space.** The T1-weighted structural MRI was processed with FreeSurfer [12] to reconstruct the cortical surfaces. An "oct6" source-space discretisation was then used to generate a cortical source grid containing approximately 8196 vertices, with 4098 vertices per hemisphere.
2. **Boundary element model and forward solution.** The geometry of the volume conductor was encapsulated by a single-layer Boundary Element Method (BEM) model. The inner skull surface boundary was automatically segmented from structural T1-weighted MRI data utilising the watershed algorithm [62] via FreeSurfer. Inside this bounding surface, the brain tissue compartment was treated as a homogeneous and isotropic volume conductor with a standard consensus conductivity assigned at 0.3 S/m [63,64]. To establish the forward solution, the MEG sensor locations were aligned with the structural MRI coordinate space via an MEG-MRI coregistration transform based on anatomical fiducials and digitised headshape points. Subsequently, the source-to-sensor gain matrix (lead-field matrix) was computed using the single-layer BEM framework [7]. To prevent mathematical singularity and numerical instability in closer sensors, a minimum source-to-sensor distance boundary of 5 mm was enforced during the forward field calculation [65]. Both the forward solution and the source space were cached to avoid unnecessary recomputation.
3. **Noise covariance and inverse operator.** The noise covariance matrix, $\hat{\Sigma}_n$, was estimated from the empty-room recording. A minimum-norm inverse operator was then constructed using a loose orientation constraint of 0.2, depth weighting of 0.8, and an assumed signal-to-noise ratio (SNR) of 3 [17,66]. The inverse solution was computed using dynamic statistical parametric mapping (dSPM) [67–69]. The regular-

isation parameter was determined from the assumed SNR as $\lambda^2 = \frac{1}{\mathrm{SNR}^2} = \frac{1}{9}$. After whitening the sensor-space data, dSPM obtains cortical current estimates through the minimum-norm solution and subsequently normalises them by their estimated noise sensitivity. The resulting estimates therefore have an approximately unit-variance t-statistic interpretation [68,69].

4. **Broadband Source Reconstruction and ROI Time Series.** The inverse operator was applied to the broadband MEG data following 0.5 Hz high-pass filtering. To constrain the inverse solution to the physiologically relevant anatomy, a cortical surface normal constraint was employed to project the estimated current vectors exclusively onto the local surface normals, reflecting the anatomical alignment of macroscopically oriented cortical pyramidal cells [66,70]. This produced continuous cortical source estimates $S(\mathbf{r}, t)$. The source estimates were then parcellated according to the Schaefer-200 cortical atlas [13]. To extract a single representative time course for each region of interest (ROI), we applied an orientation-aligned sign-flipping averaging procedure [1]. Specifically, the ROI-level time series was defined as $x_r(t) = \frac{1}{n_r} \sum_{j \in \mathrm{ROI}_r} \sigma_j \hat{S}(\mathbf{r}_j, t)$, where n_r denotes the number of vertices in ROI r, and $\sigma_j \in \{-1, +1\}$ is a sign-flip factor determined by the orientation of the surface normal at vertex j. This procedure addresses the complex folding geometry of the cortex; it mathematically flips the sign of time series from vertices whose local surface normals oppose the dominant orientation of the ROI, thereby preventing phase cancellation and ensuring an accurate, robust representation of regional population activity [1,71].

Source Leakage Correction. Because the forward-model point-spread functions of neighbouring sources overlap on the cortex, the time course reconstructed for a given parcel contains contributions from the surrounding tissue, an effect known as spatial leakage. Leakage inflates and distorts inter-regional correlation estimates, and a standard remedy in MEG connectomics is to whiten the parcel-level signals by symmetric orthogonalisation [24]. Let $\mathbf{X}$ denote the $R \times T$ matrix of band-limited ROI time courses, with the $R = 200$ rows corresponding to the Schaefer parcels and each row centred over time. The symmetric orthogonalisation applies the inverse matrix square root of the row covariance,

$$\mathbf{X}_{\mathrm{orth}} = \left(\mathbf{X}\mathbf{X}^\top\right)^{-\frac{1}{2}} \mathbf{X}, \tag{25}$$

which, for full-rank $\mathbf{X}$, decorrelates the ROI signals while rotating them as little as possible, in a least-squares sense, from the original time courses.

The full-rank construction does not apply to the present data. The 200 parcels are reconstructed from MEG measurements of limited effective rank, and the row-standardised band-limited ROI matrices show a pronounced singular-value gap: in the delta band, for instance, the relative spectrum falls from 3.7×10^{-2} to 2.0×10^{-7} between the 70th and 71st singular values. The row covariance $\mathbf{X}\mathbf{X}^\top$ is therefore numerically singular, and the inverse square root in (25) is unstable. We instead used a rank-aware, rank-truncated Löwdin orthogonalisation. Truncating the eigenvalue spectrum to the band-specific ranks selected below redistributes the variance associated with the discarded dimensions among the retained components, thereby inflating the retained eigenvalues and potentially altering the magnitude of the leakage correction [24]. Writing the singular value decomposition of the row-centred, row-standardised ROI matrix as $\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^\top$ with $s_1 \geq s_2 \geq \cdots \geq s_R$, and denoting by $\mathbf{U}_r$, $\mathbf{S}_r$, and $\mathbf{V}_r$ the leading r singular vectors and values, the leakage-corrected signals are

$$\mathbf{X}_{\mathrm{orth}} = \mathbf{U}_r \mathbf{V}_r^\top = \left(\mathbf{U}_r \mathbf{S}_r^2 \mathbf{U}_r^\top\right)^{-\frac{1}{2}} \mathbf{X}, \tag{26}$$

the Moore–Penrose analogue of (25) restricted to the retained subspace. All 200 rows are kept; the truncation confines the corrected signals to an r-dimensional subspace in

which they are mutually uncorrelated, so that the row Gram matrix $\mathbf{X}_{\text{orth}}\mathbf{X}_{\text{orth}}^{\top}$ is a rank-r projector rather than the identity. The correction is therefore best described as a rank-deficient modification of the standard symmetric orthogonalization, not as an ordinary 200-dimensional orthogonalization. Because the truncation does not preserve row norms, each corrected row was rescaled to restore the root-mean-square amplitude of the original ROI signal, so that the envelopes computed below retain the between-region amplitude structure of the source estimates.

The effective rank was selected separately for each frequency band from the archived rank diagnostics and took the values 154 (delta), 165 (theta), 162 (alpha), 166 (beta), and 170 (slow and fast gamma). The full singular spectrum, the selected rank, the explained variance, and the gap ratio were saved for every band as diagnostics alongside the envelope data.

Power envelopes were derived from the leakage-corrected, band-limited signals by the Hilbert transform: for a band-limited signal $x(t)$, the instantaneous power is $|x(t) + i\,\mathcal{H}[x(t)]|^2$ [9]. To retain the slow amplitude modulations that carry resting-state network structure [4], the power was smoothed with a Gaussian kernel of standard deviation 132.5 ms. The smoother has gain 0.71 at 1 Hz, 0.25 at 2 Hz and 1.7×10^{-4} at 5 Hz, so that fluctuations faster than about 1 Hz are progressively attenuated, and those above 3 Hz are essentially removed. Because the kernel is non-negative, the smoothed power cannot become negative, unlike a finite-impulse-response low-pass filter, which can undershoot zero where the power is small. The smoothed power was decimated to 10 Hz, which shortens each series from 574,000 to 5740 samples, and log-transformed. Because the envelopes were derived from dSPM estimates, they represent noise-normalised rather than absolute source power. For the delta band (1–4 Hz), the envelope bandwidth is comparable to the carrier frequency, so delta envelopes should be interpreted with caution. The log-envelopes remain moderately non-Gaussian (skewness from -0.42 to 4.02, excess kurtosis from -0.14 to 28.16 across regions, with per-band medians of 0.29–5.28 for the excess kurtosis; Q–Q plots in the online Supplementary Materials).

4.2. Pairwise Correlation Estimation

Standardisation and estimand. Let $y_r(t) = \log e_r(t)$ be the log-envelope of region r at the sampling times $t_k = k/10\,\text{s}$. Because the estimator assumes zero conditional mean and unit conditional variance (Section 2), each series was centred and scaled locally, $x_r(t) = \{y_r(t) - m_r(t)\}/s_r(t)$, where m_r and s_r are the Gaussian-kernel weighted mean and standard deviation with pilot scale $\sigma_p = 60\,\text{s}$, normalised at the edges. This removes slow drifts (periods longer than about seven minutes) and slow changes in variance, but leaves the local means at the much shorter time scales of the correlation bandwidths intact. The methods compared below therefore target related but distinct quantities. For a pair of regions (r, q), write $A(t) = \mathbb{E}[x_r^2 + x_q^2 \mid t]$, $B(t) = \mathbb{E}[x_r x_q \mid t]$ and $\mu_r(t) = \mathbb{E}[x_r \mid t]$ at the scale of the bandwidth. Then, CL and CE estimate the root of the population cubic with coefficients $A(t)$ and $B(t)$, which equals $B(t)$ when $A(t) = 2$; NW estimates $2B(t)/A(t)$; and the kernel local Pearson correlation (KLP) and the sliding-window correlation estimate $\{B(t) - \mu_r(t)\mu_q(t)\}/\{v_r(t)v_q(t)\}^{1/2}$, where v_r and v_q are the local variances. All coincide with $\rho(t)$ when the local means vanish and the local variances equal one; otherwise, CL measures the local second-moment association of the standardised series. Sensitivity analyses with global standardisation, with $\sigma_p = 30\,\text{s}$ and with a rank-based inverse-normal transformation are summarised in the online Supplementary Materials.

Pairwise estimation and bandwidth selection. For each of the 19,900 region pairs, we computed the boundary-corrected estimator of Definition 1: $A_n(t)$ is the local-constant average of $x_r^2 + x_q^2$, $B_n^{\text{CL}}(t)$ the local-linear estimate of $x_r x_q$ (Equation (11)), and $\hat{\rho}(t)$ the

root of (13). All kernel sums were computed by fast Fourier transforms with zero padding, so the estimates at the start and end of the recording use the truncated kernel, as in the theory. The bandwidth h_{rq} of each pair was chosen by blocked cross-validation [72]: at every fifth sample (1148 validation times, including the edges), the Gaussian predictive log-likelihood was evaluated with the kernel set to zero within ± 1.0 s of the validation time, so that the prediction does not use the serially dependent neighbouring samples (the autocorrelation of the envelope products falls below 0.1 within 0.5–0.8 s). The candidate set consisted of twelve logarithmically spaced bandwidths from 1.5×1.0 to 60 s, followed by a local refinement, together with a constant-correlation model. The constant-correlation option is a model-selection device that compares a constant ρ with a smooth $\rho(\cdot)$ by the same criterion; it is not a bandwidth covered by Theorems 3–5. Bandwidths smaller than about 1.5×1.0 cannot be assessed by this criterion.

Correlation matrices and network summaries. For each band and time point, the 19,900 estimates were assembled into a 200×200 matrix. Where this matrix was not positive definite (97–100% of time points, 99.5% overall), its eigenvalues were clipped at 10^{-6} and the matrix was rescaled to unit diagonal. This is a fast approximation of the nearest-correlation-matrix projection of Higham [73], not an implementation of it; it changed the off-diagonal entries by 0.0019–0.0114 on average across bands (overall mean 0.0056). As a summary of overall coupling, we report the mean squared off-diagonal correlation $\overline{r^2}(t)$; for a correlation matrix, the eigenvalue variance reported in the original submission equals $(p-1)\,\overline{r^2}(t)$. Yeo-7 network-level values were obtained by averaging Fisher-z-transformed correlations over all region pairs of each network pair, without thresholding.

4.3. Comparison with Static and Sliding-Window Correlation

We compared CL with the following established alternatives, all computed from the same envelopes: (i) static power-envelope correlation (PEC), the Pearson correlation of the log-envelopes over the whole recording; (ii) static amplitude-envelope correlation (AEC) of the untransformed amplitude envelopes; (iii) sliding-window Pearson correlation (SW) with windows of 10, 30, and 60 s and (iv) with a window w^* tuned for each pair; (v) the kernel local Pearson correlation (KLP), which uses the same Gaussian kernel as CL but centres and scales the series locally at the bandwidth; (vi) CL with a single bandwidth for all pairs (the median of h_{rq}); and (vii) CE, i.e., CL without boundary correction. Comparisons (v)–(vii) isolate, respectively, the effects of local centring, of pair-specific bandwidths and of the boundary correction. Near the edges of the recording, the sliding windows are truncated.

To compare methods without favouring the one whose tuning criterion is used for evaluation, we split the recording into alternating 30 s blocks A and B. All tuning parameters (h_{rq} including the constant-correlation option, the global bandwidth, w^*, and the KLP bandwidth) were selected on A by the blocked criterion of Section 4.2, and the Gaussian predictive log-likelihood was evaluated on B; the roles of A and B were then swapped and the two folds averaged. Table 3 reports, for each band, the median over pairs of the paired difference in held-out log-likelihood and the agreement of the time-averaged CL estimate with static PEC. Across bands, the time-averaged CL estimate agreed closely with static PEC ($r = 0.978$–0.994, Table 3), so the time-varying estimator reproduces the static connectivity when averaged over time. The constant-correlation model was preferred over a time-varying fit for 79.7–89.1% of region pairs. Selecting a separate bandwidth for each pair improved the held-out likelihood relative to a single global bandwidth in all six bands, by a small amount in the delta, theta, and alpha bands (median difference 0.0028–0.0042) and more substantially in the gamma bands (0.0644 in slow gamma and 0.2594 in fast gamma). Compared with the tuned sliding window, KLP and the static model,

CL performed comparably in the delta, theta, alpha and beta bands, whereas in the two gamma bands, the alternatives attained somewhat better held-out likelihoods (median differences −0.03 to −0.12), indicating that for these bands, the pair-specific smooth fits do not improve out-of-sample prediction. Because the region pairs share regions, the Wilcoxon tests over pairs are descriptive. The remaining comparisons—the agreement with static AEC, the correlation between the CL trajectory and fixed sliding windows of 10, 30, and 60 s, and the boundary-correction ablation CE—are reported in the online Supplementary Materials.

Table 3. Benchmark of CL against alternatives. Columns 2–5 report, for each band, the median over 19,900 region pairs of the held-out negative log-likelihood difference NLL(other) − NLL(CL), evaluated on a nested 30 s block split, so positive values favour CL; "global" denotes CL with a single bandwidth for all pairs. Column 6 reports the correlation between the time-averaged CL estimate and static PEC, and column 7 the percentage of pairs for which the constant-correlation model was selected.

Band	Global	SW(w^*)	KLP	Static	r(CL,PEC)	% Const.
Delta	+0.0042	+0.0000	+0.0000	+0.0000	0.992	85.3
Theta	+0.0028	+0.0000	+0.0000	+0.0000	0.978	79.7
Alpha	+0.0029	+0.0001	+0.0000	+0.0000	0.994	81.2
Beta	+0.0067	+0.0003	+0.0002	−0.0008	0.994	89.1
Slow gamma	+0.0644	−0.0328	−0.0331	−0.0360	0.990	86.1
Fast gamma	+0.2594	−0.1148	−0.1172	−0.1173	0.989	86.6

4.4. Dynamic Network Organisation

Figures 7 and 8 show, for the delta and fast-gamma bands, the eigenvalue variance of the CL correlation matrices over time. For a correlation matrix, this equals $(p-1)\,\overline{r^2}(t)$ and thus summarises the overall strength of coupling; the beta, alpha, theta and slow-gamma bands are shown in Figures S4–S7 of the online Supplementary Materials. In the delta band, the eigenvalue variance ranged from 1.91 to 2.89, and in the fast-gamma band, from 29.60 to 52.33; for theta, alpha, beta and slow gamma, the ranges were 2.46–3.75, 3.84–4.89, 20.77–27.45 and 25.20–44.34. Overall coupling therefore increased from the slow to the fast bands, and the temporal fluctuations were largest in the two gamma bands. Each figure also shows the conventional counterparts. The CL trajectory fluctuates around the static PEC value (time average 2.14 versus static 2.14 in delta and 49.35 versus 52.86 in fast gamma), and its course over time agrees with that of the tuned sliding window (correlation between the two trajectories 0.87 in delta and 0.61 in fast gamma; the sliding window shows larger excursions). The overall coupling identified by CL is therefore consistent with that of conventional approaches; the differences that are specific to CL concern the pair-specific smoothing and the estimates at the edges of the recording, which are quantified in Section 4.3.

The chord diagrams show the CL Yeo-7 network matrices at six time points. In the delta band, the strongest couplings involved the limbic network—with the default-mode, ventral-attention, and frontoparietal networks—and the weakest involved the visual network with the somatomotor and dorsal-attention networks, a pattern that was stable across the six time points. In the fast-gamma band, the same hierarchy reappeared at much higher coupling strengths and with larger temporal variation, whereas in the alpha band, the strongest couplings involved the somatomotor network and the visual network showed the weakest coupling; Figure S5 of the online Supplementary Materials shows the slow-gamma band. Averaged over time, the CL network matrix agrees with the static PEC network matrix (correlation over the 28 network pairs 0.86 to 1.00 across bands). The

differences between bands are therefore already present in static connectivity; what CL adds is how these network patterns vary within the recording.

Periodicity Analysis. To assess whether the observed temporal variations in functional connectivity exhibited significant periodicity, we applied an autoregressive (AR)-based global periodicity test to each network-pair connectivity time series. This procedure was designed to account for the substantial temporal dependence expected in dynamic functional connectivity [33,74–77]. For each frequency band and each pair of Yeo-7 functional networks, the resulting connectivity time series was first centred and standardised. An AR(p) model was then fitted to the standardised series y_t,

$$y_t = c + \sum_{\ell=1}^{p} \phi_\ell y_{t-\ell} + \varepsilon_t,$$

where the autoregressive order p was selected by minimising the Akaike information criterion (AIC) over $p = 1, \ldots, 20$. The fitted autoregressive coefficients were subsequently used to whiten both the observed connectivity series and the design matrices.

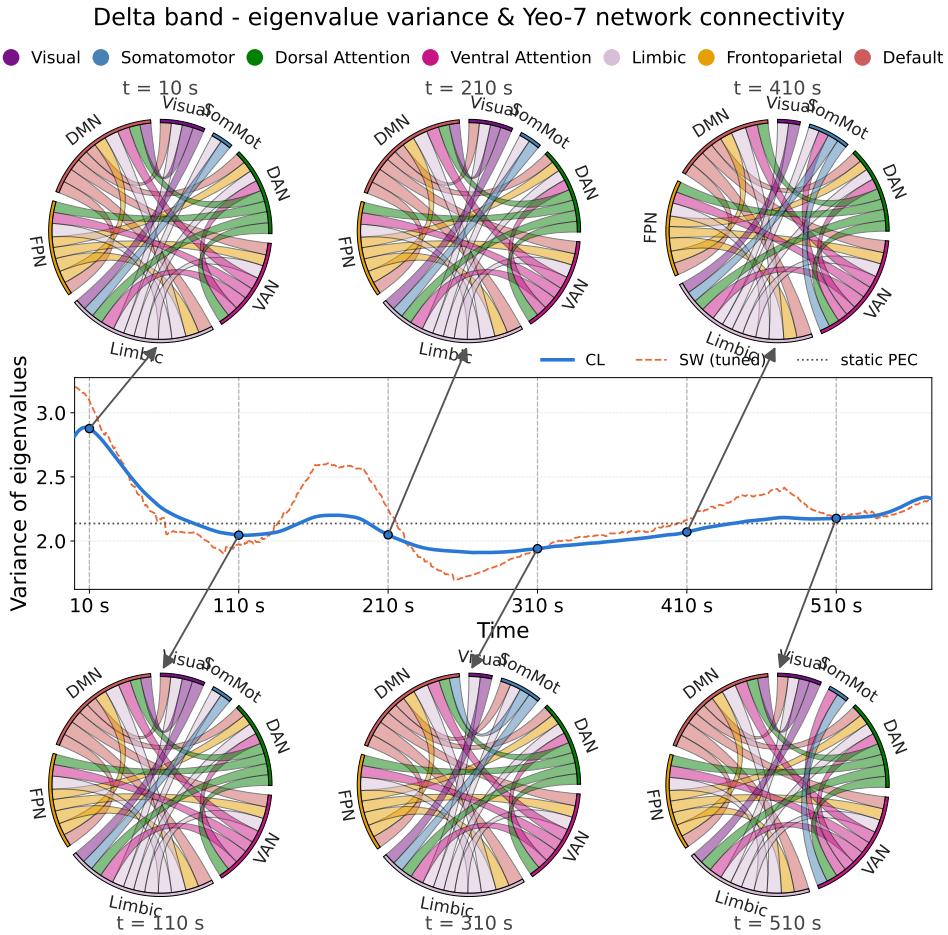


Figure 7. Eigenvalue variance of the dynamic correlation matrices in the delta band, which equals $(p-1)\,\overline{r^2}(t)$, for CL (solid), the tuned sliding window (dashed), and the static PEC matrix (dotted). Chord diagrams show the CL Yeo-7 network matrices at 10, 110, 210, 310, 410, and 510 s; for display only, links below the median of each matrix are hidden. Networks: Visual, Somatomotor (SomMot), Dorsal Attention (DAN), Ventral Attention (VAN), Limbic, Frontoparietal (FPN), and Default (DMN).

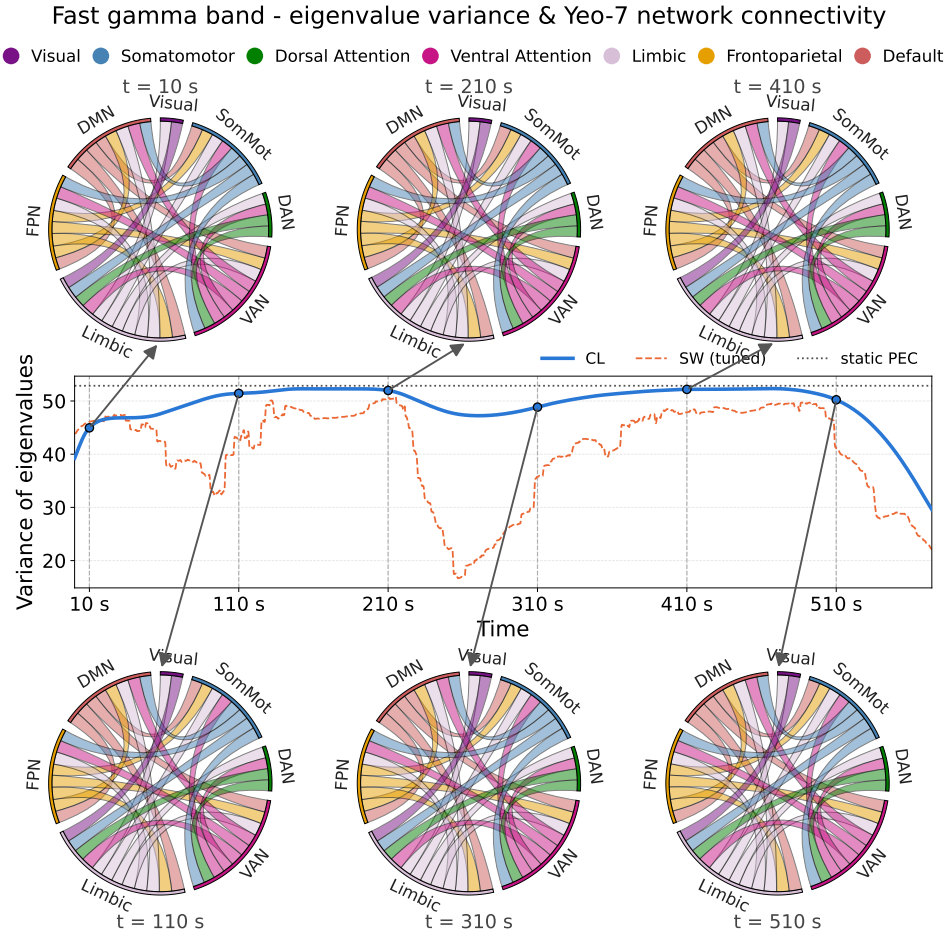


Figure 8. Eigenvalue variance of the dynamic correlation matrices in the fast gamma band, which equals $(p-1)\,\overline{r^2}(t)$, for CL (solid), the tuned sliding window (dashed), and the static PEC matrix (dotted). Chord diagrams show the CL Yeo-7 network matrices at 10, 110, 210, 310, 410, and 510 s; for display only, links below the median of each matrix are hidden. Networks: Visual, Somatomotor (SomMot), Dorsal Attention (DAN), Ventral Attention (VAN), Limbic, Frontoparietal (FPN), and Default (DMN).

To test for periodic components over a prespecified range of connectivity fluctuation frequencies, we considered a grid of 500 equally spaced frequencies within the interval corresponding to periods of 5–300 s, i.e., $f \in [1/300, 1/5]$ Hz. For each candidate frequency f, we compared a null model containing only an intercept with a harmonic model,

$$X_1(f) = \begin{bmatrix} \mathbf{1} & \cos(2\pi f t) & \sin(2\pi f t) \end{bmatrix}, \qquad X_0 = \mathbf{1}.$$

Let $L \in \mathbb{R}^{n_{eff} \times T}$ denote the whitening matrix, where n_{eff} is the effective number of observations after accounting for the AR model and T is the total number of time points. Hence, the whitened response and design matrix are as follows:

$$\tilde{y} = Ly \in \mathbb{R}^{n_{eff}}, \qquad \tilde{X}_1(f) = LX_1(f) \in \mathbb{R}^{n_{eff} \times 3}, \qquad \tilde{X}_0 = LX_0 \in \mathbb{R}^{n_{eff} \times 1}.$$

After AR whitening, the two nested models $\tilde{y} \sim \tilde{X}_1(f)$ and $\tilde{y} \sim \tilde{X}_0$ were fitted by least squares and compared using the statistic

$$F(f) = \frac{[RSS_0(f) - RSS_1(f)]/2}{RSS_1(f)/(n_{\text{eff}} - 3)},$$

which, under the corresponding model assumptions, was evaluated against an $F_{2,n_{\text{eff}}-3}$ distribution. The frequency providing the strongest evidence for a periodic component was defined as $\widehat{f}_{\text{AR}} = \arg\min_f p(f)$, $p_{\text{AR}} = \min_f p(f)$, where $p(f)$ is the p-value corresponding to the F-statistic $F(f)$. The corresponding dominant period was calculated as $\widehat{T}_{\text{AR}} = 1/\widehat{f}_{\text{AR}}$. Finally, the resulting p_{AR} values were adjusted for multiple comparisons across all frequency-band and network-pair combinations using the Benjamini–Hochberg false discovery rate (FDR) procedure, yielding q_{AR} values. Connectivity fluctuations were considered statistically significant when $q_{\text{AR}} < 0.005$.

Kernel smoothing and AR whitening can make a smooth stationary series appear periodic, so the nominal q_{AR} values are not directly interpretable. To calibrate the test, we applied the same pipeline (CL at the bandwidths selected on the observed data, network averaging, the AR test and the Benjamini–Hochberg adjustment over the 168 band–pair combinations) to 19 stationary surrogates of the log-envelopes, generated by multivariate phase randomisation, which preserves all auto- and cross-spectra and hence the static correlations [78]. A band is taken to show evidence of periodic coupling only if its number of significant network pairs exceeds that of the surrogates, and we report the Monte Carlo p-value $p_{\text{MC}} = (1 + \#\{s : n_s \geq n_{\text{obs}}\})/20$. For comparison, the same test was also applied to the network series of the tuned sliding window.

The results are illustrated in Figure 9. For CL, 80 of the 168 band–pair combinations showed a nominally significant periodic component (delta 11/28, theta 7/28, alpha 4/28, beta 24/28, slow gamma 10/28, and fast gamma 24/28), compared with a median of 5–21 significant pairs per band (range 0–28) in the stationary surrogates and all 28 pairs in every band for the tuned sliding window. The Monte Carlo p-values were $p_{\text{MC}} = 0.90$ (delta), 1.00 (theta), 1.00 (alpha), 0.35 (beta), 0.90 (slow gamma), and 0.20 (fast gamma). In no band did the number of periodic network pairs exceed that obtained from stationary surrogates. The nominal periodicity of the network series is therefore explained by the smoothing and the autocorrelation of the series, and we do not interpret the dominant periods.

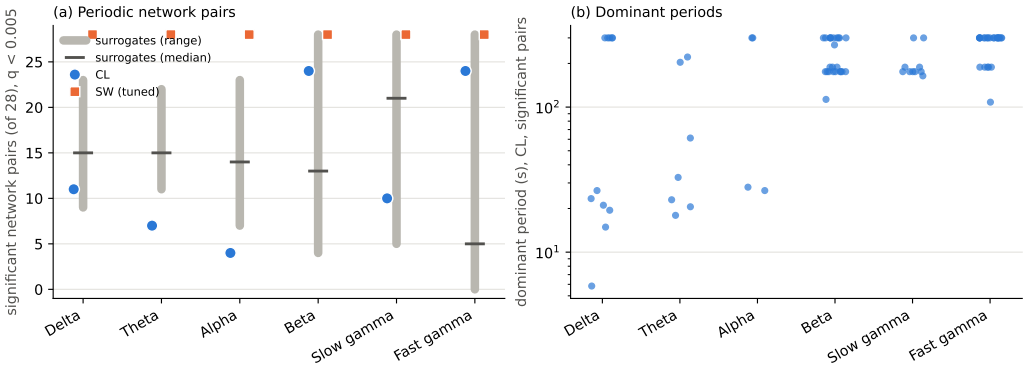


Figure 9. Periodicity analysis of the Yeo-7 network time series. An autoregressive (AR) global periodicity test over periods of 5–300 s is applied to each of the 168 frequency-band–network-pair time series, and the resulting p_{AR} values are adjusted by the Benjamini–Hochberg procedure over the 168 combinations; a network pair is declared significant when $q_{\text{AR}} < 0.005$. (**a**) Number of network pairs (out of 28) with a significant periodic component in each frequency band, for CL and the tuned sliding window (SW), together with the median and range over 19 stationary phase-randomised surrogates, which preserve all auto- and cross-spectra and hence the static correlations and are used to calibrate the test. (**b**) Dominant periods $\widehat{T}_{\text{AR}}$ of the significant CL network pairs.

5. Conclusions and Discussion

In this paper, we proposed a nonparametric correlation estimator via solving a cubic equation (CE) to carry out the time-varying correlation analysis in MEG power envelope. Furthermore, we also corrected the boundary effects by implementation of the local linear

regression (CL) method. The proposed estimator is nonparametric with respect to the temporal evolution of the correlation, imposing no parametric form on $\rho(\cdot)$; it is, however, derived from a conditional bivariate Gaussian working model with zero means and unit variances, so the data must be centred and scaled before use, and its variance formulae rely on normality (Remark 2). We not only provided the theoretical properties of the proposed estimators (CE and CL), but also conducted extensive simulation studies to demonstrate their superior performance over the existing methods. Considering the trade-off between MSE and computational cost, we recommend using the CL method for large-scale time-varying correlation analysis, as the CL method reduces computational cost relative to the LL method, while simultaneously preserving a similar level of MSE.

We applied the CL method to one subject's real MEG scan from the Cam-CAN dataset. Following a series of preprocessing steps, including channel selection, band-pass filtering, power-line noise removal, anatomical processing, boundary element modelling, and computation of the forward solution, we obtained the source-space time series for the entire recording. These source estimates were subsequently parcellated according to the Schaefer-200 cortical atlas to obtain regional time series for the 200 cortical ROIs. Before extracting the power envelopes, we applied source leakage correction following established procedures in the literature.

Once the power envelopes had been obtained, we applied CL to all 19,900 region pairs in six frequency bands to construct dynamic correlation matrices, and compared them with static and sliding-window correlations. Overall coupling was weakest in the delta, theta, and alpha bands and strongest in the beta and gamma bands, with the time-averaged $(p-1)\,\overline{r^2}(t)$ rising from 2.14 (delta), 2.97 (theta), and 4.27 (alpha) to 26.34 (beta), 41.60 (slow gamma), and 49.35 (fast gamma). The Yeo-7 network summaries varied within the recording in each band, and their strength increased from the slower to the faster bands. These results show that the approach can be applied to whole-brain MEG data. More broadly, the proposed framework offers a flexible nonparametric way of characterising temporally evolving dependence structures, without imposing a restrictive parametric form on the underlying correlation dynamics.

From statistical properties to the MEG results. The MEG analysis served two purposes, which we kept separate: showing that the estimator can be applied to whole-brain resting-state MEG (Section 4.4), and examining whether its statistical properties make a measurable difference relative to conventional approaches (Section 4.3 and the online Supplementary Materials). Each property maps onto a specific comparison. (i) Time-varying estimation. A smooth $\rho(t)$ was compared with a constant correlation selected by the same criterion: 79.7–89.1% of region pairs preferred a constant correlation, and among the remaining pairs, the smooth fit did not improve the held-out likelihood relative to the constant model (median difference -0.0013 to -0.0228, Table 3); the estimated fluctuations exceeded those of stationary surrogates in the delta, theta, alpha, and slow-gamma bands ($p = 0.01$–0.03) but not in the beta and fast-gamma bands ($p = 0.06$–0.08; online Supplementary Materials), and the periodicity of the network series did not exceed its surrogate calibration. (ii) Pair-specific bandwidths. The selected bandwidths ranged from 19.7 to 60 s (10th–90th percentiles) and differed between within- and between-network pairs only in the beta band (Kruskal–Wallis $p < 10^{-6}$); relative to a single bandwidth for all pairs, held-out prediction improved in all six bands. (iii) Boundary correction. CL provides estimates within the first and last 30 s of the recording, where the sliding window must be truncated; the price is a larger variance at the edges (Theorem 5, Remark 6 and the MEG-like simulation in the online Supplementary Materials), so estimates near the edges should be read with this in mind. Where CL and the conventional approaches agree—for the time-averaged connectivity and the overall course of coupling—the MEG results corrob-

orate known resting-state structure rather than reveal new phenomena; where they differ, the comparisons above indicate which property of the estimator is responsible. Because only one participant was analysed, none of these results is a population-level finding.

Limitations. The proposed methodology also has several limitations that should be acknowledged. First, the proposed framework is derived under a conditional bivariate Gaussian working model with zero conditional means and unit conditional variances. The zero-mean and unit-variance conditions are required for identification and must be enforced by standardisation; Gaussianity itself is used for the stated asymptotic variances and for the uniform consistency argument, whereas pointwise consistency holds under the moment conditions alone (Remark 2). Although a logarithmic transformation and local standardisation are applied before correlation estimation, we do not perform an additional formal assessment of normality. In practical applications, we recommend that the distributional assumptions be carefully assessed, and the normality of the transformed data be examined where appropriate, before applying the proposed method. If the normality test fails on the transformed data, a further discussion on transformation for heavy-tailed data should be considered. We also acknowledge that the current theoretical analysis is developed under an i.i.d. framework with unit marginal variances. These assumptions are primarily adopted to make the asymptotic analysis of the cubic estimating equation tractable and to isolate the statistical properties of the proposed nonparametric correlation estimator. Some non-i.i.d cases were fully discussed in Zhang and Li [79], Li [80]. The theoretical results are based on an i.i.d. unit variance setting and do not fully account for temporal dependence, bandwidth estimation, and multiple pairwise estimation effects, which we leave as important extensions for future work.

Second, because the proposed correlation estimator is constructed independently for each pair of variables, the resulting correlation matrix is not guaranteed to be positive definite. Enforcing positive definiteness at the estimation stage would require the same bandwidth to be used across all matrix entries. Although such a strategy could preserve the positive-definite structure of the estimated correlation matrix, it would introduce other important drawbacks. In particular, the computational cost would increase substantially because the bandwidth would need to be jointly selected across all variable pairs, and the resulting bandwidth could be dominated by the sparsest or most difficult-to-estimate pairs [80]. We therefore adopt a compromise in which the correlations are estimated independently for each pair, allowing the pairwise estimation problems to be readily parallelised, and the resulting matrix is subsequently projected onto the nearest correlation matrix. While this strategy provides a practical balance between computational efficiency and matrix validity, the projection step inevitably introduces some distortion to the estimated pairwise correlations [73]. The extent to which this projection affects downstream statistical inference and the preservation of the underlying dependence structure has not been fully investigated in the present study and warrants further investigation.

Third, the proposed nonparametric correlation estimator is based on a local smoothing framework and may therefore become less stable when the sample size is small. This limitation is less consequential for high-frequency data, where numerous observations are typically available, such as financial transaction data, MEG, and EEG recordings.

Fourth, the bandwidth selection procedure employed in this study is based on cross-validation. Although we have introduced several computational optimisations to improve its efficiency, bandwidth selection may still impose a considerable computational burden, particularly for large-scale datasets involving numerous variable pairs or observations.

Fifth, in terms of estimation accuracy, the CL method remains slightly less accurate than the LL method in some settings. This represents a trade-off between computational

efficiency and statistical accuracy that should be considered when selecting an estimator for a particular application.

Finally, the proposed nonparametric correlation estimator was developed for continuously distributed data. Its direct application to discrete-valued data is therefore not currently supported and would require appropriate modifications to the estimation procedure. This limitation is inherent to the current formulation of the estimator and represents an important direction for future methodological development.

Scope, Generalisability, and Future Directions. In this study, we focused on time-varying correlation analysis based on power envelopes, as this framework has been widely used in neuroscience and provides a useful measure of functional interactions between brain regions. Nevertheless, the proposed nonparametric correlation estimation method is not restricted to power envelope connectivity and can, in principle, be extended to other forms of time-varying dependence analysis. For example, it could be applied to time-varying measures of phase synchronisation, coherence, phase-locking value (PLV), weighted phase-lag index (wPLI), imaginary coherence, and other measures of frequency-specific functional connectivity [81]. It could also be adapted to time-varying amplitude-envelope connectivity and cross-frequency coupling measures. Future work could investigate these extensions to assess the general applicability and robustness of the proposed methodology across different representations of dynamic functional connectivity.

We used the variance of the eigenvalues as a summary measure because it provides a compact characterisation of the overall connectivity strength of a time-varying correlation matrix. Other graph-theoretic and matrix-based measures, such as mean connectivity strength, mean clustering coefficient, average path length, network density, and global efficiency, could also be used to provide a more comprehensive characterisation of the temporal evolution of functional connectivity. Owing to space limitations, we do not report these additional analyses in detail here. They nevertheless provide natural extensions for future analyses and could offer complementary perspectives on the temporal organisation of the estimated correlation matrices.

The analysis of the time-varying correlation matrices across the six frequency bands revealed clear frequency-dependent differences in the overall connectivity structure. At the same time, some similarities in connectivity patterns were also observed across different frequency bands, suggesting that frequency-specific network organisation may coexist with common features of large-scale functional architecture. These observations are consistent with the view that functional brain networks exhibit both frequency-specific and frequency-general characteristics. However, the present analysis was based on MEG data from a single participant, and the observed patterns should therefore not be interpreted as representative of the general population. Analyses based on a substantially larger cohort will be required to assess the robustness, reproducibility, and generalisability of these findings, as well as to determine whether the observed frequency-specific patterns are consistent across individuals.

Finally, the MEG data analysed in this study were acquired during the resting state. An important direction for future research would therefore be to apply the proposed methodology to task-based MEG or EEG recordings. Such analyses could provide an opportunity to investigate how functional connectivity between brain regions changes dynamically in response to different cognitive demands and whether the proposed nonparametric framework can capture task-related transitions in functional network organisation. Extending the methodology to task-evoked and other experimentally controlled settings would also provide a further test of its ability to characterise dynamic dependence structures under different neural and behavioural conditions.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/app162111083/s1, Figure S1: MEG-like simulation; Figure S2: Q–Q plots of the standardised log-envelopes against the standard normal quantiles, one panel per frequency band; Figure S3: Heatmap of the median logarithmic MSEs across different correlation functions and kernel functions for each combination of estimation method, sampling strategy, and sample size; Figure S4: Eigenvalue variance of the dynamic correlation matrices in the beta band; Figure S5: Eigenvalue variance of the dynamic correlation matrices in the alpha band; Figure S6: Eigenvalue variance of the dynamic correlation matrices in the theta band; Figure S7: Eigenvalue variance of the dynamic correlation matrices in the slow gamma band; Table S1: Monte Carlo check of the boundary variance of the CL estimator at $u_0 = -1 + ch$; Table S2: Fraction of CL evaluations (in %) at which the modified cubic equation had no root in $(-1, 1)$ and the truncation fallback with $\varepsilon = 10^{-3}$ was applied, by sample size n and sampling strategy; Table S3: Diagnostics of the standardised log-envelopes by frequency band: median (and range over the 200 regions) of the skewness and excess kurtosis, and the smallest ratio of the smoothed power to its region-wise median; Table S4: Sensitivity of the MEG network results to the standardisation of the log-envelopes, relative to the main analysis (local, $\sigma_p = 60$ s); Table S5: Additional benchmark results complementing Table 3 of the main text. References [82,83] are mentioned in the supplementary materials.

Author Contributions: Conceptualisation, S.Y., G.G., J.Z., A.G., and J.L.; methodology, S.Y., G.G., J.Z., A.G., and J.L.; software, S.Y. and J.L.; validation, S.Y. and J.L.; formal analysis, S.Y. and J.L.; investigation, S.Y. and J.L.; resources, S.Y. and J.L.; data curation, S.Y.; writing—original draft preparation, S.Y. and J.L.; writing—review and editing, S.Y., G.G., J.Z., A.G., and J.L.; visualisation, J.L.; supervision, J.L.; project administration, J.L.; funding acquisition, S.Y. and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Fourth Cohort of Young and Middle-aged Academic and Technical Talent Candidates at Lijiang Culture and Tourism College, grant number: 2026xshb02. This research was aslo supported by the EPSRC project: From Covariance Regressions to Nonparametric Dynamic Causal Modelling (CoreDCM), grant number: EP/X038297/1.

Data Availability Statement: The subject data come from the Cambridge Centre for Ageing and Neuroscience (Cam-CAN) dataset https://cam-can.mrc-cbu.cam.ac.uk/dataset/ (accessed on 2 October 2025) [51]. The simulation code is available on the GitHub repository: https://github.com/Jieli12/nonparametric_correlation_estimator_cubic (accessed on 24 October 2026).

Acknowledgments: During the preparation of this manuscript/study, the authors used ChatGPT 5.3 and DeepSeek 4.1 for the purposes of deploying AI-assisted tools for language translation, language editing and polishing, and grammar checking. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Woolrich, M.W.; Baker, A.; Luckhoo, H.; Mohseni, H.; Barnes, G.; Brookes, M.; Rezek, I. Dynamic State Allocation for MEG Source Reconstruction. *NeuroImage* **2013**, *77*, 77–92. [CrossRef] [PubMed]
2. Favaretto, C.; Allegra, M.; Deco, G.; Metcalf, N.V.; Griffis, J.C.; Shulman, G.L.; Brovelli, A.; Corbetta, M. Subcortical-Cortical Dynamical States of the Human Brain and Their Breakdown in Stroke. *Nat. Commun.* **2022**, *13*, 5069. [CrossRef] [PubMed]
3. Jin, H.; Ranasinghe, K.G.; Prabhu, P.; Dale, C.; Gao, Y.; Kudo, K.; Vossel, K.; Raj, A.; Nagarajan, S.S.; Jiang, F. Dynamic Functional Connectivity MEG Features of Alzheimer's Disease. *NeuroImage* **2023**, *281*, 120358. [CrossRef] [PubMed]
4. Hipp, J.F.; Hawellek, D.J.; Corbetta, M.; Siegel, M.; Engel, A.K. Large-Scale Cortical Correlation Structure of Spontaneous Oscillatory Activity. *Nat. Neurosci.* **2012**, *15*, 884–890. [CrossRef] [PubMed]
5. O'Neill, G.C.; Barratt, E.L.; Hunt, B.A.E.; Tewarie, P.K.; Brookes, M.J. Measuring Electrophysiological Connectivity by Power Envelope Correlation: A Technical Review on MEG Methods. *Phys. Med. Biol.* **2015**, *60*, R271–R295. [CrossRef] [PubMed]
6. Brookes, M.J.; Woolrich, M.; Luckhoo, H.; Price, D.; Hale, J.R.; Stephenson, M.C.; Barnes, G.R.; Smith, S.M.; Morris, P.G. Investigating the Electrophysiological Basis of Resting State Networks Using Magnetoencephalography. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 16783–16788. [CrossRef] [PubMed]

7. Hämäläinen, M.S.; Sarvas, J. Realistic Conductivity Geometry Model of the Human Head for Interpretation of Neuromagnetic Data. *IEEE Trans. Biomed. Eng.* **1989**, *36*, 165–171. [CrossRef] [PubMed]
8. Gramfort, A. MEG and EEG Data Analysis with MNE-Python. *Front. Neurosci.* **2013**, *7*, 267. [CrossRef] [PubMed]
9. Bruns, A. Fourier-, Hilbert- and Wavelet-Based Signal Analysis: Are They Really Different Approaches? *J. Neurosci. Methods* **2004**, *137*, 321–332. [CrossRef] [PubMed]
10. Coelli, S.; Corda, M.; Bianchi, A.M. The Time-Varying Brain: A Comprehensive Review of Dynamic Functional Connectivity Analysis in EEG and MEG. *J. Neural Eng.* **2025**, *22*, 051005. [CrossRef] [PubMed]
11. Carrasco-Gómez, M.; García-Colomo, A.; Cabrera-Álvarez, J.; Bruña, R.; Santos, A.; Maestú, F. Dynamics of Brain Connectivity across the Alzheimer's Disease Spectrum through Magnetoencephalography. *Sci. Rep.* **2026**, *16*, 3905. [CrossRef] [PubMed]
12. Fischl, B. FreeSurfer. *NeuroImage* **2012**, *62*, 774–781. [CrossRef] [PubMed]
13. Schaefer, A.; Kong, R.; Gordon, E.M.; Laumann, T.O.; Zuo, X.N.; Holmes, A.J.; Eickhoff, S.B.; Yeo, B.T.T. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb. Cortex* **2018**, *28*, 3095–3114. [CrossRef] [PubMed]
14. Wen, S.; Wang, J.; Liu, W.; Meng, X.; Jiao, Z. Spatio-Temporal Dynamic Functional Brain Network for Mild Cognitive Impairment Analysis. *Front. Neurosci.* **2025**, *19*, 1597777. [CrossRef] [PubMed]
15. Zhang, J.; Lyu, Y.; Yu, X.; Zhang, L.; Cao, C.; Chen, T.; Chen, M.; Zhuang, Y.; Liu, T.; Zhu, D. Classification of Mild Cognitive Impairment Based on Dynamic Functional Connectivity Using Spatio-Temporal Transformer. In *Proceedings of the 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI)*; IEEE: Piscataway, NJ, USA, 2025; Volume 2025, pp. 1–12. [CrossRef] [PubMed]
16. Kajimura, S.; Margulies, D.; Smallwood, J. Frequency-Specific Brain Network Architecture in Resting-State fMRI. *Sci. Rep.* **2023**, *13*, 2964. [CrossRef] [PubMed]
17. Lin, F.H.; Witzel, T.; Hämäläinen, M.S.; Dale, A.M.; Belliveau, J.W.; Stufflebeam, S.M. Spectral Spatiotemporal Imaging of Cortical Oscillations and Interactions in the Human Brain. *NeuroImage* **2004**, *23*, 582–595. [CrossRef] [PubMed]
18. O'Neill, G.C.; Tewarie, P.; Vidaurre, D.; Liuzzi, L.; Woolrich, M.W.; Brookes, M.J. Dynamics of Large-Scale Electrophysiological Networks: A Technical Review. *NeuroImage* **2018**, *180*, 559–576. [CrossRef] [PubMed]
19. Lachaux, J.P.; Rodriguez, E.; Martinerie, J.; Varela, F.J. Measuring Phase Synchrony in Brain Signals. *Hum. Brain Mapp.* **1999**, *8*, 194–208. [CrossRef]
20. Nolte, G.; Bai, O.; Wheaton, L.; Mari, Z.; Vorbach, S.; Hallett, M. Identifying True Brain Interaction from EEG Data Using the Imaginary Part of Coherency. *Clin. Neurophysiol.* **2004**, *115*, 2292–2307. [CrossRef] [PubMed]
21. Stam, C.J.; Nolte, G.; Daffertshofer, A. Phase Lag Index: Assessment of Functional Connectivity from Multi Channel EEG and MEG with Diminished Bias from Common Sources. *Hum. Brain Mapp.* **2007**, *28*, 1178–1193. [CrossRef] [PubMed]
22. Vinck, M.; Oostenveld, R.; van Wingerden, M.; Battaglia, F.; Pennartz, C.M.A. An Improved Index of Phase-Synchronization for Electrophysiological Data in the Presence of Volume-Conduction, Noise and Sample-Size Bias. *NeuroImage* **2011**, *55*, 1548–1565. [CrossRef] [PubMed]
23. Baselice, F.; Sorriso, A.; Rucco, R.; Sorrentino, P. Phase Linearity Measurement: A Novel Index for Brain Functional Connectivity. *IEEE Trans. Med. Imaging* **2019**, *38*, 873–882. [CrossRef] [PubMed]
24. Colclough, G.; Brookes, M.; Smith, S.; Woolrich, M. A Symmetric Multivariate Leakage Correction for MEG Connectomes. *Neuroimage* **2015**, *117*, 439–448. [CrossRef] [PubMed]
25. Colclough, G.L.; Woolrich, M.W.; Tewarie, P.K.; Brookes, M.J.; Quinn, A.J.; Smith, S.M. How Reliable Are MEG Resting-State Connectivity Metrics? *NeuroImage* **2016**, *138*, 284–293. [CrossRef] [PubMed]
26. Siems, M.; Siegel, M. Dissociated Neuronal Phase- and Amplitude-Coupling Patterns in the Human Brain. *NeuroImage* **2020**, *209*, 116538. [CrossRef] [PubMed]
27. Baker, A.P.; Brookes, M.J.; Rezek, I.A.; Smith, S.M.; Behrens, T.; Probert Smith, P.J.; Woolrich, M. Fast Transient Networks in Spontaneous Human Brain Activity. *eLife* **2014**, *3*, e01867. [CrossRef] [PubMed]
28. Vidaurre, D.; Hunt, L.T.; Quinn, A.J.; Hunt, B.A.E.; Brookes, M.J.; Nobre, A.C.; Woolrich, M.W. Spontaneous Cortical Activity Transiently Organises into Frequency Specific Phase-Coupling Networks. *Nat. Commun.* **2018**, *9*, 2987. [CrossRef] [PubMed]
29. Tewarie, P.; Liuzzi, L.; O'Neill, G.C.; Quinn, A.J.; Griffa, A.; Woolrich, M.W.; Stam, C.J.; Hillebrand, A.; Brookes, M.J. Tracking Dynamic Brain Networks Using High Temporal Resolution MEG Measures of Functional Connectivity. *NeuroImage* **2019**, *200*, 38–50. [CrossRef] [PubMed]
30. Leonardi, N.; Van De Ville, D. On Spurious and Real Fluctuations of Dynamic Functional Connectivity during Rest. *NeuroImage* **2015**, *104*, 430–436. [CrossRef] [PubMed]
31. Hindriks, R.; Adhikari, M.H.; Murayama, Y.; Ganzetti, M.; Mantini, D.; Logothetis, N.K.; Deco, G. Can Sliding-Window Correlations Reveal Dynamic Functional Connectivity in Resting-State fMRI? *NeuroImage* **2016**, *127*, 242–256. [CrossRef] [PubMed]
32. Zalesky, A.; Breakspear, M. Towards a Statistical Test for Functional Connectivity Dynamics. *NeuroImage* **2015**, *114*, 466–470. [CrossRef] [PubMed]

33. Liégeois, R.; Laumann, T.O.; Snyder, A.Z.; Zhou, J.; Yeo, B.T.T. Interpreting Temporal Fluctuations in Resting-State Functional Connectivity MRI. *NeuroImage* **2017**, *163*, 437–455. [CrossRef] [PubMed]
34. Rahimi, S.; Jackson, R.; Farahibozorg, S.R.; Hauk, O. Time-Lagged Multidimensional Pattern Connectivity (TL-MDPC): An EEG/MEG Pattern Transformation Based Functional Connectivity Metric. *NeuroImage* **2023**, *270*, 119958. [CrossRef] [PubMed]
35. Fan, J.; Heckman, N.E.; Wand, M.P. Local Polynomial Kernel Regression for Generalized Linear Models and Quasi-Likelihood Functions. *J. Am. Stat. Assoc.* **1995**, *90*, 141–150. [CrossRef]
36. Fan, J.; Gijbels, I. Data-Driven Bandwidth Selection in Local Polynomial Fitting: Variable Bandwidth and Spatial Adaptation. *J. R. Stat. Soc. Ser. B (Methodol.)* **1995**, *57*, 371–394. [CrossRef]
37. Fan, J.; Yao, Q.; Tong, H. Estimation of Conditional Densities and Sensitivity Measures in Nonlinear Dynamical Systems. *Biometrika* **1996**, *83*, 189–206. [CrossRef]
38. Fan, J.; Yao, Q. Efficient Estimation of Conditional Variance Functions in Stochastic Regression. *Biometrika* **1998**, *85*, 645–660. [CrossRef]
39. Staniswalis, J.G.; Lee, J.J. Nonparametric Regression Analysis of Longitudinal Data. *J. Am. Stat. Assoc.* **1998**, *93*, 1403–1418. [CrossRef] [PubMed]
40. Kendall, M.G.; Stuart, A.; Ord, J.K. *Inference and Relationship*, 3rd ed.; Number 2 in the Advanced Theory of Statistics; Griffin: London, UK, 1973.
41. Fan, J.; Gijbels, I. *Local Polynomial Modelling and Its Applications*, 1st ed.; Number 66 in Monographs on Statistics and Applied Probability; Chapman & Hall/CRC: Boca Raton, FL, USA, 1996.
42. Ahlfors, L.V. *Complex Analysis: An Introduction to the Theory on Analytic Functions of One Complex Variable*, 3rd ed.; International Series in Pure and Applied Mathematics; McGraw-Hill: New York, NY, USA, 2007.
43. Fan, J. Design-Adaptive Nonparametric Regression. *J. Am. Stat. Assoc.* **1992**, *87*, 998–1004. [CrossRef]
44. Ruppert, D.; Wand, M.P. Multivariate Locally Weighted Least Squares Regression. *Ann. Stat.* **1994**, *22*, 1346–1370. [CrossRef]
45. Müller, H.G. Smooth Optimum Kernel Estimators near Endpoints. *Biometrika* **1991**, *78*, 521–530. [CrossRef]
46. Gasser, T.; Sroka, L.; Jennen-steinmetz, C. Residual Variance and Residual Pattern in Nonlinear Regression. *Biometrika* **1986**, *73*, 625–633. [CrossRef]
47. Silverman, B.W. *Density Estimation for Statistics and Data Analysis*; CRC Press: Boca Raton, FL, USA, 1986.
48. Schoffelen, J.M.; Gross, J. Source Connectivity Analysis with MEG and EEG. *Hum. Brain Mapp.* **2009**, *30*, 1857–1865. [CrossRef] [PubMed]
49. Bastos, A.M.; Schoffelen, J.M. A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls. *Front. Syst. Neurosci.* **2016**, *9*, 175. [CrossRef] [PubMed]
50. Liu, Z.; Fukunaga, M.; de Zwart, J.A.; Duyn, J.H. Large-Scale Spontaneous Fluctuations and Correlations in Brain Electrical Activity Observed with Magnetoencephalography. *NeuroImage* **2010**, *51*, 102–111. [CrossRef] [PubMed]
51. Cam-CAN; Shafto, M.A.; Tyler, L.K.; Dixon, M.; Taylor, J.R.; Rowe, J.B.; Cusack, R.; Calder, A.J.; Marslen-Wilson, W.D.; Duncan, J.; et al. The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) Study Protocol: A Cross-Sectional, Lifespan, Multidisciplinary Examination of Healthy Cognitive Ageing. *BMC Neurol.* **2014**, *14*, 204. [CrossRef] [PubMed]
52. Desikan, R.S.; Ségonne, F.; Fischl, B.; Quinn, B.T.; Dickerson, B.C.; Blacker, D.; Buckner, R.L.; Dale, A.M.; Maguire, R.P.; Hyman, B.T.; et al. An Automated Labeling System for Subdividing the Human Cerebral Cortex on MRI Scans into Gyral Based Regions of Interest. *NeuroImage* **2006**, *31*, 968–980. [CrossRef] [PubMed]
53. Guglielmi, A.V.; Cisotto, G.; Erseghe, T.; Badia, L. Frequency-Dependent Functional Connectivity of Brain Networks at Resting-State. In *Proceedings of the 2022 14th Biomedical Engineering International Conference (BMEiCON)*; IEEE: Piscataway, NJ, USA, 2022; pp. 1–5. [CrossRef]
54. Zink, N.; Mückschel, M.; Beste, C. Resting-State EEG Dynamics Reveals Differences in Network Organization and Its Fluctuation between Frequency Bands. *Neuroscience* **2021**, *453*, 43–56. [CrossRef] [PubMed]
55. Conti, M.; D'Onofrio, V.; Lorenzon, L.; Grassi, L.L.; Mascioli, D.; Ferrari, V.; Simonetta, C.; Pavan, S.; Di Giuliano, F.; Pierantozzi, M.; et al. Frequency- and Network-Specific Changes in Functional Connectivity Reflect Pathophysiological Mechanisms across Parkinson's Disease Stages. *Ann. Neurol.* **2026**, *100*, 319–333. [CrossRef] [PubMed]
56. Colgin, L.L.; Denninger, T.; Fyhn, M.; Hafting, T.; Bonnevie, T.; Jensen, O.; Moser, M.B.; Moser, E.I. Frequency of Gamma Oscillations Routes Flow of Information in the Hippocampus. *Nature* **2009**, *462*, 353–357. [CrossRef] [PubMed]
57. Buzsáki, G. *Rhythms of the Brain*, 1st ed.; Oxford University Press: New York, NY, USA, 2011.
58. Zheng, C.; Bieri, K.W.; Hsiao, Y.T.; Colgin, L.L. Spatial Sequence Coding Differs during Slow and Fast Gamma Rhythms in the Hippocampus. *Neuron* **2016**, *89*, 398–408. [CrossRef] [PubMed]
59. Taulu, S.; Simola, J.; Kajola, M. Applications of the Signal Space Separation Method. *IEEE Trans. Signal Process.* **2005**, *53*, 3359–3372. [CrossRef]
60. Taulu, S.; Simola, J. Spatiotemporal Signal Space Separation Method for Rejecting Nearby Interference in MEG Measurements. *Phys. Med. Biol.* **2006**, *51*, 1759–1768. [CrossRef] [PubMed]

61. Widmann, A.; Schröger, E.; Maess, B. Digital Filter Design for Electrophysiological Data—A Practical Approach. *J. Neurosci. Methods* **2015**, *250*, 34–46. [CrossRef] [PubMed]
62. Ségonne, F.; Dale, A.M.; Busa, E.; Glessner, M.; Salat, D.; Hahn, H.K.; Fischl, B. A Hybrid Approach to the Skull Stripping Problem in MRI. *NeuroImage* **2004**, *22*, 1060–1075. [CrossRef] [PubMed]
63. Geddes, L.A.; Baker, L.E. The Specific Resistance of Biological Material—A Compendium of Data for the Biomedical Engineer and Physiologist. *Med. Biol. Eng.* **1967**, *5*, 271–293. [CrossRef] [PubMed]
64. Vorwerk, J.; Cho, J.H.; Rampp, S.; Hamer, H.; Knösche, T.R.; Wolters, C.H. A Guideline for Head Volume Conductor Modeling in EEG and MEG. *NeuroImage* **2014**, *100*, 590–607. [CrossRef] [PubMed]
65. Huang, M.X.; Yurgil, K.A.; Robb, A.; Angeles, A.; Diwakar, M.; Risbrough, V.B.; Nichols, S.L.; McLay, R.; Theilmann, R.J.; Song, T.; et al. Voxel-Wise Resting-State MEG Source Magnitude Imaging Study Reveals Neurocircuitry Abnormality in Active-Duty Service Members and Veterans with PTSD. *NeuroImage Clin.* **2014**, *5*, 408–419. [CrossRef] [PubMed]
66. Lin, F.H.; Witzel, T.; Ahlfors, S.P.; Stufflebeam, S.M.; Belliveau, J.W.; Hämäläinen, M.S. Assessing and Improving the Spatial Accuracy in MEG Source Localization by Depth-Weighted Minimum-Norm Estimates. *NeuroImage* **2006**, *31*, 160–171. [CrossRef] [PubMed]
67. Hämäläinen, M.S.; Ilmoniemi, R.J. Interpreting Magnetic Fields of the Brain: Minimum Norm Estimates. *Med. Biol. Eng. Comput.* **1994**, *32*, 35–42. [CrossRef] [PubMed]
68. Dale, A.M.; Liu, A.K.; Fischl, B.R.; Buckner, R.L.; Belliveau, J.W.; Lewine, J.D.; Halgren, E. Dynamic Statistical Parametric Mapping: Combining fMRI and MEG for High-Resolution Imaging of Cortical Activity. *Neuron* **2000**, *26*, 55–67. [CrossRef] [PubMed]
69. Marinkovic, K.; Dhond, R.P.; Dale, A.M.; Glessner, M.; Carr, V.; Halgren, E. Spatiotemporal Dynamics of Modality-Specific and Supramodal Word Processing. *Neuron* **2003**, *38*, 487–497. [CrossRef] [PubMed]
70. Dale, A.M.; Sereno, M.I. Improved Localizadon of Cortical Activity by Combining EEG and MEG with MRI Cortical Surface Reconstruction: A Linear Approach. *J. Cogn. Neurosci.* **1993**, *5*, 162–176. [CrossRef] [PubMed]
71. Ahlfors, S.P.; Han, J.; Belliveau, J.W.; Hämäläinen, M.S. Sensitivity of MEG and EEG to Source Orientation. *Brain Topogr.* **2010**, *23*, 227–232. [CrossRef] [PubMed]
72. Burman, P.; Chow, E.; Nolan, D. A Cross-Validatory Method for Dependent Data. *Biometrika* **1994**, *81*, 351–358. [CrossRef]
73. Higham, N.J. Computing the Nearest Correlation Matrix—A Problem from Finance. *IMA J. Numer. Anal.* **2002**, *22*, 329–343. [CrossRef]
74. Xu, L.; Choi, S.; Zhao, Y.; Li, M.; Rogers, B.P.; Anderson, A.; Gore, J.C.; Gao, Y.; Ding, Z. Seasonal Variations of Functional Connectivity of Human Brains. *Sci. Rep.* **2023**, *13*, 16898. [CrossRef] [PubMed]
75. Eklund, A.; Andersson, M.; Knutsson, H. Fast Random Permutation Tests Enable Objective Evaluation of Methods for Single-Subject fMRI Analysis. *Int. J. Biomed. Imaging* **2011**, *2011*, 627947. [CrossRef] [PubMed]
76. Choe, A.S.; Jones, C.K.; Joel, S.E.; Muschelli, J.; Belegu, V.; Caffo, B.S.; Lindquist, M.A.; van Zijl, P.C.M.; Pekar, J.J. Reproducibility and Temporal Structure in Weekly Resting-State fMRI over a Period of 3.5 Years. *PLoS ONE* **2015**, *10*, e0140134. [CrossRef] [PubMed]
77. van Es, M.W.J.; Higgins, C.; Gohil, C.; Quinn, A.J.; Vidaurre, D.; Woolrich, M.W. Large-Scale Cortical Functional Networks Are Organized in Structured Cycles. *Nat. Neurosci.* **2025**, *28*, 2118–2128. [CrossRef] [PubMed]
78. Prichard, D.; Theiler, J. Generating Surrogate Data for Time Series with Several Simultaneously Measured Variables. *Phys. Rev. Lett.* **1994**, *73*, 951–954. [CrossRef] [PubMed]
79. Zhang, J.; Li, J. Factorized Estimation of High-dimensional Nonparametric Covariance Models. *Scand. J. Stat.* **2022**, *49*, 542–567. [CrossRef]
80. Li, J. Statistical Inference for High-Dimensional Nonparametric Models. Ph.D. Thesis, University of Kent, Canterbury, UK, 2021. [CrossRef]
81. Liu, Z.Q.; Luppi, A.I.; Hansen, J.Y.; Tian, Y.E.; Zalesky, A.; Yeo, B.T.T.; Fulcher, B.D.; Misic, B. Benchmarking Methods for Mapping Functional Connectivity in the Brain. *Nat. Methods* **2025**, *22*, 1593–1602. [CrossRef] [PubMed]
82. Silverman, B.W. Weak and Strong Uniform Consistency of the Kernel Estimate of a Density and Its Derivatives. *Ann. Stat.* **1978**, *6*, 177–184. [CrossRef]
83. Mack, Y.P.; Silverman, B.W. Weak and Strong Uniform Consistency of Kernel Regression Estimates. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* **1982**, *61*, 405–415. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.