Supplementary methods, archive audit and sensitivity analyses for conditional trial-count estimation of table-tennis displacement waveforms

Trial-count estimates for table tennis displacement waveforms vary with action, representation and analytical assumptions

Mingke Han[1†], Yupeng Shen[1†], Lianzhen Ma[1], Pengfei Jin[2], Chen Yang[3], Chaochong Deng[1], Liang Wang[1*]

[1] School of Physical Education and Sports Science, South China Normal University, Guangzhou 510006, China

[2] Department of Physical Education and Sport, Shanghai Ocean University, Shanghai 201306, China

[3] SWJTU-Leeds Joint School and College of Sports and Health, Southwest Jiaotong University, Chengdu 611756, China

† Mingke Han and Yupeng Shen contributed equally to this work and share first authorship.

* Correspondence: Liang Wang, School of Physical Education and Sports Science, South China Normal University, 378 Outer Ring Road, Guangzhou University Town, Panyu District, Guangzhou, Guangdong 510006, China. Email: 20051116@m.scnu.edu.cn. ORCID: https://orcid.org/0009-0003-9906-0803

This Supplementary Information documents TTMD6 source verification, archive-code boundaries, the 14-label mapping, structural-zero handling, functional-summary estimands, residual-dependence scenarios, temporal registration, coordinate-quality rules, archive-code influence, bootstrap diagnostics and pointwise marginal uncertainty. The empirical scope is limited to the six fixed English labels and two archive displacement-magnitude representations in TTMD6. Third-party coordinates are not redistributed in the submission package [1,2].

S1. Terminology, action codes and inferential target

The original labels are forehand attack, forehand drive, forehand push, backhand attack, backhand drive and backhand push [1]. Public materials contain no trial-level ball speed, spin, placement, impact video or specialist adjudication, so equivalence to narrower technique constructs cannot be confirmed. The manuscript and Supplementary Information therefore use the English labels and codes as fixed archive labels.

Supplementary Table S1 | Codes and operational descriptions for the six fixed action labels

Code	Original English label	Operational description
FA	forehand attack	Fixed archive label
FD	forehand drive	Fixed archive label; not assigned to a narrower topspin subtype
FP	forehand push	Fixed archive label
BA	backhand attack	Fixed archive label
BD	backhand drive	Fixed archive label; not assigned to a narrower topspin subtype
BP	backhand push	Fixed archive label

The source paper's direct label-validation evidence comprised a separate sample of 30 racket-trajectory animations, five per class, identified by 40 table-tennis specialists with a mean accuracy of 92.63%. Push had the highest recognition, and the principal confusion was between attack and drive [1]. This supports coarse recognisability of a small racket-trajectory sample, but is not a trial-level review of the 9,000 primary records and does not validate the body labels or code-to-participant mapping. We therefore do not interpret 92.63% as the label accuracy of the present analysis sample.

The primary estimand is the L^2 trace-type relative reliability of the event-unregistered mean displacement-magnitude waveform from n trials under the included 30 archive codes, a fixed action label, the same acquisition session, the public observable-interval rule, 0–100% time normalisation and a specified coordinate-quality rule. It is not between-day absolute agreement, identity-recognition probability, technical level, competition performance or movement mechanism.

S2. Data source, verification and inclusion boundary

The local TTMD6.rar file was 341,074,031 bytes, with MD5 1c9ce9cbf79dd35dd22f16a7199e2a8c and SHA-256 93d1b52a470f14b9dc0ba0600959bff921be891a3da1b71e609bd328224b354d. The complete Springer ESM object linked to the source Scientific Reports article was downloaded on 15 July 2026. Byte count and both hashes matched, and direct bytewise comparison found no difference. The input analysed here was therefore identical to that publisher-hosted object [1].

A 2026 figshare record, 'Table Tennis MOCAP Dataset', version 2 (version DOI 10.6084/m9.figshare.31746358.v2; concept DOI 10.6084/m9.figshare.31746358), is 341,085,926 bytes and was not verified as byte-identical to the 341,074,031-byte archive analysed here. It is cited as a later data record and is not treated as a 2024 publication object or as the repository for this study's code and derived data [2].

Complete extraction yielded 12,000 body files and 12,000 racket files for 40 codes, six action labels and 50 trials per cell. The source article reported only 30 participants, and codes 1–30 formed the first complete contiguous block. The primary cohort was therefore set deterministically to $30 \times 6 \times 50 = 9{,}000$ matched

pairs; 3,000 pairs under codes 31–40 were quarantined because their provenance was unexplained. This rule does not establish a one-to-one mapping between codes 1–30 and the reported participants. Eight complete racket–body duplicates linked the extra codes to the primary cohort; none occurred wholly within the primary cohort.

The source article further reports 55 consecutive recordings after five familiarisation strokes, stored in a large file and manually segmented. Only 50 trials per code–action cell are public. The selection or exclusion rule for the remaining five trials cannot be verified, so results apply to the public 50, not to the complete recorded sequence [1].

All 50 racket files for code 13–FP contained 400 rows in two element-wise identical 200-row blocks, so only the first block was retained under a result-independent deterministic rule. The same pattern occurred in 50 body files for the quarantined code 31–BA.

All 180 primary code–action cells contained 50 files with consecutive global indices (Table_S_ArchiveOrderAudit.csv). Within-cell archive ranks 1–50 were defined by ascending global index. These are not independently supplied trial numbers and are not verified acquisition order. Associations with rank are therefore archive-rank proxy diagnostics and cannot be interpreted as fatigue, learning or true sequential-trial dependence.

Because no code-to-participant crosswalk was available, three alternative scenarios were evaluated: the equally wide contiguous window 11–40, all codes 1–40, and all codes after removal of the codes 31–40 member of each of eight exact duplicate pairs. The de-duplicated scenario used unbalanced profile REML; other scenarios retained 50 trials per code–action cell.

Supplementary Table S2 | Sensitivity to the cohort window, all 40 codes and removal of exact duplicates

Code scenario	Representation	L^2 trace-ratio integer (FA–BP)	Phase-averaged pointwise integer (FA–BP)
Codes 1–30 primary	Racket	6, 7, 8, 17, 27, 10	7, 8, 9, 16, 17, 11
Codes 1–30 primary	Aggregated body	6, 6, 4, 7, 8, 4	6, 6, 4, 7, 8, 5
Codes 11–40	Racket	10, 8, 9, 20, 27, 12	11, 9, 10, 19, 20, 11
Codes 11–40	Aggregated body	7, 6, 4, 7, 9, 4	8, 6, 5, 8, 9, 5
Codes 1–40	Racket	7, 8, 9, 18, 25, 10	8, 8, 9, 17, 17, 11
Codes 1–40	Aggregated body	6, 6, 4, 7, 9, 4	6, 6, 4, 8, 8, 4
Codes 1–40, de-duplicated	Racket	7, 8, 9, 18, 25, 10	8, 8, 9, 17, 17, 11
Codes 1–40, de-duplicated	Aggregated body	6, 6, 4, 7, 9, 4	6, 6, 4, 8, 8, 4

Across the 12 action–representation conditions, the codes 11–40 window changed the L^2 threshold by 0–4 trials relative to the primary cohort. Exact de-duplication and the unmodified 40-code analysis were identical at the integer level under both functional summaries. The BD racket L^2 thresholds were 27, 27, 25

and 25. This reduces concern that the eight duplicates drive the main conclusion, but cannot prove that codes 1–30 correspond to the source participants or that all 40 codes are confirmed independent participants. Complete outputs are in Table_S_CohortAndDedupSensitivity.csv, Table_S_ExactDuplicateRemovalAudit.csv and Table_S_CohortBootstrapThresholds.csv.

S3. The 200-row boundary, structural zeros and waveform construction

Nominal filename lengths ranged from 104 to 313. In the primary cohort, 1,180/9,000 trials (13.1%) exceeded 200 frames; percentages for FA, FD, FP, BA, BD and BP were 13.7%, 21.1%, 20.3%, 3.9%, 9.5% and 10.2%. The public archive cannot distinguish downsampling, resampling, clipping or another transformation, so 200 rows is termed an archive storage/transformation boundary, not a verified hard truncation.

The body observable intervals contained 9,594 exact (0, 0, 0) triplets across 190 trials; counts by action were 1, 1, 17, 47, 109 and 15. Of the triplets, 99.6% occurred in the eighth label block, identified as RightForeArm by the source figure. No finer marker-centre definition was available, so this pattern was not assigned to a confirmed anatomical failure. No all-zero racket triplets were observed. Under the primary rule, exact (0, 0, 0) values were structural-zero or missingness candidates. A landmark contributed to a frame interval only when both endpoints were non-zero; at least 10 of 14 landmarks remained available for every interval. Raw coordinates were not imputed, trials were not removed by outcome magnitude and structural zeros were not treated as true stationarity.

The 42 unlabelled body columns were operationally mapped to 14 consecutive three-coordinate blocks in the order shown in Fig. 1 of the source paper. Axis names and directions were not independently documented.

Supplementary Table S3 | Operational mapping of the 14 body labels

No.	Source label	Operational anatomical description	CSV columns (1-based)
1	Hips	pelvic-region label	1–3
2	Head	head label	4–6
3	LeftShoulder	left shoulder	7–9
4	LeftArm	left upper arm	10–12
5	LeftForeArm	left forearm	13–15
6	RightShoulder	right shoulder	16–18
7	RightArm	right upper arm	19–21
8	RightForeArm	right forearm	22–24
9	LeftUpLeg	left thigh	25–27
10	LeftLeg	left lower leg	28–30
11	LeftFoot	left foot	31–33

No.	Source label	Operational anatomical description	CSV columns (1-based)
12	RightUpLeg	right thigh	34–36
13	RightLeg	right lower leg	37–39
14	RightFoot	right foot	40–42

Because Euclidean frame-interval magnitude is invariant to coordinate-axis permutation or reversal, this mapping is sufficient for that scalar measure, but it cannot recover an undisclosed axis orientation or anatomical marker definition. The body aggregate equally weights available labels because no validated segment mass, inertia or external biomechanical weighting standard was available. Magnitude averaging prevents direct cancellation of signed coordinates but smooths local differences and is not a centre of mass, joint angle, coordination, torque or power measure.

Racket and aggregated-body displacement magnitudes were defined as

$$d^{R}_{p,a,r}(t) = \left\| \mathbf{x}^{R}_{p,a,r}(t+1) - \mathbf{x}^{R}_{p,a,r}(t) \right\|_2 \tag{S1}$$

$$d^{H}_{p,a,r}(t) = |V_{p,a,r,t}|^{-1} \sum_{j \in V_{p,a,r,t}} \left\| \mathbf{x}_{j,p,a,r}(t+1) - \mathbf{x}_{j,p,a,r}(t) \right\|_2 \tag{S2}$$

Here p, a, r and t denote archive code, fixed action label, within-cell trial and valid frame interval, and $V_{par,t}$ contains landmarks that were non-zero at both endpoints. Each waveform was linearly mapped to 200 equally spaced nodes over 0–100%. The racket waveform discards direction, absolute position and racket-face orientation; the body waveform is not a centre of mass, joint configuration, coordination, torque or power.

Structural-zero checks gave body $n_{0\cdot90}$ values of 5, 6, 4, 7, 9 and 4 after complete deletion of 190 affected trials; 5, 6, 3, 6, 8 and 4 using a fixed 13-landmark set after removal of the eighth landmark and four remaining affected trials; and 6, 6, 3, 6, 8 and 4 using the eight landmarks visible in all 9,000 trials (labels 1–6, 11 and 14). The latter analyses show that the body pattern was not caused solely by a variable denominator. Some complete-case BA and BD cells retained only four and three trials, so those results are directional checks only.

S4. Finite-set contrast dispersion for six fixed action levels

The six action labels were fixed conditions, not a random sample from all table-tennis actions [3]. The across-action model supplied only a variance background for the current label set:

$$Y_{p,a,r}(u) = \mu(u) + \alpha_a(u) + P_p(u) + (P\alpha)_{p,a}(u) + \varepsilon_{p,a,r}(u), \qquad \sum_{a=1}^{A} \alpha_a(u) = 0 \tag{S3}$$

The finite-set action-contrast dispersion was

$$D(A)(u) = \frac{1}{A-1}\sum_{a=1}^{A}\alpha_a^2(u) \tag{S4}$$

with the non-negative, noise-corrected balanced-design estimator

$$\widehat{D}(A)(u) = \max\left\{\frac{MS(A)(u) - MS(P\times A)(u)}{PR},\ 0\right\} \tag{S5}$$

This is not a random action variance sampled from an action population. The archive-code main effect, archive-code × fixed-action component and within-trial component were estimated from the same balanced mean-square system. Integrated proportions describe only these six labels, 30 archive codes and representations. Racket proportions were 31.2%, 1.5%, 29.0% and 38.3% for finite-set action contrasts, across-action archive code, archive code × fixed action and within-trial dispersion; body proportions were 50.0%, 4.4%, 27.1% and 18.5%. The supported description is that, within the included labels, cross-action-consistent archive-code differences were small and relative code differences were more strongly action dependent. These results do not establish random action variance or a stable personal fingerprint.

S5. Probability definition of the L^2 trace-type relative-reliability coefficient

For fixed action a, a single waveform was treated as a square-integrable random element in L^2 over the unit interval:

$$Y_{p,r}^{(a)}(u) = \mu_a(u) + P_p^{(a)}(u) + \varepsilon_{p,r}^{(a)}(u) \tag{S6}$$

$P_p^{(a)}$ and $\varepsilon_{pr}^{(a)}$ were assumed mean zero, second-order integrable and mutually independent, with covariance operators C(P,a) and C(ε,a) [4,5]. Define

$$B_a = \mathrm{tr}\{C_{P,a}\} = \mathbb{E}\left\|P_p^{(a)}\right\|_{L^2}^2 = \int_0^1 \sigma_{P,a}^2(u)\,\mathrm{d}u \tag{S7}$$

$$W_a = \mathrm{tr}\{C_{\varepsilon,a}\} = \mathbb{E}\left\|\varepsilon_{p,r}^{(a)}\right\|_{L^2}^2 = \int_0^1 \sigma_{\varepsilon,a}^2(u)\,\mathrm{d}u \tag{S8}$$

If trial residuals within a code–action cell are identically distributed with common marginal covariance C(ε,a) and zero or negligible cross-trial covariance, the error covariance of the mean of n trials is C(ε,a)/n. We then define

$$R_{L^2,a}(n) = \frac{B_a}{B_a + W_a/n} \tag{S9}$$

For a target reference g under negligible residual dependence,

$$n_{a,g}^{*} = \frac{gW_a}{(1-g)B_a}, \qquad n_{a,g} = \left\lceil n_{a,g}^{*}\right\rceil \tag{S10}$$

At g = 0.90, $n^*_{a,0.90} = 9W_a/B_a$. A larger $n_{0.90}$ may reflect larger W_a, smaller B_a or both, and does not independently establish that an action is more variable or difficult.

Supplementary Table S4 | Action-specific integrated variance components and 0.90 trial-count thresholds

Representation	Action	B_a	W_a	W_a/B_a	Continuous $n^*_{a,0.90}$	Integer $n_{0\cdot90}$
Racket	FA	56.546	36.740	0.650	5.848	6
Racket	FD	183.268	138.580	0.756	6.805	7
Racket	FP	17.479	15.027	0.860	7.737	8
Racket	BA	26.309	48.484	1.843	16.586	17
Racket	BD	72.758	212.798	2.925	26.323	27
Racket	BP	16.410	18.154	1.106	9.956	10
Aggregated body	FA	1.154	0.665	0.576	5.187	6
Aggregated body	FD	4.687	2.824	0.602	5.422	6
Aggregated body	FP	1.432	0.525	0.367	3.299	4
Aggregated body	BA	0.648	0.458	0.707	6.364	7
Aggregated body	BD	1.410	1.239	0.878	7.906	8
Aggregated body	BP	1.101	0.426	0.387	3.481	4

This table reports point estimates. Percentile intervals for continuous thresholds are in main-text Table 2 and Supplementary Table S8 and must accompany interpretation of these values.

As a functional-summary sensitivity, pointwise relative reliability $R_a(u;n) = \sigma^2P_{,a}(u)/[\sigma^2P_{,a}(u) + \sigma^2e_{,a}(u)/n]$ was calculated before trapezoidal phase averaging and integer search. Racket thresholds were 7, 8, 9, 16, 17 and 11; body thresholds were 6, 6, 4, 7, 8 and 5. Most conditions differed from the trace ratio by 0–1 trial, but the BD racket changed from 27 to 17. The L^2 trace ratio integrates variances separately before taking their ratio, thereby weighting higher-variance phases more heavily; the phase average weights phase intervals equally. These are different functional estimands, and 17 is not a correction to 27. Full results are in Table_S_FunctionalAggregationComparison.csv.

The coefficient is invariant to orthogonal basis changes in L^2 and common scalar unit transformations, but generally not to non-isometric linear transformations, time warping, temporal weighting, replacement of the representation or quality rules [4,5]. It adopts the repeated-trial error-reduction logic of a generalizability-theory decision study [3], but is not a complete standard G coefficient with all random facets and interactions and is not termed a ‘functional G coefficient’.

S6. Node estimation, numerical integration and self-checks

A non-negative constrained random-intercept REML model was fitted at each of 200 nodes for each fixed action and waveform, giving 12 models [6–9]. All nodes in the complete 30 × 50 balanced design had interior solutions. Archive-code and trial variances were trapezoidally integrated over normalised time before $R(L^2,a;n)$ was calculated; this is not an arithmetic mean of 200 nodewise reliability values.

All 13 numerical self-checks passed. The sufficient-statistic REML objective for synthetic unbalanced data differed from a dense-matrix implementation by at most 1.78×10^{-14}; analytical constrained solutions for the 12 actual balanced models differed from profile REML by a maximum relative error below 7.2×10^{-7}. The cold-reproduction environment was Python 3.12.13, NumPy 2.3.5, pandas 2.2.3, SciPy 1.18.0, Matplotlib 3.11.0 and Pillow 12.2.0. Seeds were 20260712 for the primary analysis and 20260713 for added assumption and influence diagnostics. Release v1.0.2 fixes the final environment.

S7. Archive-order proxy ACF and AR(1) residual-dependence scenarios

S7.1. Archive-order proxy diagnostic

Each waveform was centred by its archive-code–action cell mean, and residuals were ordered by archive rank 1–50 as defined by global file index. The L^2 trace-type lag-k correlation proxy divided the integrated cross-products of node residuals by the corresponding zero-lag square-integral terms. Five thousand within-cell random permutations established a lag-1 reference distribution.

All 12 lag-1 proxies were positive (0.075–0.328). For each condition, the unadjusted two-sided Monte Carlo p-value from 5,000 permutations was $1/5001 \approx 0.0002$. This indicates that residual similarity is associated with numeric archive order. It does not verify acquisition order and cannot establish temporal autocorrelation, fatigue or learning; it is not a multiplicity-adjusted confirmatory test. Lags 1–10 are in Table_S_ArchiveOrderResidualACF.csv.

Supplementary Table S5 | Lag-1 L^2 trace-type correlation proxy calculated by archive rank

Representation	FA	FD	FP	BA	BD	BP
Racket	0.141	0.137	0.286	0.114	0.075	0.208
Aggregated body	0.194	0.138	0.328	0.232	0.136	0.282

S7.2. AR(1) design scenarios

As a design-effect scenario, we assumed equally spaced adjacent archive ranks, weak stationarity of the marginal residual trace and $\mathrm{trCov}(\varepsilon_r, \varepsilon_{r+k}) = \varphi^k W_a$, equivalently $\rho_k = \varphi^k$ for the normalised scalar lag association [10,11]. This is neither a nodewise AR(1) model nor an AR(1) fit to the complete covariance operator. Under this working assumption, the error trace of an n-trial mean is $(W_a/n)D(n,\varphi)$, with finite-sample design effect

$$D(n,\phi) = 1 + 2\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)\phi^k \tag{S11}$$

and relative reliability

$$R_{L^2,a}^{(\phi)}(n) = \frac{B_a}{B_a + \frac{W_a}{n} D(n,\phi)} \quad \text{(S12)}$$

For φ = 0, 0.10, 0.20 and 0.30, we searched the positive integers for the smallest n satisfying $R_{L^2,a}^{(\varphi)}(n) \geq 0.90$. The scenario retained the primary B_a and W_a and changed only the rate at which mean error decreased. It was not fitted to a verified time series and does not quantify how ignoring correlation could shift the variance-component estimates themselves.

Supplementary Table S6 | Minimum integer trial count reaching 0.90 under AR(1) working scenarios

Representation	Action	φ = 0	φ = 0.10	φ = 0.20	φ = 0.30
Racket	FA	6	7	9	11
Racket	FD	7	9	10	12
Racket	FP	8	10	12	14
Racket	BA	17	21	25	31
Racket	BD	27	32	40	49
Racket	BP	10	12	15	18
Aggregated body	FA	6	7	8	9
Aggregated body	FD	6	7	8	10
Aggregated body	FP	4	4	5	6
Aggregated body	BA	7	8	10	12
Aggregated body	BD	8	10	12	14
Aggregated body	BP	4	5	5	6

The results must be read in two layers. The primary scenario assumes negligible residual dependence and gives 4–8 body trials and 6–27 racket trials. The second layer retains B_a and W_a from the first and alters only the error-reduction term under a condition-specific positive-dependence scenario. φ was not fitted from verified acquisition order, and the variance components were not re-estimated under correlated residuals. The values must not be combined into a general 6–49 trial range. Neither 6–27 nor 49 for BD at φ = 0.30 is a field prescription, empirical upper bound, safety margin or recommendation.

S8. Single proxy-peak registration sensitivity

The primary W_a combines amplitude variation, phase-proportion differences, peak-timing displacement, segmentation-start differences and coordinate error. To examine temporal displacement, a transparent main-peak registration was added as a sensitivity analysis, not as ground truth. Within the central 10–90% of each racket waveform, the main peak was found on an 11-node-smoothed Hampel-type QC copy. The reference for an action was the median peak node across all trials. Endpoint-preserving piecewise-linear time warping mapped each peak to the reference and was applied to the original racket waveform without Hampel replacement and its paired body waveform. This protects landmark search from a single high jump

while keeping quality replacement out of the registered result. Registration changes the estimand and does not demonstrate alignment to actual ball impact.

Supplementary Table S7 | Trial-count thresholds before and after single proxy-peak registration

Representation	Action	Unregistered continuous n*	Unregistered integer	Registered continuous n*	Registered integer
Racket	FA	5.848	6	3.785	4
Racket	FD	6.805	7	4.017	5
Racket	FP	7.737	8	3.743	4
Racket	BA	16.586	17	5.539	6
Racket	BD	26.323	27	30.669	31
Racket	BP	9.956	10	3.947	4
Aggregated body	FA	5.187	6	5.237	6
Aggregated body	FD	5.422	6	3.268	4
Aggregated body	FP	3.299	4	2.661	3
Aggregated body	BA	6.364	7	3.742	4
Aggregated body	BD	7.906	8	4.344	5
Aggregated body	BP	3.481	4	2.705	3

Most racket thresholds decreased, but BD increased from 27 to 31. The body FA result was almost unchanged and most others decreased. Amplitude–phase separation is therefore not negligible, and no registration rule can be proven uniquely correct from this archive [12–14]. Landmark distributions are in Table_S_PeakRegistrationLandmarks.csv.

S9. Hampel-type local high-value rule and BD archive-code influence

The Hampel-type rule was a sensitivity analysis, not an instrument truth. Parameters were fixed before the run and were not tuned by action. The rule operated on derived frame-interval displacement before phase normalisation. For an interior node, at most five neighbouring points on each side, excluding the centre, formed the reference set. With m = median(xref) and s = 1.4826 median|xref − m|, x_i was flagged only if x_i > max(m + 8s, 4m), then interpolated linearly from the nearest unflagged displacement values [15–17]. The rule flagged 423/1,647,435 racket values (0.0257%; 379 trials) and 11/1,647,435 aggregated-body values (0.000668%; seven trials).

The BD racket accounted for 309 flags across 292/1,500 trials and 27/30 archive codes, or 73.0% of all racket flags. Codes 13, 18, 29 and 3 contained 38, 37, 30 and 27 flags, affecting 68%, 62%, 56% and 48% of their BD trials. Codes are used only for influence diagnosis, not interpretation of individual technique.

For the BD racket under the primary rule, B = 72.758, W = 212.798, W/B = 2.925, continuous $n^*_{0\cdot90}$ = 26.323 and integer n_{90} = 27. Under the Hampel-type rule, B = 67.456, W = 134.625, W/B = 1.996,

continuous $n^*_{0.90}$ = 17.961 and integer n_{90} = 18; the 95% percentile interval from 5,000 archive-code cluster bootstrap replicates was 12.87–29.01. Most of the change arose from lower W.

After leaving out one archive code at a time, BD racket n_{90} was 24–29 under the primary rule and 17–20 under the Hampel-type rule; continuous thresholds were 23.379–28.788 and 16.659–19.593. No single omission reconciled the two rules, so quality-rule dependence is itself a primary finding. Machine-readable archive-code variance contributions, BD influence and flags, and leave-one-archive-code-out results are provided in the repository.

S10. Cluster bootstrap, boundary solutions and Monte Carlo diagnostics

Each bootstrap replicate sampled 30 complete archive-code clusters with replacement and retained all six actions and trials; 5,000 replicates were used [18,19]. Replicates contained a mean of 19.135 distinct archive codes, with minimum 13, 2.5th percentile 16, median 19, 97.5th percentile 22 and maximum 26. Increasing the number of resamples does not increase cross-code information.

Across all 12 × 5,000 = 60,000 action-specific fits, B ≤ 0 and undefined continuous $n^*_{a},0.90$ each occurred zero times. Only two FD racket replicates contained an archive-code boundary estimate at one or more nodes.

Primary reporting used percentile intervals for continuous $n^*_{a},0.90$ [18–22]. Basic bootstrap and BCa intervals, kernel-density quantile Monte Carlo standard-error approximations, and quantile stability across five consecutive 1,000-replicate blocks within the same random stream were also calculated. The five blocks are not independent runs.

Supplementary Table S8 | Archive-code cluster bootstrap diagnostics for the BD racket waveform

Metric	Result
Continuous $n^*_{a},0.90$ point estimate	26.323
Percentile 95% interval	17.647–43.477
Basic bootstrap 95% interval	9.168–34.998
BCa 95% interval	16.474–40.037
Monte Carlo SE of 2.5th percentile	0.135
Monte Carlo SE of 97.5th percentile	0.427
Range of 2.5th percentiles across five 1,000-replicate blocks	17.365–18.178
Range of 97.5th percentiles across five 1,000-replicate blocks	41.686–44.649
Maximum single-replicate continuous threshold	65.234

Differences among interval methods reflect skewness of the threshold distribution with N = 30. Interpretation should jointly consider the point estimate, continuous percentile interval, leave-one-archive-

code-out range and the endpoints under the two pre-specified quality rules. Complete results for all 12 conditions are in Table_S_BootstrapDiagnostics.csv.

Continuous threshold differences for all 15 action pairs were also calculated within each paired archive-code cluster bootstrap replicate. Intervals were not multiplicity adjusted and are descriptive, not evidence that all six actions differ statistically. Results are in Table_S_PairedActionBootstrapContrasts.csv.

S11. Pointwise marginal bootstrap intervals

For each representation and each of the four across-action nodewise curves (finite-set action contrast, across-action archive code, archive code × fixed action and within-cell residual), 5,000 archive-code cluster bootstrap replicates provided 2.5%, 50% and 97.5% pointwise quantiles. Table_S_GlobalPointwiseBootstrapBands.csv contains two representations × four curves × 200 nodes.

These are pointwise marginal intervals, not simultaneous confidence bands or local significance tests [9,18,19]. An 11-node moving average is used for display only. All integrals and $R(L^2)$ values use unsmoothed node estimates. Local interval overlap must not be used to claim significance between nodes or actions.

S12. Other archive-boundary and balanced-resampling analyses

Excluding 1,180 trials whose nominal length exceeded 200 left 7,820 trials: 1,295 FA, 1,183 FD, 1,196 FP, 1,441 BA, 1,358 BD and 1,347 BP. Minimum cell sizes were 29, 18, 26, 39, 35 and 20. Non-negative unbalanced profile REML completed for every condition. Racket n_{90} values were 5, 6, 7, 16, 23 and 10; body values were 5, 5, 3, 7, 7 and 4. The largest absolute change in $R(L^2,a;50)$ was 0.0075 for the BD racket.

Within this exclusion set, 1,000 balanced samples without replacement used the minimum cell size for each action and shared the same trial indices between the two waveforms. The largest absolute difference between the median resampled $R(L^2,a;50)$ and the unbalanced REML estimate was 0.0021. This resampling diagnoses imbalance; it is not a population bootstrap and its range is not a population confidence interval.

S13. Machine-readable results, code and the reproducibility chain

The permanent repository contains: the input archive name, size, checksums, inclusion and quarantine rules, and file manifest; all scripts from the third-party archive to analysis arrays and trial-level QC audits; finite-set decomposition, 12 action-specific REML models, cluster bootstrap, code-window and exact-de-duplication analyses, functional-summary comparison, AR(1) scenarios, registration, leave-one-archive-code-out influence and pointwise marginal-band scripts; tidy CSV files supporting manuscript values, figure source data and random seeds; a one-command reproduction guide, locked dependencies, logs, self-

checks and release hash; Python code for all main and supplementary figures and SVG/PDF/PNG/600-dpi TIFF exports; and a permanent version, licences and DOI. Software sources are cited in references [23–26].

Third-party TTMD6 coordinates are accessed only from the original public source and are not redistributed. The frozen software and derived results are available at https://github.com/hmk-228674/TTMD6-functional-G-theory-analysis, tag v1.0.2; version DOI https://doi.org/10.5281/zenodo.21599397; concept DOI https://doi.org/10.5281/zenodo.21382966; release/tag commit 28e143e03a7e51d1797e30b9b4bc1a0d2f5b3040. Code is released under the MIT License; author-generated documentation, figures, figure source data and derived tables are CC BY 4.0. Releases v1.0.0 and v1.0.1 remain immutable. Release v1.0.2 does not contain the original TTMD6 RAR, and the frozen software and original data source should be cited separately.

Supplementary references

1. Zhang, J., Ren, Y., Lin, L., Xing, Y. & Ren, J. Table tennis motion recognition based on the bat trajectory using

varying-length-input convolution neural networks. Sci. Rep. 14, 3549 (2024). https://doi.org/10.1038/s41598-024-54150-5

2. Zhang, J. Table Tennis MOCAP Dataset, version 2. figshare (2026). https://doi.org/10.6084/m9.figshare.31746358.v2

3. Brennan, R. L. Generalizability Theory (Springer, 2001). https://doi.org/10.1007/978-1-4757-3456-0

4. Hsing, T. & Eubank, R. Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators (Wiley, 2015). https://doi.org/10.1002/9781118762547

5. Di, C.-Z., Crainiceanu, C. M., Caffo, B. S. & Punjabi, N. M. Multilevel functional principal component analysis. Ann. Appl. Stat. 3, 458–488 (2009). https://doi.org/10.1214/08-AOAS206

6. Patterson, H. D. & Thompson, R. Recovery of inter-block information when block sizes are unequal. Biometrika 58, 545–554 (1971). https://doi.org/10.1093/biomet/58.3.545

7. Harville, D. A. Maximum likelihood approaches to variance component estimation and to related problems. J. Am. Stat. Assoc. 72, 320–338 (1977). https://doi.org/10.1080/01621459.1977.10480998

8. Self, S. G. & Liang, K.-Y. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J. Am. Stat. Assoc. 82, 605–610 (1987). https://doi.org/10.1080/01621459.1987.10478472

9. Morris, J. S. & Carroll, R. J. Wavelet-based functional mixed models. J. R. Stat. Soc. B 68, 179–199 (2006). https://doi.org/10.1111/j.1467-9868.2006.00539.x

10. Brockwell, P. J. & Davis, R. A. Introduction to Time Series and Forecasting 3rd edn (Springer, 2016). https://doi.org/10.1007/978-3-319-29854-2

11. Pinheiro, J. C. & Bates, D. M. Mixed-Effects Models in S and S-PLUS (Springer, 2000). https://doi.org/10.1007/b98882

12. Helwig, N. E., Hong, S., Hsiao-Wecksler, E. T. & Polk, J. D. Methods to temporally align gait cycle data. J. Biomech. 44, 561–566 (2011). https://doi.org/10.1016/j.jbiomech.2010.09.015

13. Marron, J. S., Ramsay, J. O., Sangalli, L. M. & Srivastava, A. Functional data analysis of amplitude and phase variation. Stat. Sci. 30, 468–484 (2015). https://doi.org/10.1214/15-STS524

14. Ramsay, J. O. & Li, X. Curve registration. J. R. Stat. Soc. B 60, 351–363 (1998). https://doi.org/10.1111/1467-9868.00129

15. Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. & Stahel, W. A. Robust Statistics: The Approach Based on Influence Functions (Wiley, 2005). https://doi.org/10.1002/9781118186435

16. Rousseeuw, P. J. & Croux, C. Alternatives to the median absolute deviation. J. Am. Stat. Assoc. 88, 1273–1283 (1993). https://doi.org/10.1080/01621459.1993.10476408

17. Pearson, R. K. Mining Imperfect Data: Dealing with Contamination and Incomplete Records (SIAM, 2005). https://doi.org/10.1137/1.9780898717884

18. Efron, B. & Tibshirani, R. J. An Introduction to the Bootstrap (Chapman & Hall, 1993).

19. Field, C. A. & Welsh, A. H. Bootstrapping clustered data. J. R. Stat. Soc. B 69, 369–390 (2007). https://doi.org/10.1111/j.1467-9868.2007.00593.x

20. Efron, B. Better bootstrap confidence intervals. J. Am. Stat. Assoc. 82, 171–185 (1987). https://doi.org/10.1080/01621459.1987.10478410

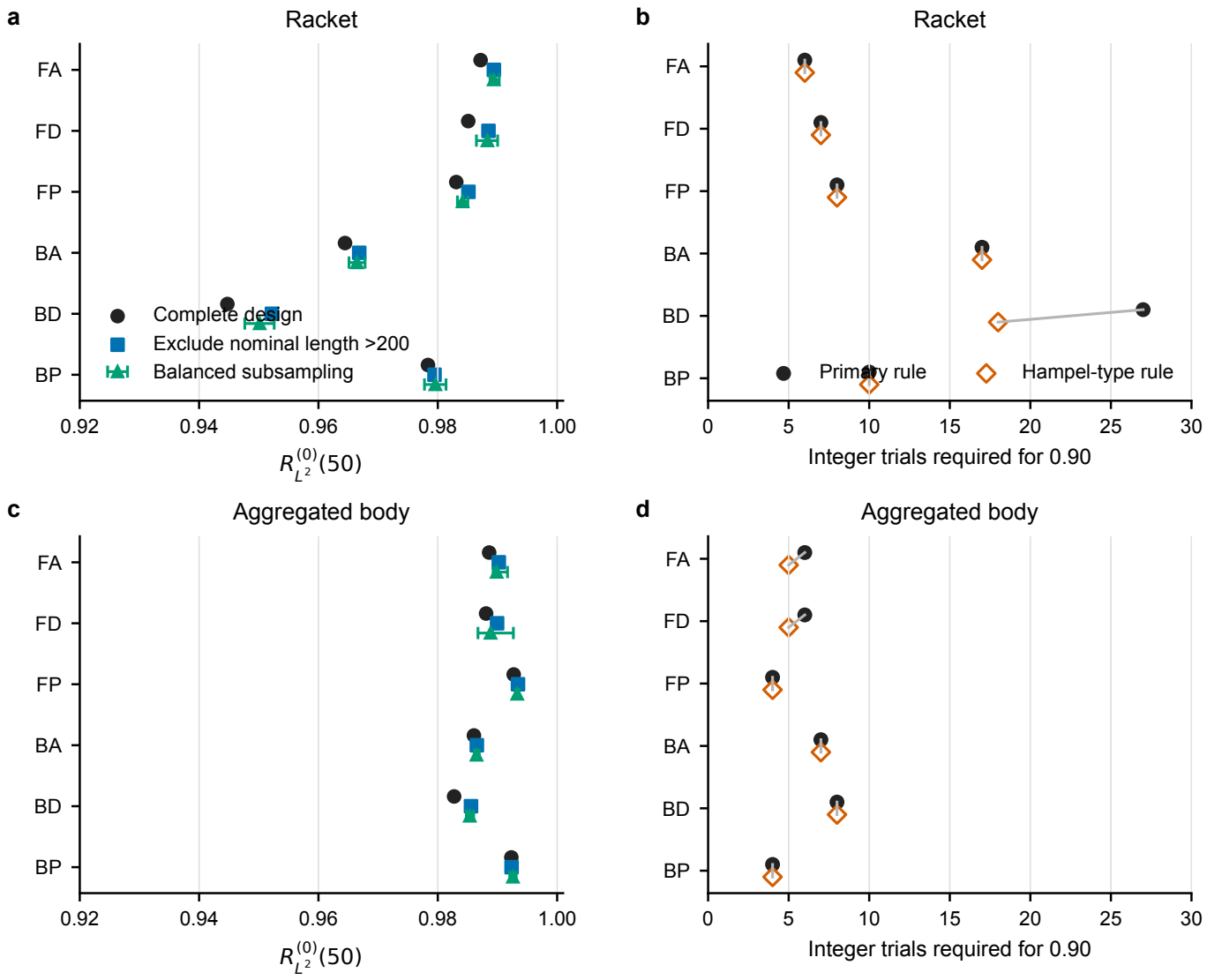
21. Ukoumunne, O. C., Davison, A. C., Gulliford, M. C. & Chinn, S. Non-parametric bootstrap confidence intervals for the intraclass correlation coefficient. Stat. Med. 22, 3805–3821 (2003). https://doi.org/10.1002/sim.1643

22. Andrews, D. W. K. & Buchinsky, M. A three-step method for choosing the number of bootstrap repetitions. Econometrica 68, 23–51 (2000). https://doi.org/10.1111/1468-0262.00092

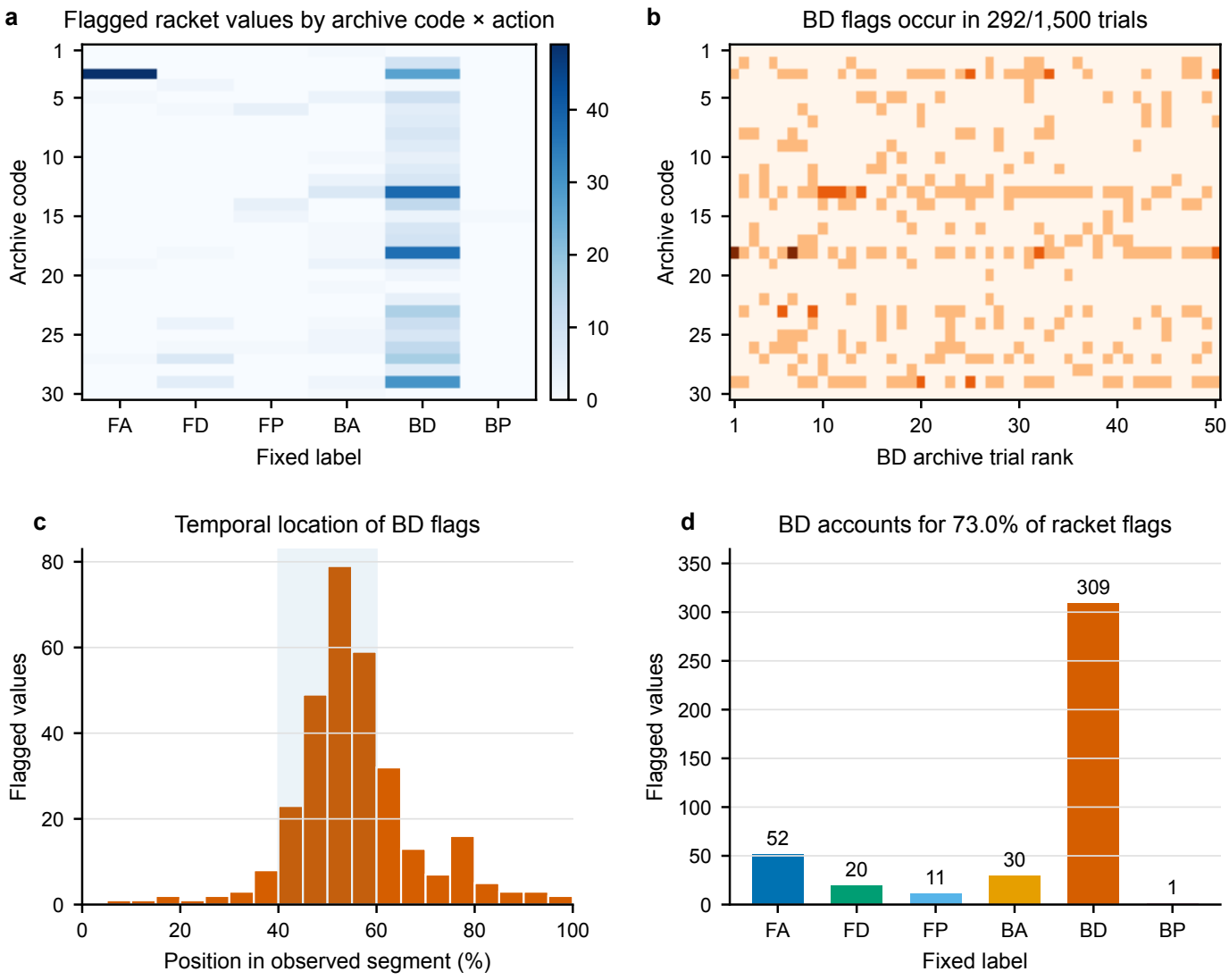
23. McKinney, W. Data structures for statistical computing in Python. In Proc. 9th Python in Science Conference 56–61 (2010). https://doi.org/10.25080/Majora-92bf1922-00a

24. Hunter, J. D. Matplotlib: a 2D graphics environment. Comput. Sci. Eng. 9, 90–95 (2007). https://doi.org/10.1109/MCSE.2007.55

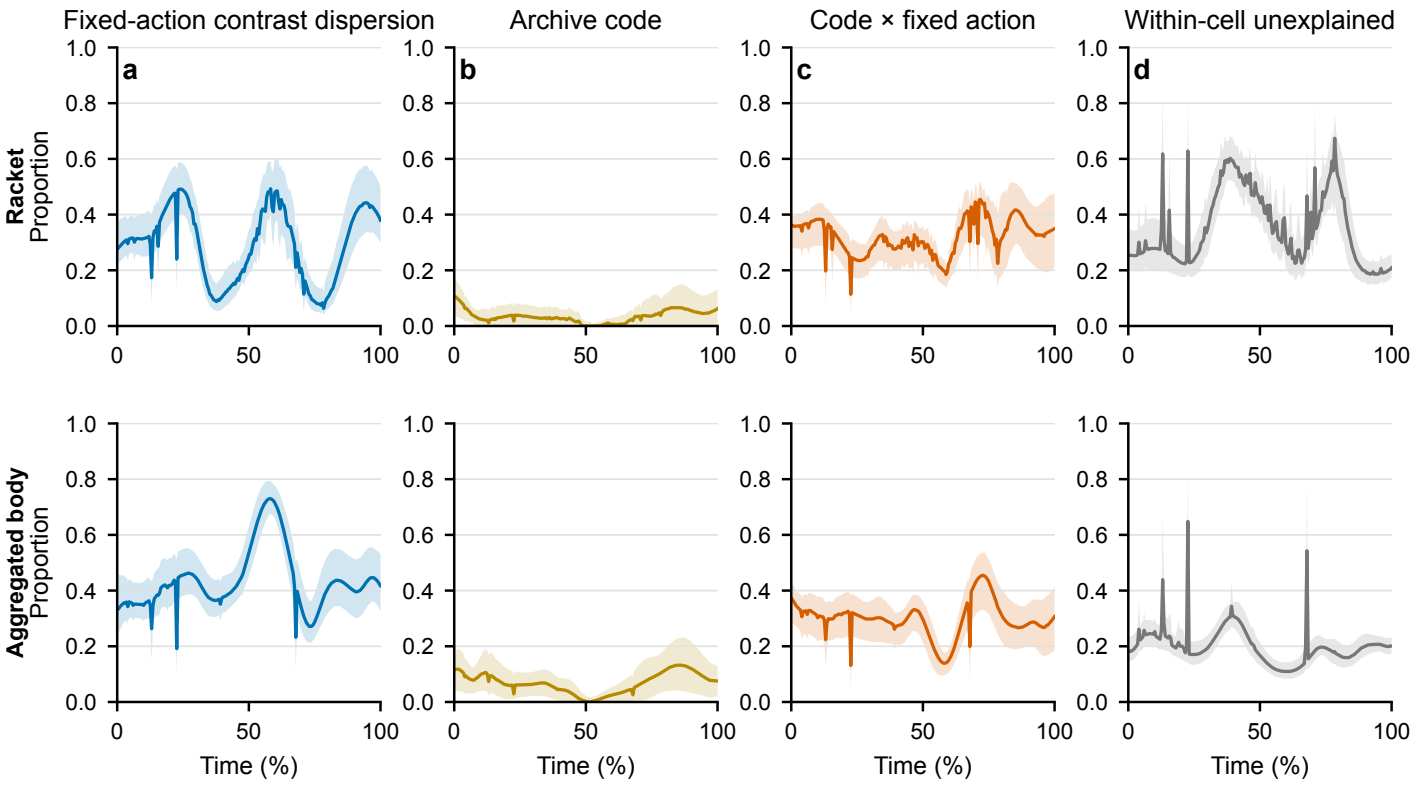
348 25. Harris, C. R. et al. Array programming with NumPy. Nature 585, 357–362 (2020).
349 https://doi.org/10.1038/s41586-020-2649-2
350 26. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods
351 17, 261–272 (2020). https://doi.org/10.1038/s41592-019-0686-2



Supplementary Fig. S1 | Archive boundary, balanced resampling and high-value-rule sensitivity. The left column compares the complete design, the unbalanced estimate after exclusion of nominal lengths above 200 frames and 1,000 action-wise balanced subsamples of $R(L^2,a;50)$. The right column compares integer counts reaching 0.90 under the primary and Hampel-type rules. Racket waveforms are shown above and aggregated-body waveforms below.



Supplementary Fig. S2 | Distribution of Hampel-type high-value flags in racket waveforms. a, Number of flagged values by archive code × fixed action. b, BD flags across 30 archive codes and 50 within-cell archive ranks. c, Temporal positions of BD flags within the original observable interval. d, Total flags under the six labels. BD contained 309/423 racket flags across 292/1,500 BD trials.



Supplementary Fig. S3 | Archive-code cluster bootstrap marginal intervals for pointwise dispersion proportions under the six fixed labels. Solid lines are original-sample pointwise estimates and shading denotes 2.5–97.5% marginal quantile intervals from 5,000 archive-code cluster bootstrap replicates. Nodewise intervals are not simultaneous confidence bands and cannot support local significance decisions.