Unsupervised methods for detection of neural states: case study of hippocampal-amygdala interactions

EXTENDED DATA

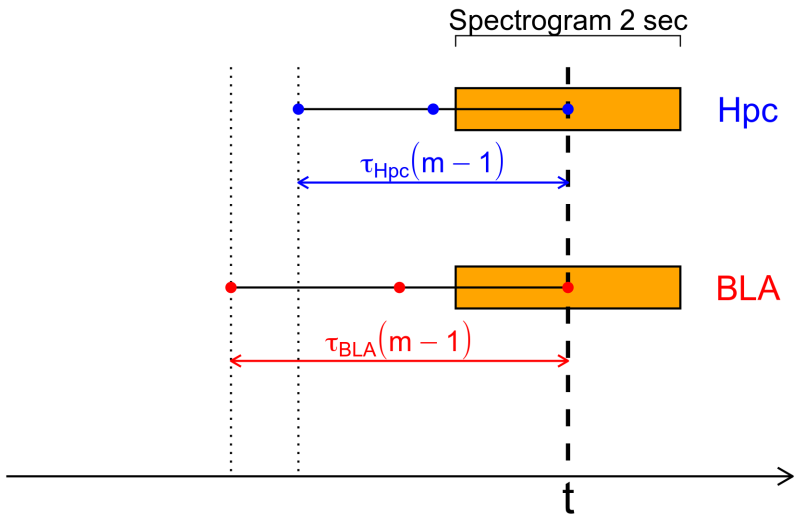

Figure 1-1: **Embedded time series and alignment settings.** We illustrate here the alignment settings between TC, spectrograms, and external labels described in Sec. 2.4. For a given point in time, the "cursor" t, we show the spectrogram windows (in orange), and two time-lagged embeddings with dimension $m = 3$. The time lag between points in the embedding, τ, is dependent on the underlying time series but m is homogenized to the larger of the two values suggested by the simplex projection method (Sec. 2.3). For the TC calculation, the two embedded time series are aligned to each other to the right. We assign to the computed TC value the time t corresponding to the middle of the spectrogram window (thick dotted line). Note that this last operation is relevant only when comparing TC to other time-series characterizing events or behavioral stages, *e.g.,* air puffs, SWRs, awake phase. The implications of these choices are analyzed in Sec. 2.4 and Figs. 3-3, 3-4.

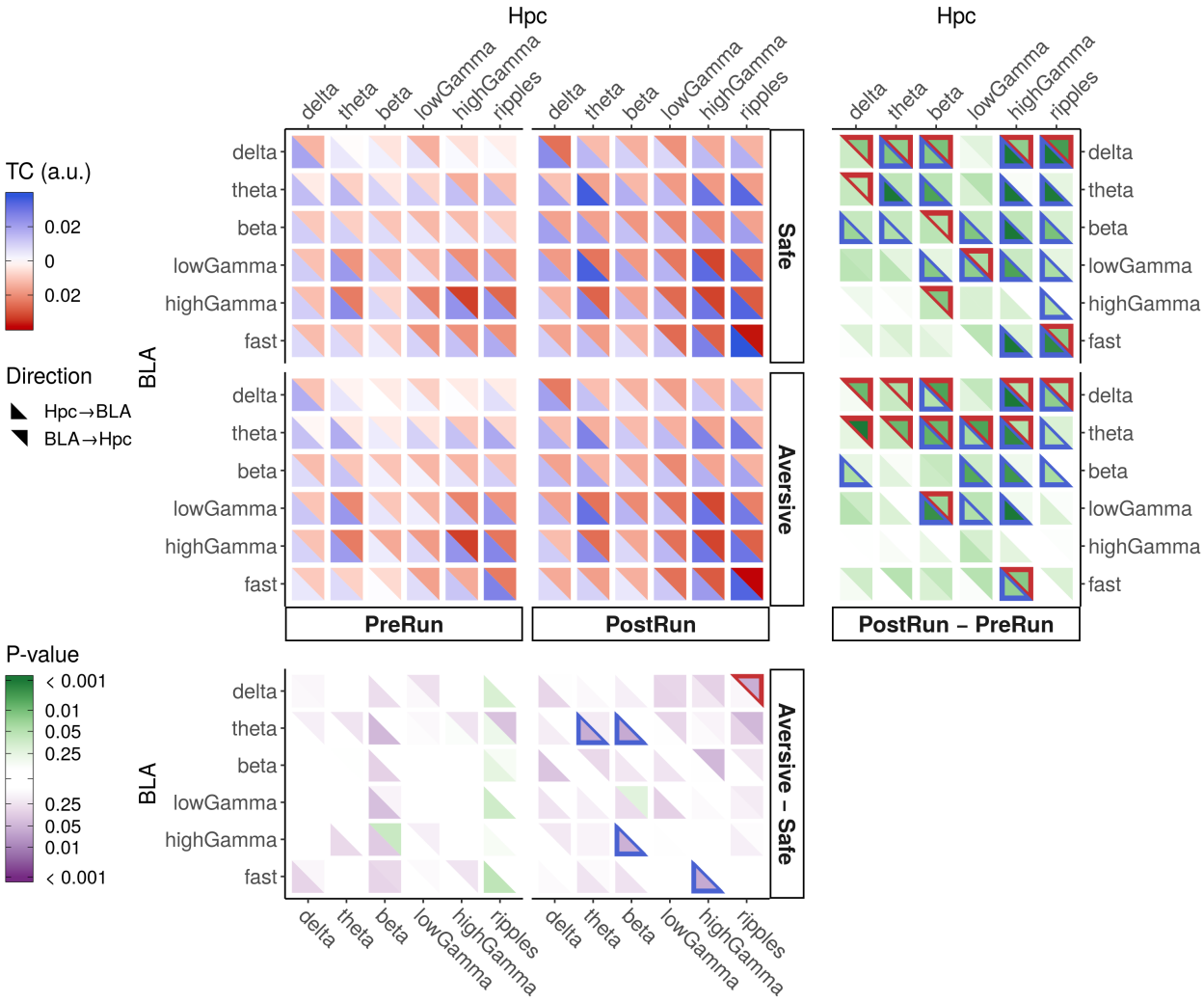

Figure 2-1: **The influence of the aversive stimulus on TC during post- and pre-run epochs.** Mean TC values (n = 17 sessions, 3 rats) were calculated separately for time points from two behavioral epochs, *viz.*, the pre- or post-run phases, and the two running directions. All sessions have at least 20 seconds of total time in each of the four different data subsets shown. The plot layout and interpretation are the same as in Fig. 2a (see main text for details). The characterization of a trajectory, as either safe or aversive, depends on the direction where the air puff was delivered in the run period that preceded the epoch under examination. This means that, for example, in a given pre-run phase, we consider the safe and aversive directions of the run epoch of the preceding session.

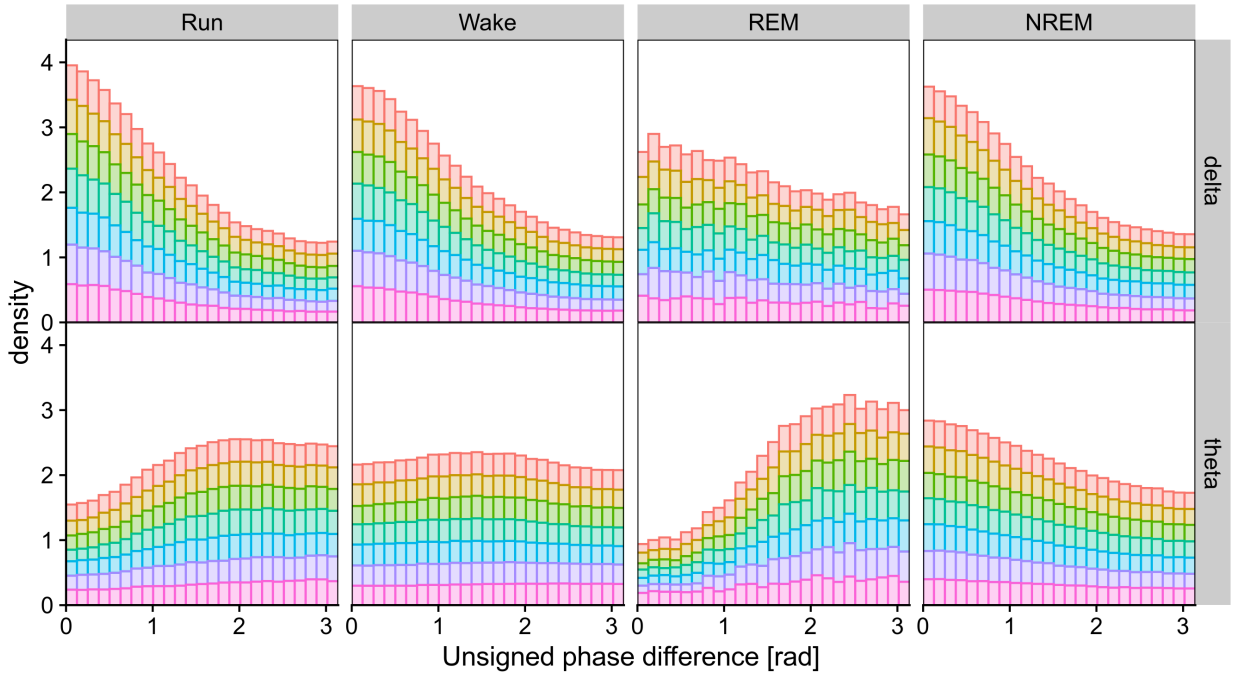

Figure 2-2: **Phase differences between hippocampus and BLA for delta and theta bands.** Phases were extracted from the continuous wavelet transform of the LFPs performed with analytic Morlet wavelets (*cwt.m* in Matlab toolbox 'Wavelet') in the delta (0.5-4 Hz) and theta (4-12 Hz) bands of both regions. Circular means across the electrodes in each anatomical area were computed, and unsigned phase differences were extracted from heterogenous time windows. 'Wake' refers to periods in pre- and post-sleep epochs where the animal is awake. The histograms show data accumulated from six sessions of Rat2. The high similarity across sessions (separate colors, stacked bars) supports that these data are robust.

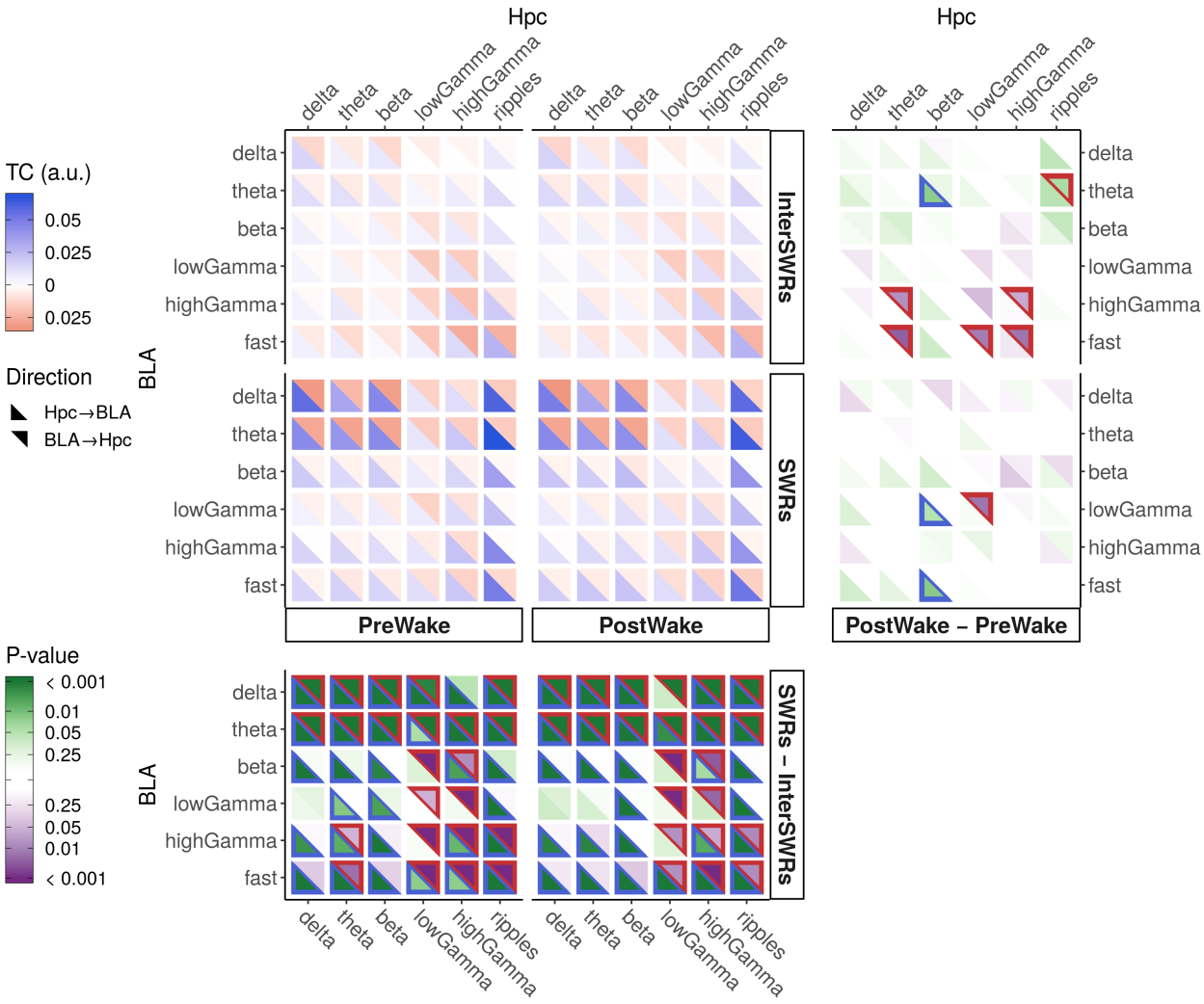

Figure 3-1: **TC analysis for awake phases.** From the data corresponding to awake phases, mean TC values (n = 20 sessions, 3 rats) were calculated separately for the four combinations of the two behavioral epochs, pre- or post-, and the presence or absence of sharp-wave ripples (SWRs). All sessions show at least 5 seconds of time featuring SWRs in both pre- and post-epochs. The plot layout and interpretation are the same as that in Fig. 2a (see main text for details).

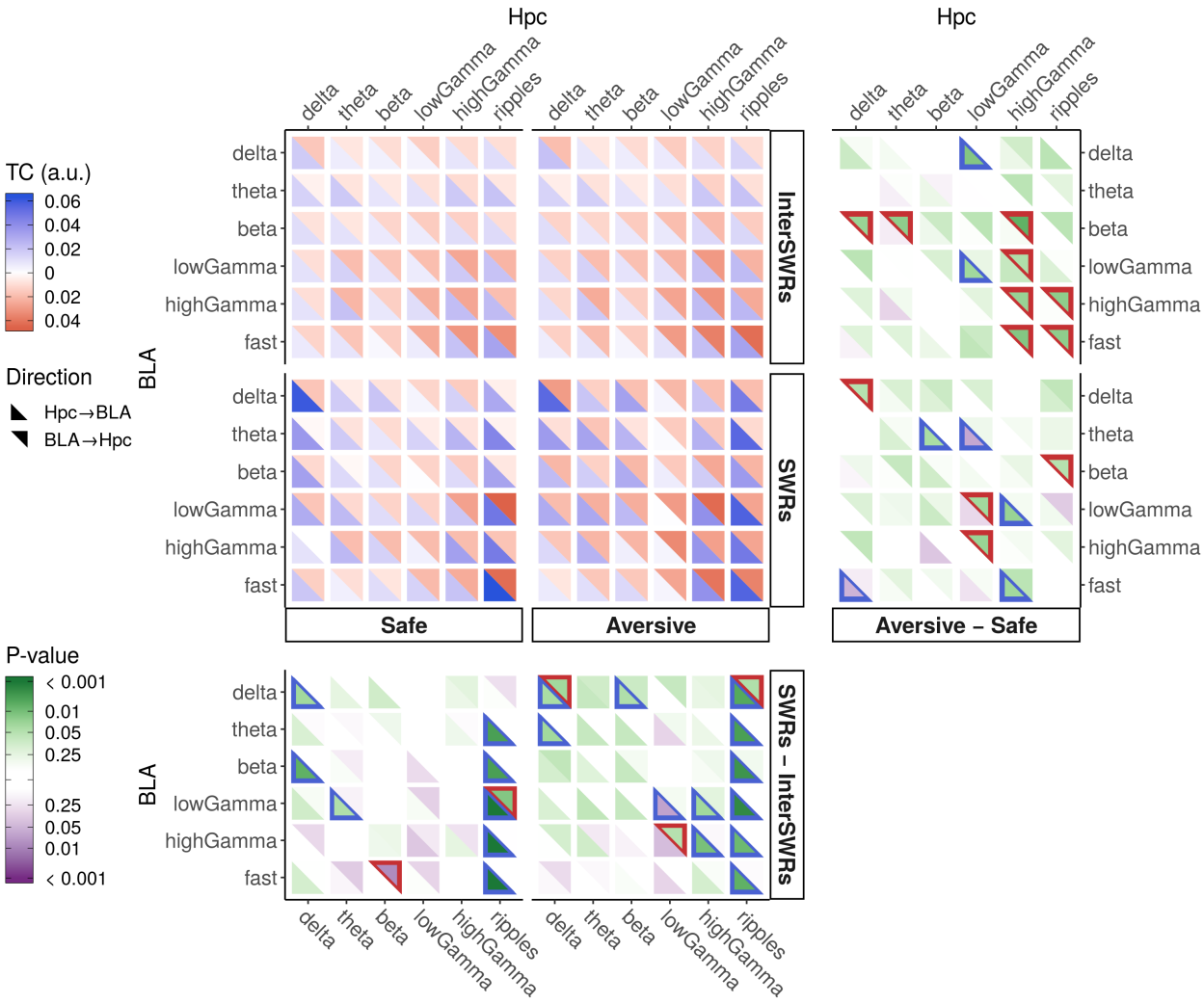

Figure 3-2: **Occurrence of SWRs during the run phases.** We calculated mean TC values (n = 13 sessions, 3 rats) separately for the four cases resulting from the combinations of the labels safe/aversive direction and presence/absence of SWRs during the run epoch. All sessions show at least 2 seconds of time featuring SWRs: this holds independently for both directions. The plot layout and interpretation are the same as in Fig. 2a (see main text for details).

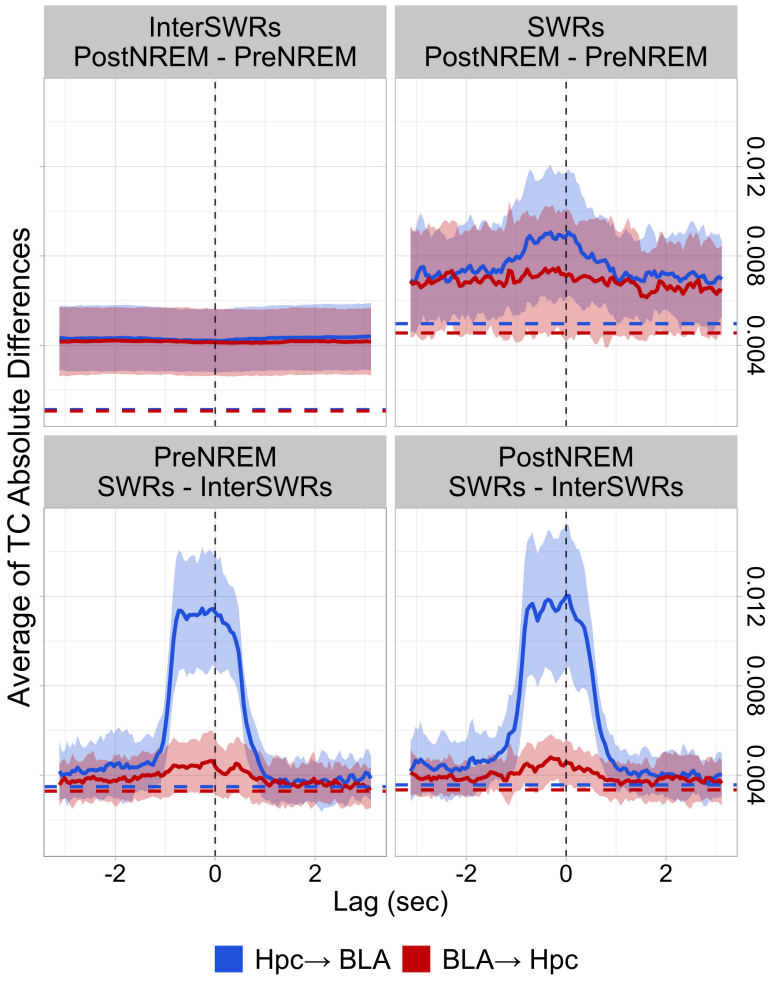

Figure 3-3: **Variation of TC absolute differences with respect to time lag during NREM sleep.** For each lag τ on the x-axis, TC values are shifted forward ($\tau > 0$) or backward ($\tau < 0$) in time (see Sec. 2.4 for details). The overall differences between the different sleep/SWRs phases were computed based on values similar to those used to perform the Wilcoxon signed-rank test in the right and bottom panels of Fig. 3. For each of the combinations of sleep/SWRs phases and for both directions of influence, we show the average of the absolute values (solid lines) along with their respective standard deviations (shaded areas) ($n = 20$ sessions, 3 rats). Dashed lines represent the average across sessions of the 95% quantiles of the chance level distributions (block permutation of time bins). *(continued)*

Figure 3-3: *(continued)* When we differentiate between pre- and post-NREM, the total variation, in both directions of influence, is generally higher than the underlying chance level. Thus, this derived global measure of interaction indicates a robust difference even when the individual TC differences may appear marginal. As expected, within the short time windows of SWRs, these values become sensitive to the lag of the TC, especially for hippocampus-driven interactions. In fact, if we consider the difference matrices between SWRs and interSWR phases (bottom panels), there is a clear regime of more or less constant differences within a window of $\sim$ 1 s. While this plateau is slightly asymmetric around the choice we made, this choice nonetheless appears reasonable.

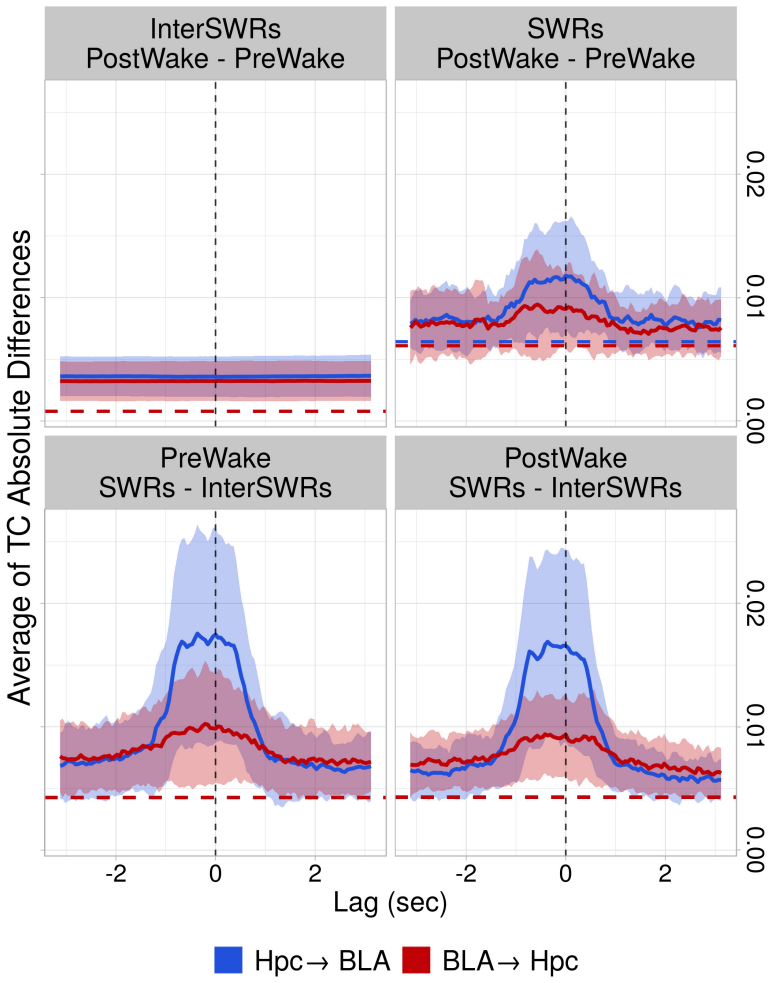

Figure 3-4: **Variation of TC absolute differences for awake phases as a function of lag.** The data were separated per behavioral epoch, pre- or post-, and the presence or absence of SWRs during awake periods ($n = 20$ session, 3 Rats). The figure is analogous to Fig. 3-3, and, as in that figure, our choice of relative alignment appears reasonable by observing that the absolute differences between SWRs and interSWR phases are sustained around our working point ($\tau = 0$). In addition, these variations are larger and appreciable also for the BLA→hippocampus direction.

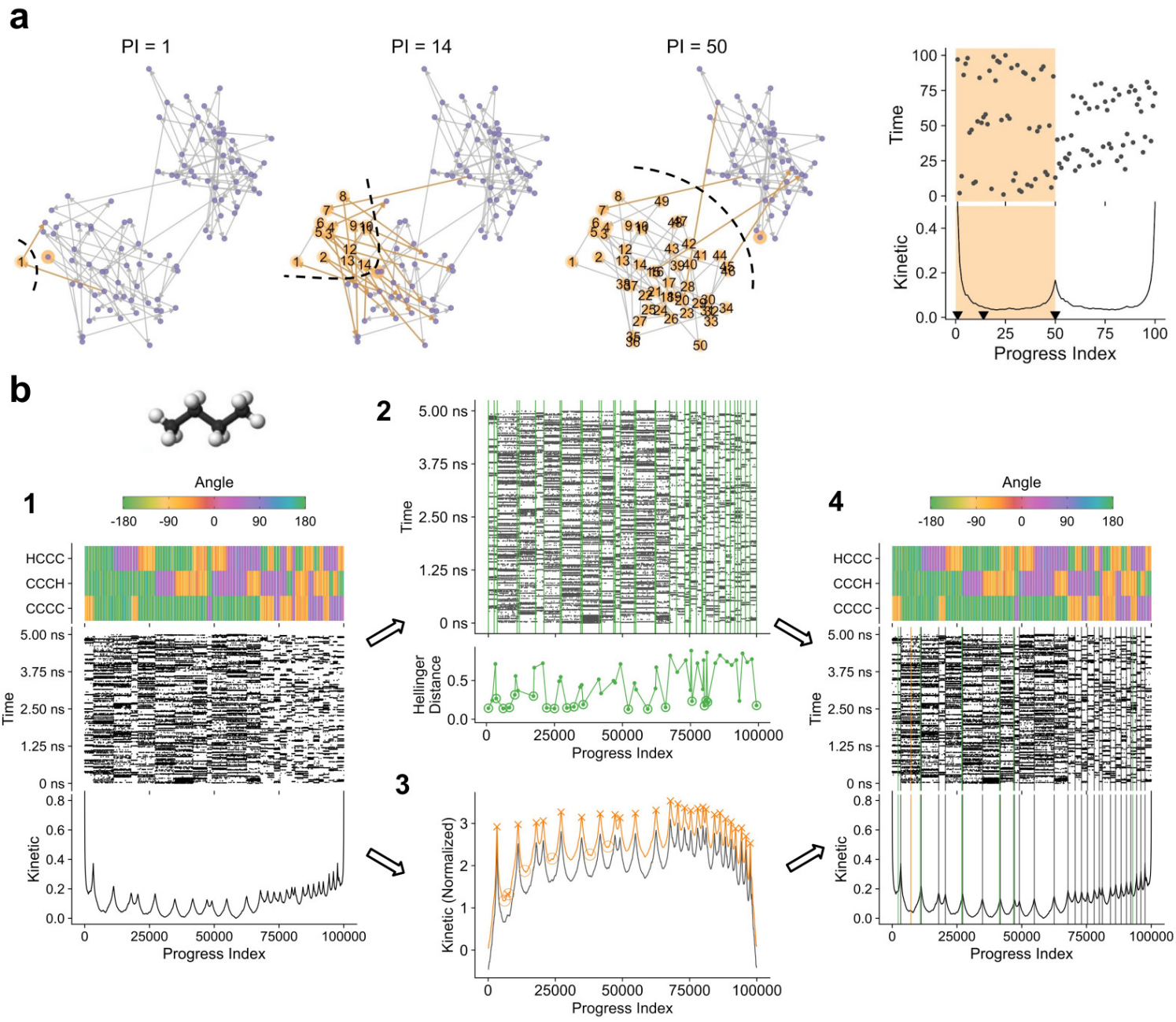

Figure 4-1: **Schematic of SAPPHIRE methodology.** (a) Progress Index (PI) construction. The first snapshot (time point) is arbitrarily selected and given PI = 1 (left panel). At any stage, the next index in sequence (violet dots with orange halos) is assigned to the snapshot that is closest to any of the points already added to the PI (orange and numbered dots). The SAPPHIRE plot (right) is composed by plotting two annotations ordered according to the PI. On the upper panel, the original time values per snapshot are shown (gray arrows in the left panels denote the sequence of time). This represents the "Time" or temporal annotation. On the lower panel, the kinetic annotation, in its simplest version, indicates the kinetic distance of the group of snapshots to the left of the specific PI value from that to the right. *(continued)*

Figure 4-1: *(continued)* Focusing on PI = 50, the group on the left (orange area in the SAPPHIRE plot) is connected by only four temporal transitions to the other cluster (orange edges on the respective PI = 50 panel). In graph-theory this is equivalent to find the minimum cut between the two subsets (dashed line). The kinetic annotation is inversely proportional to this number of connecting edges. For the lowest and for the highest values of PI (see *e.g.*, PI = 1), the kinetic function is necessarily large due to the small size of one of the two subsets. (b) SAPPHIRE-based clustering. A summary of the method is provided in Sec. 2.8 while full details can be found in (Cocina et al., 2020). Below we present the key aspects. A simulation of n-butane of 10^5 snapshots with a sampling time of 50 fs serves to illustrate the method (molecular cartoon above (1)). The system is described by the three dihedral angles associated with the three carbon-carbon bonds (denoted as HCCC, CCCH, CCCC). These have been used for the construction of the PI along with a proper angular metric. The resulting SAPPHIRE plot is shown in (1). The annotation on top shows the actual dihedral angle values. The clustering method proceeds by identifying putative partitions along the PI. This is done separately in the "Time" (2) and in the kinetic annotation (3). In (2), the analysis is based on an underlying 2D histogram ("PI" and "Time" axes) that serves to identify densely clustered 'blocks' of dots (gray). These correspond to individual residence periods in a given putative state, which is delimited by two partitions. An initial set of partitions (green vertical lines) is inferred from the end points of these 'blocks'. The temporal distributions of adjacent candidate states are quantified by their Hellinger distance, which is used as a test statistic (green points in the lower panel; these are aligned to the partitions delimiting two adjacent states). A null distribution is created from the repeated shuffling of the PI values across the two adjacent states within restricted temporal windows (*i.e.*, similar to the aforementioned 'blocks'). The partition is then kept if the respective Hellinger distance is statistically significant (P-value < 0.01), otherwise it is rejected (circled dots). The kinetic annotation, see (3), is normalized by subtracting the analytical curve derived from a random exploration of the phase space (black line), see Blöchliger et al. (2013). A smoothing filter is applied (orange curve) and peaks, corresponding to the partitions, are identified (crosses). The peaks that do not satisfy specific criteria of prominence are rejected (circles). Partitions from (2) and (3) are eventually combined (green and orange lines respectively) and merged if they overlap (black lines). Reprinted (adapted) with permission from Cocina et al. (2020).

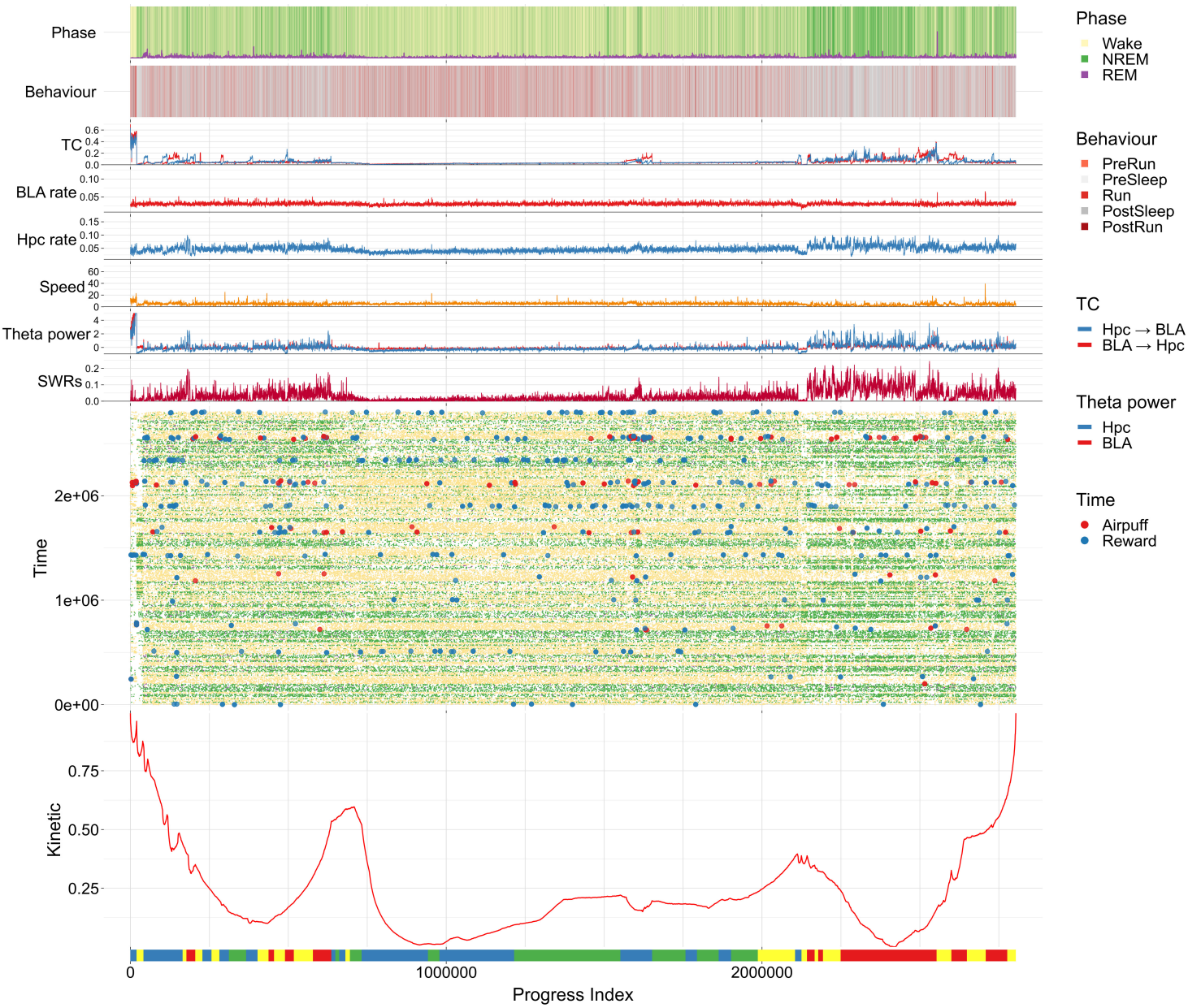

Figure 4-2: **SAPPHIRE plot for Rat1.** For the description of the labels and layout, please refer to Fig. 4 in the main text. The colored bars at the bottom indicate the results of the matching procedure between the four coarse states of Rat3 (Fig. 4) and the clusters that can be derived from the SAPPHIRE plot shown here (Rat1). As for Rat3, the clusters are extracted with the SAPPHIRE-based clustering (SbC) method that is described in Sec. 2.8. The matching procedure is explained in Sec. 2.11. We discarded one of the seven sessions due to its poor compatibility with all the other sessions. The raw data were the 18 principal components extracted from the 84-dimensional space of band powers and TCs (see Methods, Section 2.7). The remaining sessions comprise all five behavioral epochs. Note that Rat1 not only showed a reduced rate of crossings in the run phase compared to Rat2/3 ($\sim$0.3, $\sim$1.0, and $\sim$2.0, respectively) but also exhibited a $\sim$30-fold difference in this rate across sessions ($\sim$2 to 3-fold for Rat2/3).

Figure 4-3: **SAPPHIRE plot for Rat2.** For the description of the labels and layout, please refer to Fig. 4 in the main text. As for Fig. 4-2, the colored bars at the bottom are used to indicate the clusters that were matched to the four coarse states identified for Rat3 (Fig. 4; see Secs. 2.7 and 2.11). All seven sessions for this rat and all five behavioral epochs were used in the analysis. The raw data were the 18 principal components extracted from the 84-dimensional space of band powers and TCs (see Methods, Section 2.7).

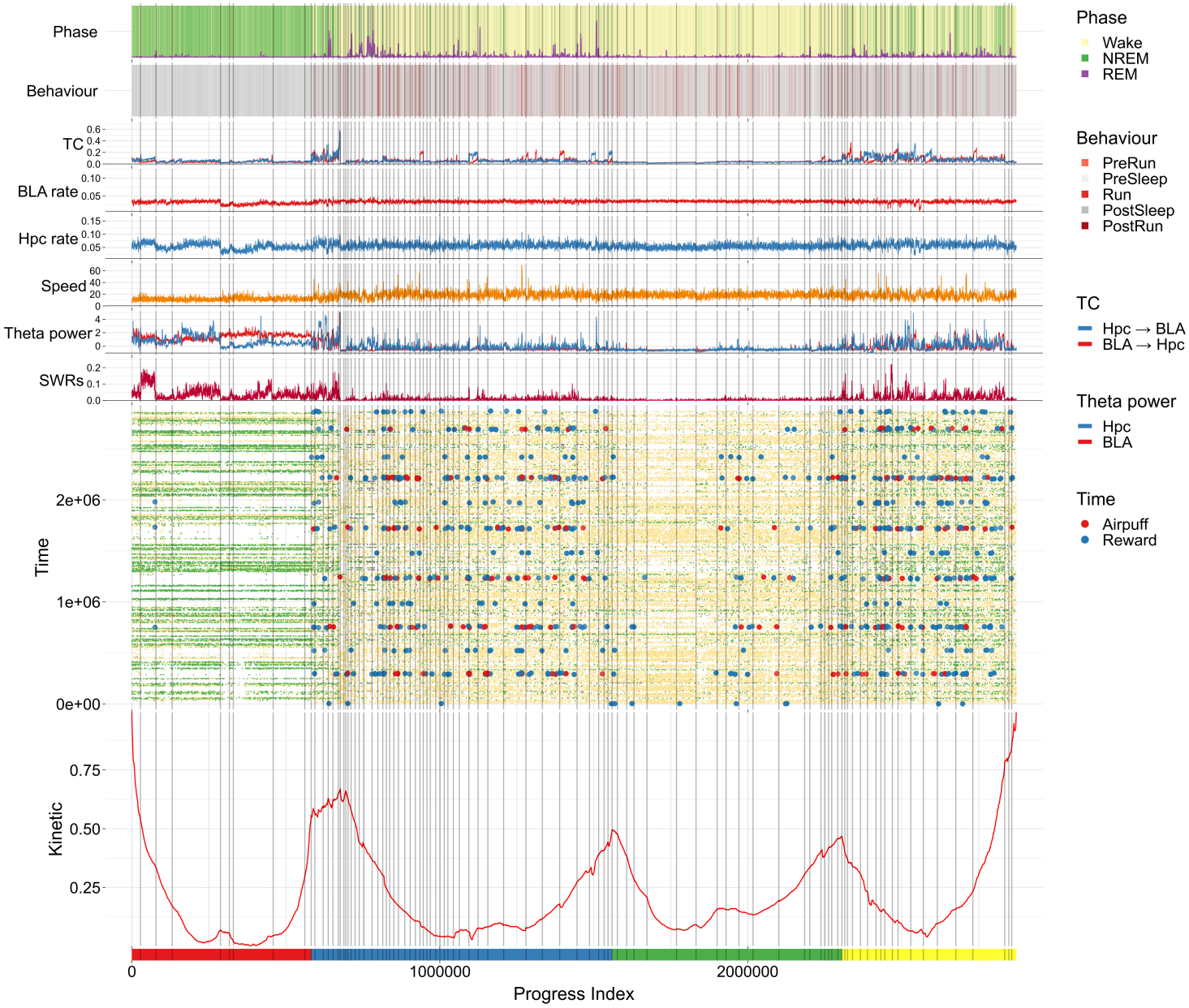

Figure 4-4: **SAPPHIRE plot for Rat3 with highlighted clusters.** This is the same as Fig. 4 in the main text only that the separations between identified clusters are shown as vertical lines (see Methods, Section 2.8 for details).

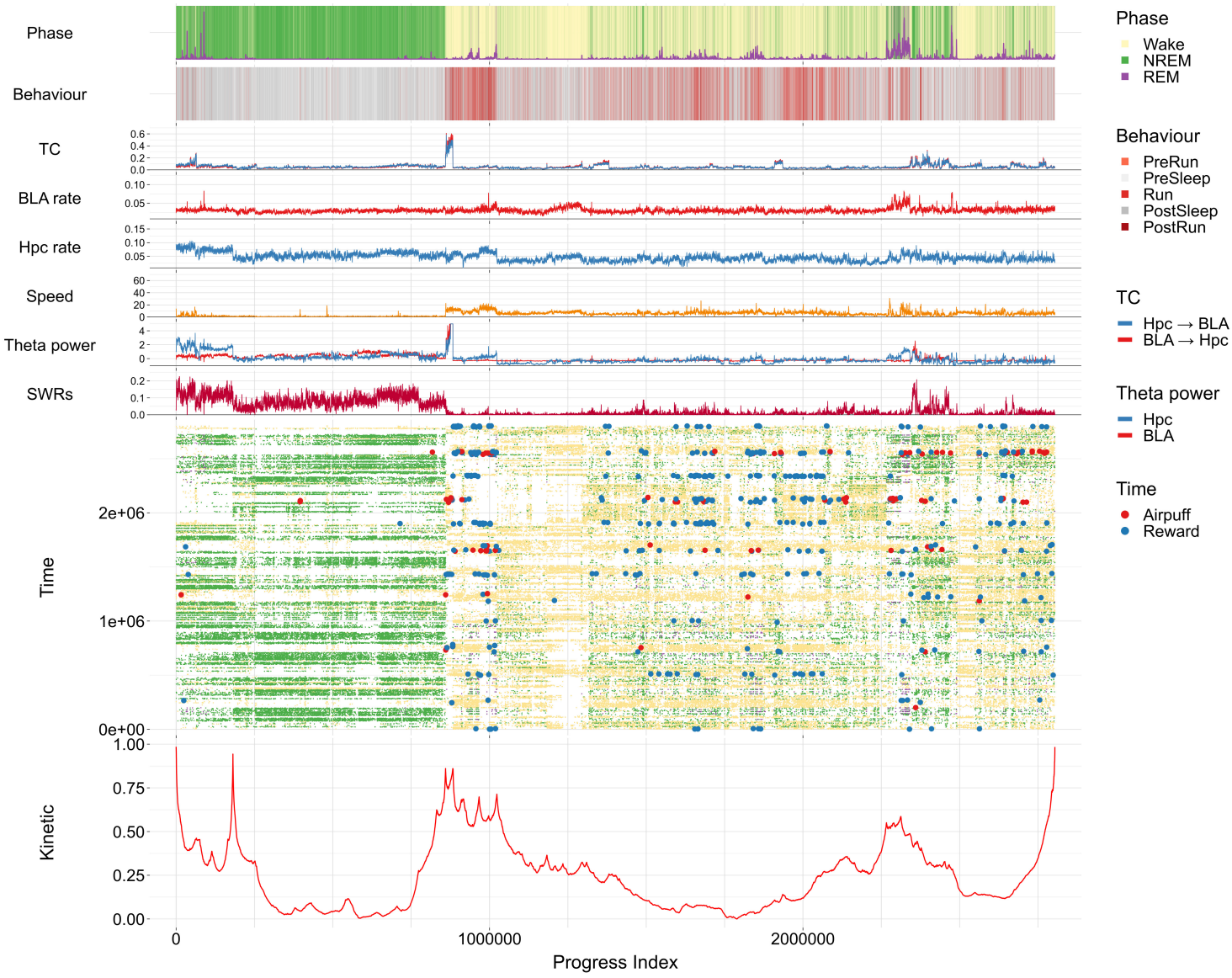

Figure 4-5: **SAPPHIRE plot for Rat1 with alternative preprocessing.** Differing from Fig. 4-2, we included all six sessions, and the source data were preprocessed differently. Briefly, we computed locally adaptive weights from the 84 combined band powers and TC values and used them to strengthen the impact of temporally stable signals (see Methods section 2.6 for details). This is expected to be more sensitive to episodes of sustained activity in particular power bands. Only the first five principal components (by variance) of the resulting data set were retained. As a result of this preprocessing pipeline, the partitioning of labels is much more localized compared to Fig. 4-2. This discrepancy is a particular feature of the data for Rat1, and it arises primarily because the partitioning in Fig. 4-2 is poor.

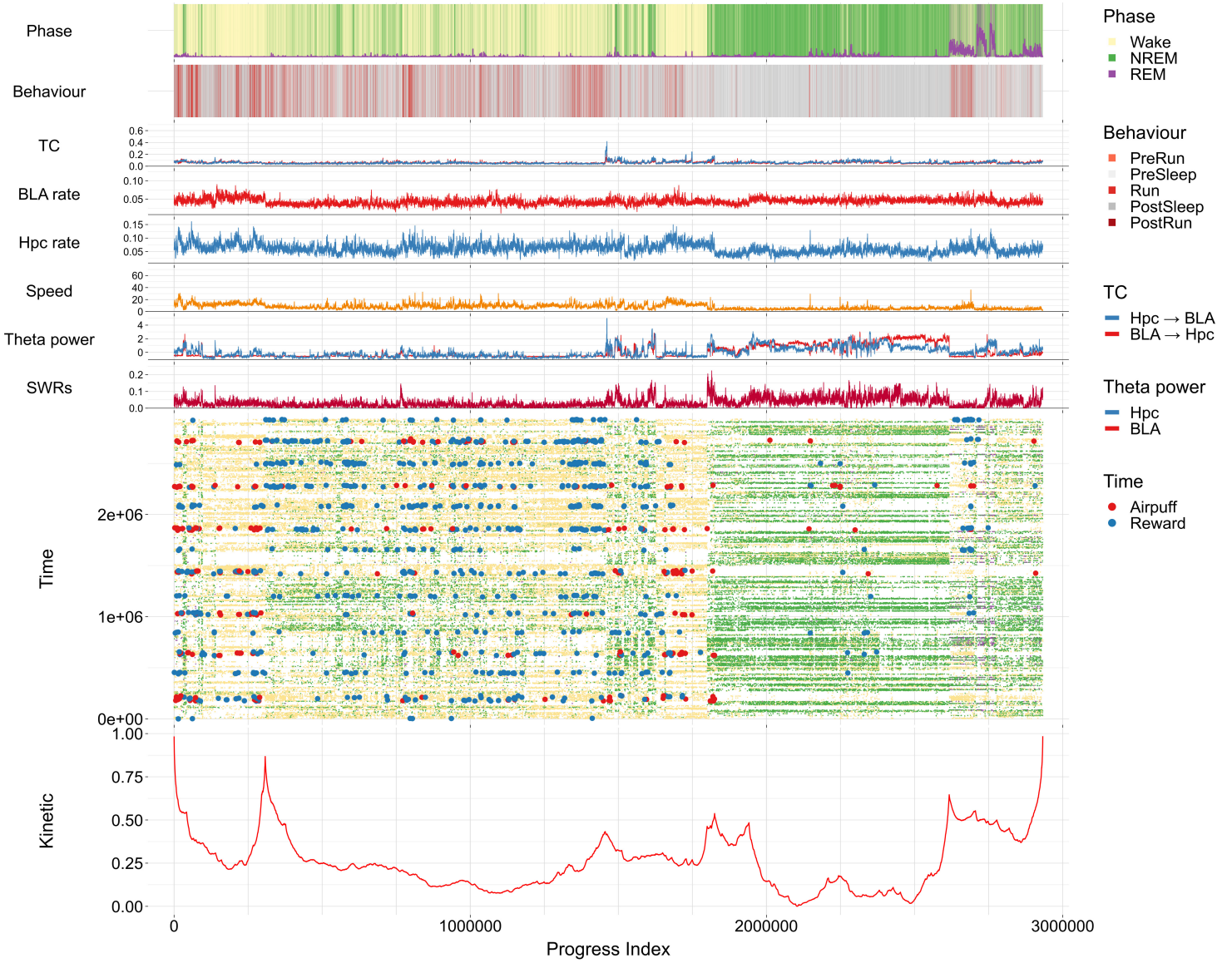

Figure 4-6: **SAPPHIRE plot for Rat2 with alternative preprocessing.** The raw data were the same as in Fig. 4-3, and the preprocessing is the same as that described for Fig. 4-5. Note that this plot is much more similar to Fig. 4-3 than Fig. 4-5 is to Fig. 4-2.

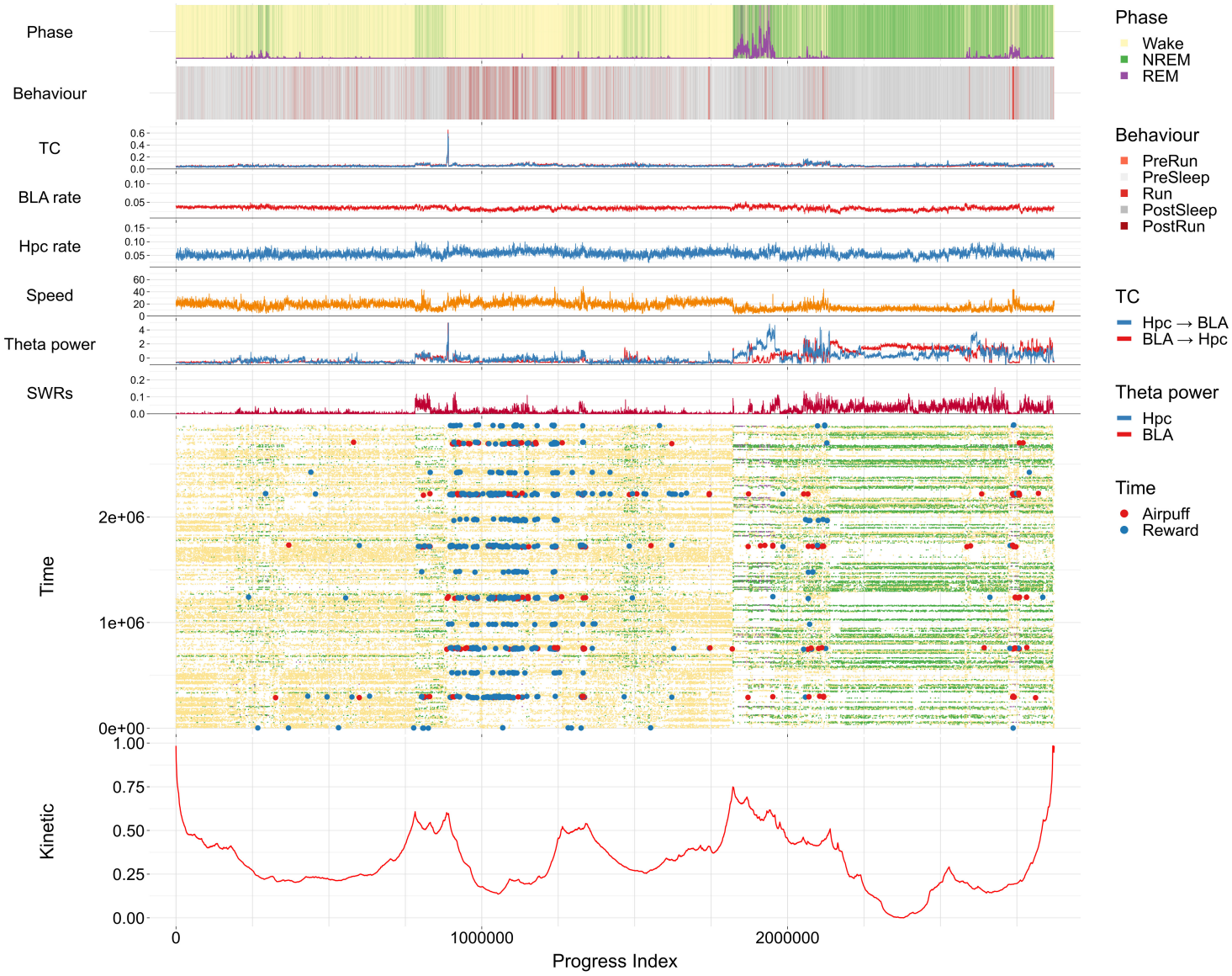

Figure 4-7: **SAPPHIRE plot for Rat3 with alternative preprocessing.** The raw data were the same as in Fig. 4 in the main text, and the preprocessing is the same as that described for Fig. 4-5.

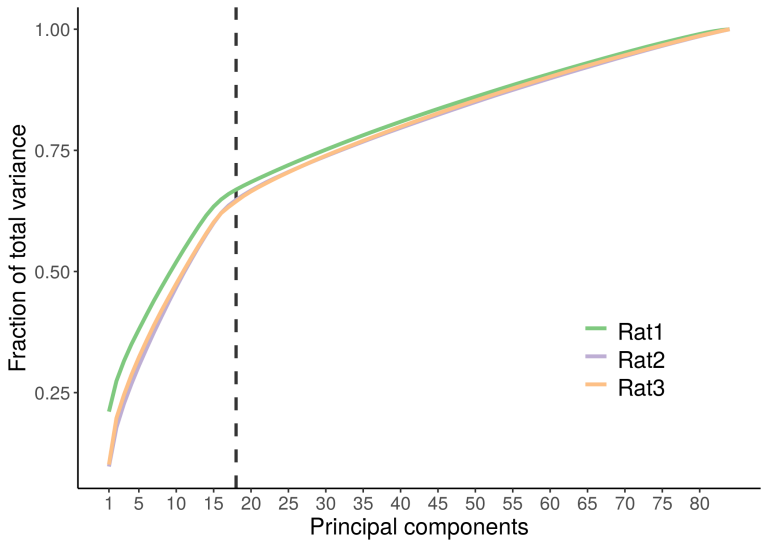

Figure 4-8: **Explained variance with respect to the number of principal components.** This figure refers to the data set composed of the 12 band powers (6 each for hippocampus and amygdala) and the 72 TC values (36 per direction of influence).

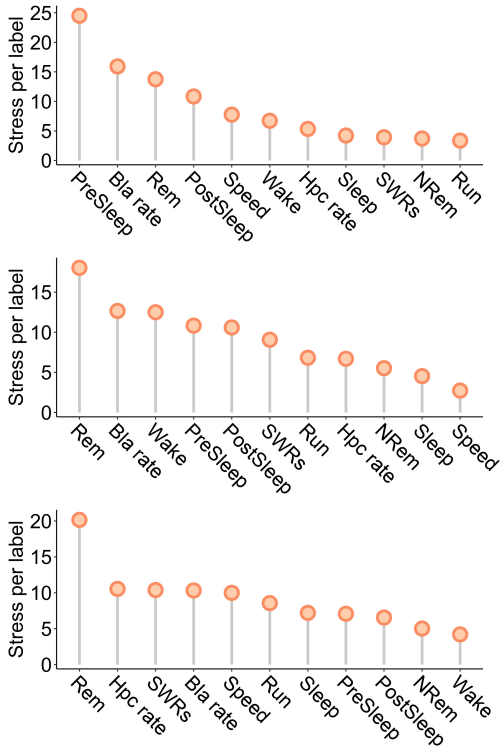

Figure 6-1: **Analysis of stress values of the unfolding projection.** (Top, middle, and bottom) Results for Rat1, Rat2, and Rat3, respectively. Percentage of total stress values per unfolded label.

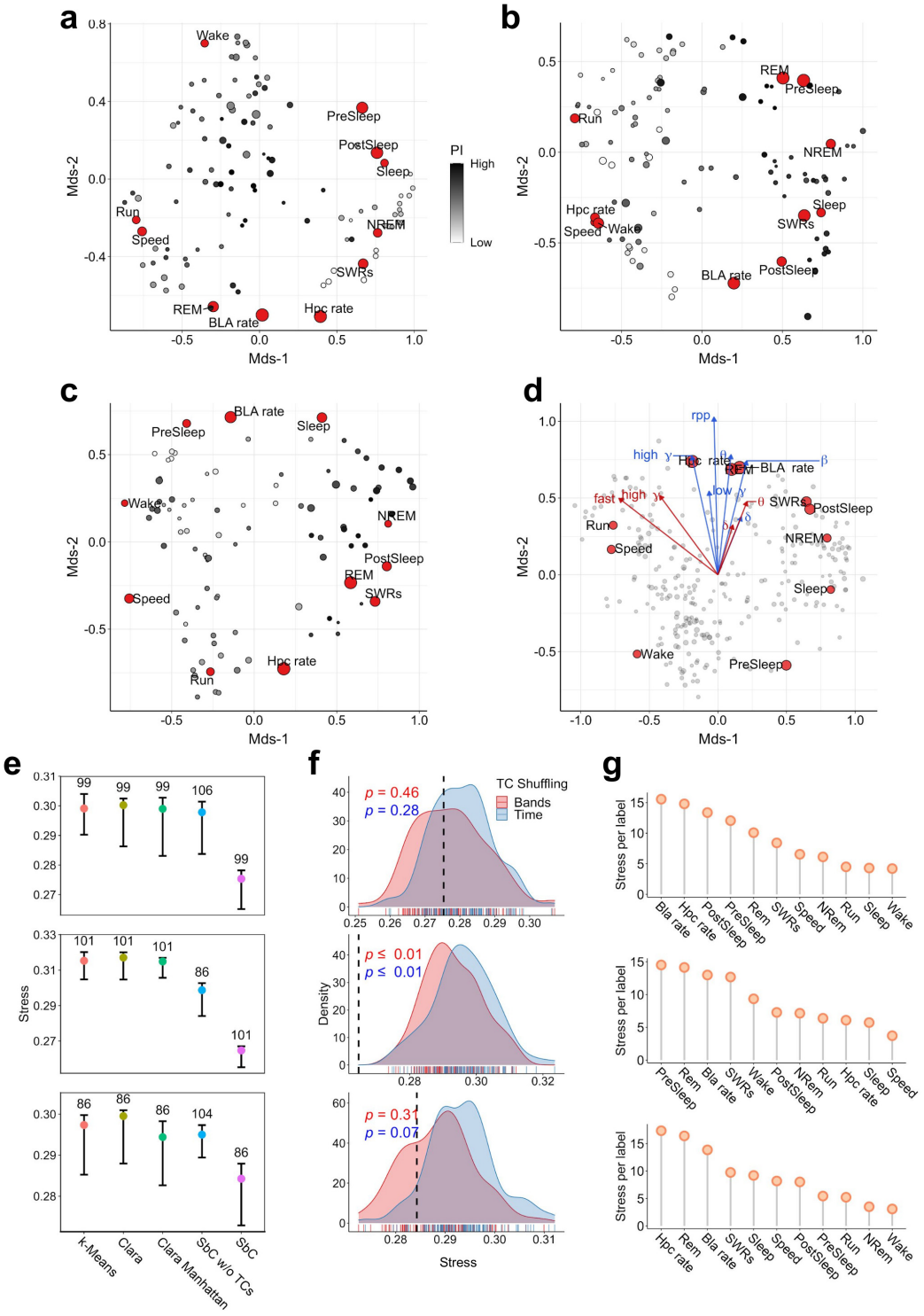

Figure 6-2: **Results of the unfolding projection applied to the combined data set with alternative preprocessing.** The SAPPHIRE plots from which the clusters are derived are shown in Figs. 4-5, 4-6, and 4-7. (a-c) Unfolding plots for Rat1, Rat2, and Rat3, respectively (analogous to Fig. 5b-d in the main text). *(continued)*

Figure 6-2: *(continued)* (d) Unfolding projection for all three rats and biplot of the relevant mean TC components (analogous to Fig. 5e in the main text). (e-g) Stress per label, comparison with other clustering methods, and robustness with respect to TC shuffling (analogous to Fig. 5f and Figs. 6-1a,b).