PNAS
www.pnas.org

Supplementary Information for

Matched-filter coding of sky polarization results in an internal sun compass in the brain of the desert locust

Frederick Zittrell, Keram Pfeiffer, Uwe Homberg

Corresponding author:
Uwe Homberg
Faculty of Biology, Department of Animal Physiology, Philipps-University Marburg, Karl-von-Frisch-Street 8, 35032 Marburg, Germany
E-mail: homberg@biologie.uni-marburg.de

This PDF file includes:

Supplementary Methods

Visualization of spherical data. An illustration of the coordinate system used throughout the paper is shown in Fig. S1. Spherical coordinates were transformed on a polar-coordinate-like grid for figures where data are shown on a flattened hemisphere from above. This transformation has the advantage that the spherical coordinate grid is not distorted as it would occur when simply viewed from above.

Estimating visual fields. The algorithm used for estimating the extent of visual fields (Figs. 3 and S4) is illustrated step-by-step in Fig. S2. We chose the R^2 threshold of 0.75 based on visual inspection of the resulting visual fields.

Fig. S3 shows the effect of the threshold height on an example data set: A lower threshold typically results in larger fields, a higher threshold results in smaller fields. The number of resulting fields may also be affected because the threshold affects the number of data points that are considered for clustering (Fig. S2*C*). A lower threshold typically leads to larger clouds of data points, which may be clustered into the same field. A higher threshold leads to more data points being discarded, which may include all data points that otherwise would have formed a separate cloud.

The threshold used in this study results in a reasonable number of fields per neuron (on average 1.35 fields per neuron). The fields that were estimated with this threshold also reflected our impression of the raw data sets well. It is apparent however that this estimation is very coarse and should only give a qualitative impression of the visual fields of the recorded neurons. A much denser grid of sample points would be necessary for a finer analysis of field size and shape. A denser sampling grid would, e. g., allow for the calculation of the centroid coordinate of each field and the geometrical extent of each field could be measured with precision.

Calculation of the best-matching polarization pattern. Fig. S5 illustrates the evaluation steps that yield the best-matching polarization pattern of a neuron (same data set as in Fig. 4). Fig. S5*A* shows the final result of the evaluation: The neuron's measured response pattern of preferred angles of polarization (AoP) and the calculated best-matching sky polarization pattern with the corresponding solar coordinates. This polarization pattern has the lowest deviation (D) from the pattern of preferred AoPs, i. e., this polarization pattern fits the response pattern best. D was calculated for 32,760 equally distributed solar positions. As a comparison, Fig. S5*B,C* shows two selected solar positions that produce polarization patterns with intermediate and high deviation.

The deviation between the response pattern and each polarization pattern, which depends on the solar coordinates, was calculated using a deviation matrix Θ and a weight matrix W. The contents of these matrices are illustrated in Fig. S5*D* by means of three selected stimulus positions, for the polarization pattern shown in Fig. S5*A*. Θ contains the angular deviations between the response pattern and the given polarization pattern at all tested stimulus positions. We calculated this deviation at each stimulus position using

$$\delta(\alpha,\Phi) = \begin{cases} ||\alpha-\Phi|-180|, & \text{if } |\alpha-\Phi|>90^\circ \\ |\alpha-\Phi|, & \text{otherwise} \end{cases},$$

where α is the AoP that was calculated at this position and Φ is the neuron's preferred AoP at this position. The angular deviations for each generated polarization pattern were stored in a deviation matrix

$$\Theta(az,el) = \begin{pmatrix} \delta(\alpha(az,el)_1,\Phi_1) \\ \vdots \\ \delta(\alpha(az,el)_i,\Phi_i) \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_i \end{pmatrix},$$

where $\alpha(az,el)_i$ is the AoP at stimulus position i of the polarization pattern generated by the sun at azimuth az and elevation el and Φ_i is the neuron's preferred AoP at position i. The weight matrix for each polarization pattern contained

$$W(az,el) = \begin{pmatrix} DoP(az,el)_1 \cdot R_1^2 \cdot \omega_1 \\ \vdots \\ DoP(az,el)_i \cdot R_i^2 \cdot \omega_i \end{pmatrix},$$

where az and el are the solar coordinates used to generate the polarization pattern, $DoP(az,el)_i$ is the degree of polarization at the stimulus position i of the generated polarization pattern (illustrated by the length of the black bars in Figs. 4, S5), R_i^2 is the R^2 value of the neuron's response when stimulated from position i (illustrated by the length of the red double arrows in Figs. 4, S5) and ω_i is the sum of the arc distances between position i and the nearest 22 % of stimulus positions that were used in the experiment, normalized over the range of all ω values in the data set.

We chose the value of 22 % in accordance with previous experiments (1), where every data set contained data from 37 positions and the nearest 8 positions were used to calculate ω_i. In this study, the total number of stimulus positions varied due to the instability of intracellular recordings. Hence, we chose $8/37 \approx 22\,\%$ of the total number of stimulus positions of each data set as the number of nearest positions used to calculate ω_i. The actual number of nearest positions used for calculation was

$$k = \begin{cases} 1, & \text{if } Round(i\cdot 0.22)<1 \\ Round(i\cdot 0.22), & \text{otherwise} \end{cases},$$

where i is the number of stimulus positions in the given data set. This weighting factor was introduced, because the stimulus positions were spatially not homogeneously distributed, owing to constraints of the setup. Therefore we sought to counterbalance the overrepresentation of values from densely sampled areas using the weighting factor ω. Fig. S6 however shows that the impact of ω on the best-matching polarization pattern is rather small, or even absent in many cases irrespective of how many neighbors are used for the calculation.

Finally, we calculated the weighted average deviation D between the response pattern of a single neuron (red double arrows in Figs. 4, S5) and a polarization pattern (black bars in Figs. 4, S5) by

$$D(az, el) = \overline{\Psi}(\Theta(az, el), W(az, el)),$$

where az is the azimuth and el is the elevation of the solar position that produces the polarization pattern, Θ is the deviation matrix, W is the weight matrix and $\overline{\Psi}(\Theta, W)$ returns the average angle of Θ weighted by W. We used `circ_mean` from the *CircStat* package (2) for $\overline{\Psi}(\Theta, W)$.

D was calculated for 32,760 equally distributed solar positions and the one with the lowest $D(az, el)$ was considered the solar position that would produce the sky polarization pattern that matches the neuron's response pattern best. Fig. S5*E,F* shows the deviation values as a function of solar position for one example data set. For clarification, the sky polarization patterns generated by two selected solar positions are additionally visualized: Fig. S5*C* shows a polarization pattern that matches the neuron's response pattern badly (high D value); Fig. S5*B* shows a polarization pattern that yields a medium D value. Note that absolute D values have no meaning and that D values cannot be compared between neurons. Hence, we qualified the goodness-of-fit using a bootstrap approach (see Methods in the main text).

Robustness analysis of the compass regression. We investigated the effect of pooling different neuron types for the regression analysis (Fig. 5B) by excluding each of the four neuron types followed by regression analysis on the resulting subset. Fig. S9 shows that the compass regression based on the complete data set (Fig. 5B) is robust against the exclusion of any single neuron type. Excluding tangential cells (Fig. S9*A*) had a minor effect on the regression intercept and slope steepness, but did not affect the sign of the slope and the correlation statistics. Exclusion of any other single type had no qualitative effect on the regression parameters and the correlation coefficient ρ. Please note that care should be taken interpreting the P values of these reduced data sets, because the recommended sample size of $N \geq 11$ for calculating P values (3) is hardly exceeded in most of them.

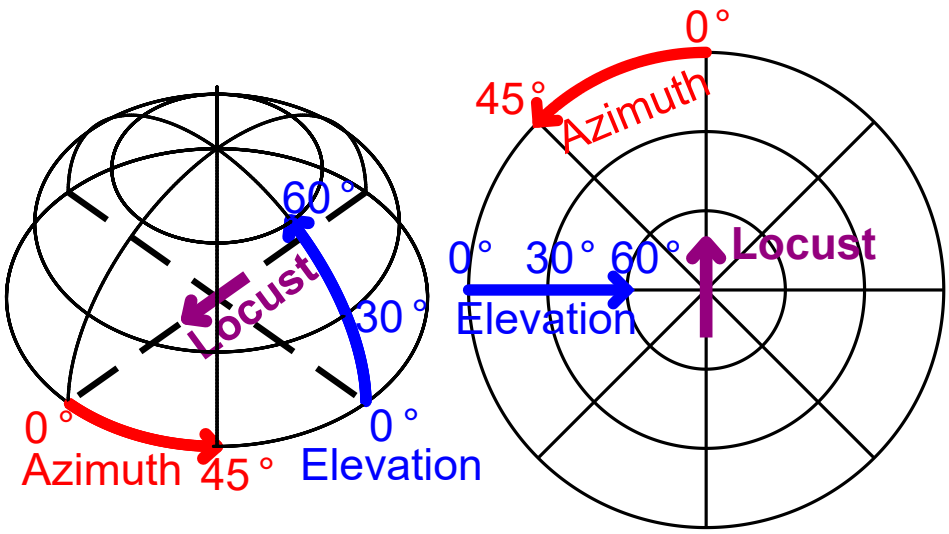

Fig. S1. Spherical coordinate system used in the paper. Left: 3D view. Right: Flattened view from above after transformation on a polar-coordinate-like grid.

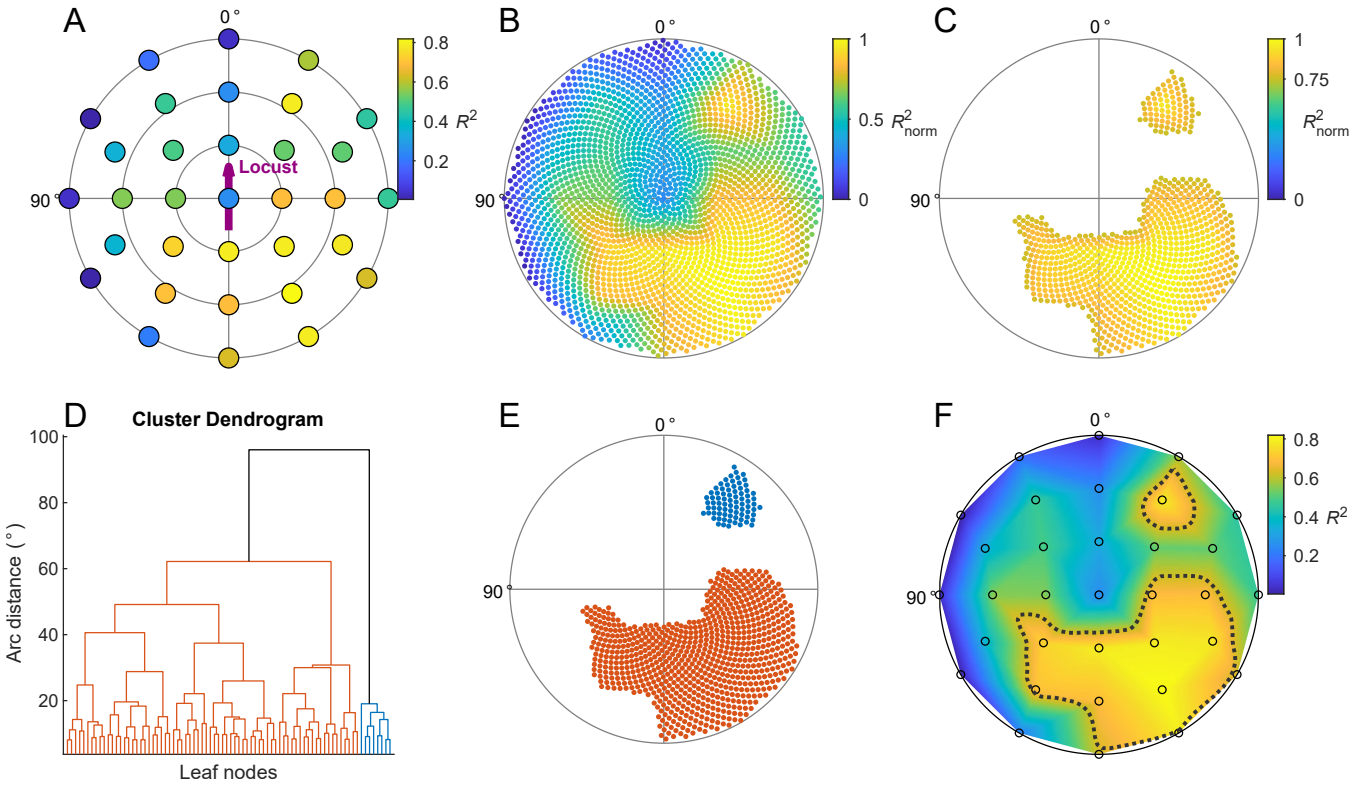

Fig. S2. Illustration of the procedure to estimate visual fields. Related to Fig. 3. *(A)* R^2 values of the responses of a single neuron to polarized-light stimulation at different positions in the dorsal visual field (cf. Fig. S1 right). All responses are plotted irrespective of whether the neuron responded with a statistically significant activity modulation when stimulated from the respective position. *(B)* A dense data-point grid was created and the R^2 values of the responses were used for interpolation across the grid. The scale was normalized to the range in the data set. *(C)* All data points with $R^2_{norm} < 0.75$ were discarded, producing, in this example, two separated clouds of relatively high R^2 values, which indicates high polarization sensitivity of the neuron at these coordinates. *(D)* The remaining data points were classified into continuous clouds with a hierarchical cluster analysis using the arc distance as distance measure. The dendrogram shows the hierarchical cluster tree. Data points are grouped into leaf nodes for clearer visualization, i. e., each leaf node contains ≥ 1 data points. The two colored groups of nodes are far apart, indicated by the length of the connecting branch (black line). *(E)* Data points that belong to the cluster groups, same coloring as in the dendrogram. We considered each group to constitute a receptive field of this neuron. *(F)* Final result of the visual field estimation of this response pattern. The dotted lines are the smoothed convex hulls around each cluster. Similar to *B*, the R^2 values are interpolated between the stimulus positions (black circles), but, for aesthetic reasons, over a surface generated by Delaunay triangulation of the positions.

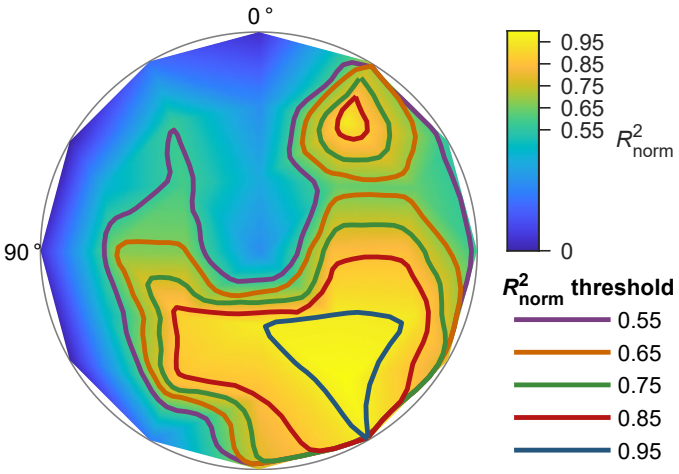

Fig. S3. Impact of visual field separation threshold on visual field estimation. Related to Figs. 3, S4 and S2. Colored lines show the visual field extents that result from visual field estimation when different thresholds are used (cf. Fig. S2, same data set).

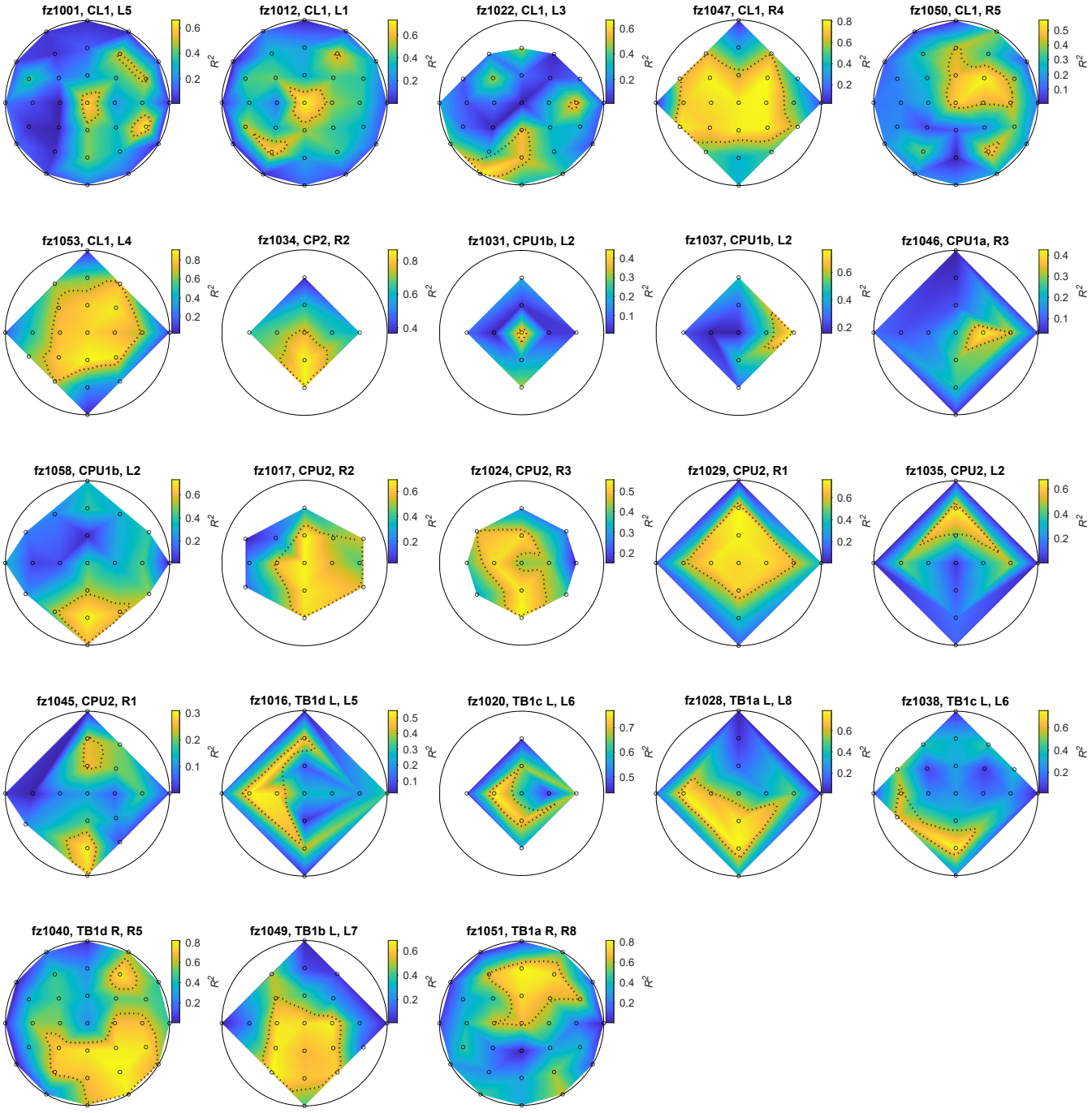

Fig. S4. AoP sensitivity maps of all recorded neurons; cf. Fig. 3. Flattened spherical top view (cf. Fig. S1). The R^2 value of the AoP response at each sampled position (black circles) is color coded and interpolated in between. Dotted lines are smoothed 75 % isolines of the normalized R^2 values (cf. Fig. S2 for calculation procedure). For each plot the data set ID, the neuron type and the arborization column in the PB are indicated.

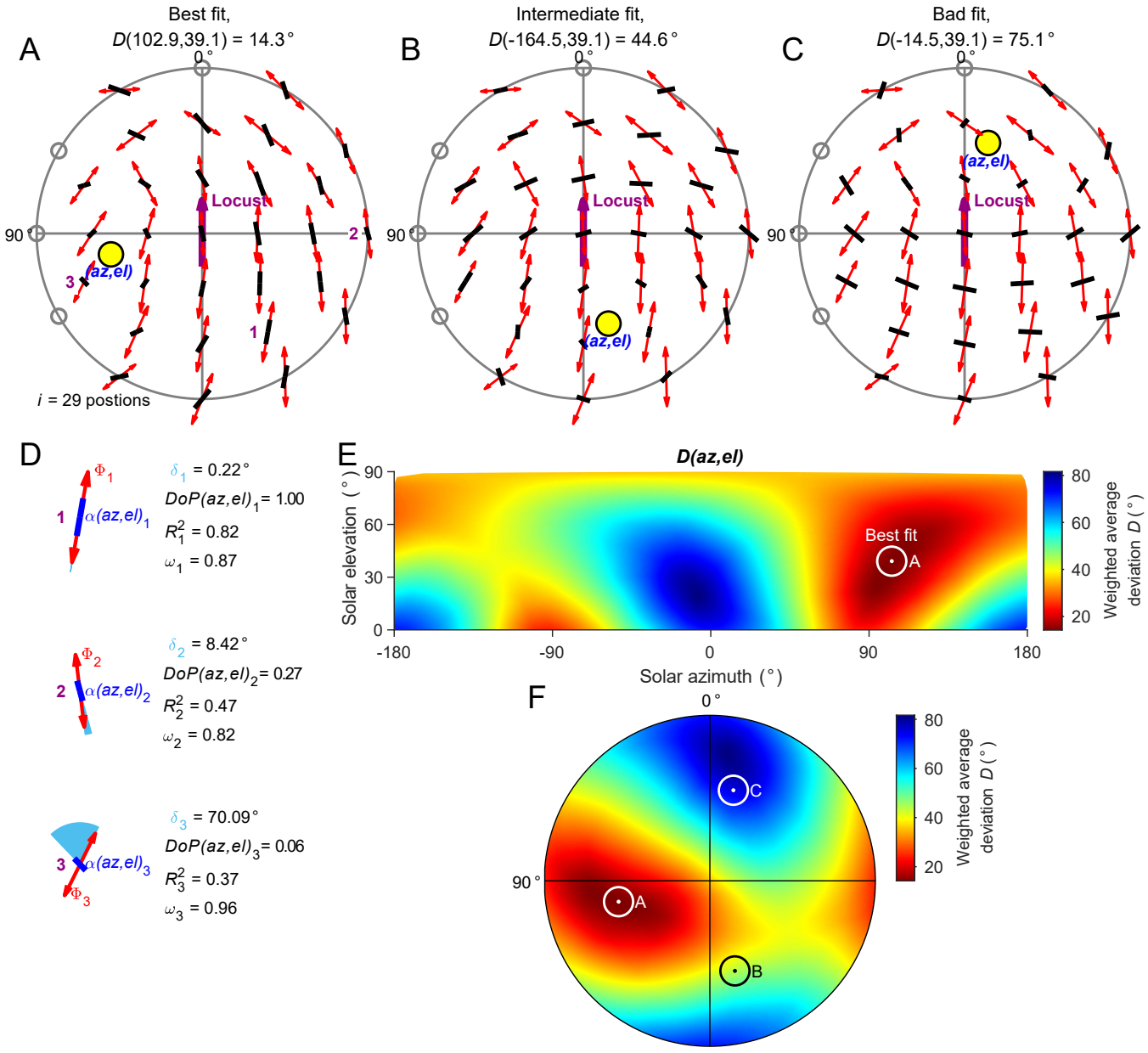

Fig. S5. Illustration of the procedure to calculate the best-matching polarization pattern. *(A)* Measured pattern of preferred AoPs (red double arrows) of a single cell (same data set as in Fig. 4) and the best-matching polarization pattern (black bars) with the sun at the corresponding position. The red double arrows are scaled by the R^2 value of the responses and the black bars are scaled by degree of polarization (DoP). The weighted average deviation $D(az, el)$ between the response pattern and the polarization pattern for this solar position is denoted in the title. *(B,C)* Same plots as in *A*, but for intermediately *(B)* and badly *(C)* matching polarization patterns. *(D)* Illustration of the numerical data contained at every stimulus position by means of three example positions. The indices correspond to the numbered positions in *A*. At each stimulus position i, the AoP (blue bars, $\alpha(az, el)_i$ and $DoP(az, el)_i$ were calculated depending on the solar position (at azimuth az and elevation el). DoP values here are relative to the natural scale, i. e., a value of 1 corresponds to a DoP value of 0.75, which is the maximum DoP in a naturalistic polarization pattern. Φ_i denotes the neuron's preferred AoP when stimulated from position i. δ_i denotes the angular deviation between α_i and Φ_i. ω_i denotes the normalized sum of the arc distances to the neighboring positions. See text for details on the calculation of ω_i and $D(az, el)$. *(E)* Weighted average deviation color coded as a function of solar position. The marker indicates the solar position that generates the polarization pattern in *A*, which has the lowest weighted average deviation $D(az, el)$ of all solar positions used for the fitting procedure (N = 32,760 equally distributed solar positions). *(F)* Same data as in *E*, but visualized in flattened spherical coordinates (cf. Fig. S1 right). The markers labeled B and C indicate the solar positions for which the polarization patterns are plotted in *B* and *C*, respectively.

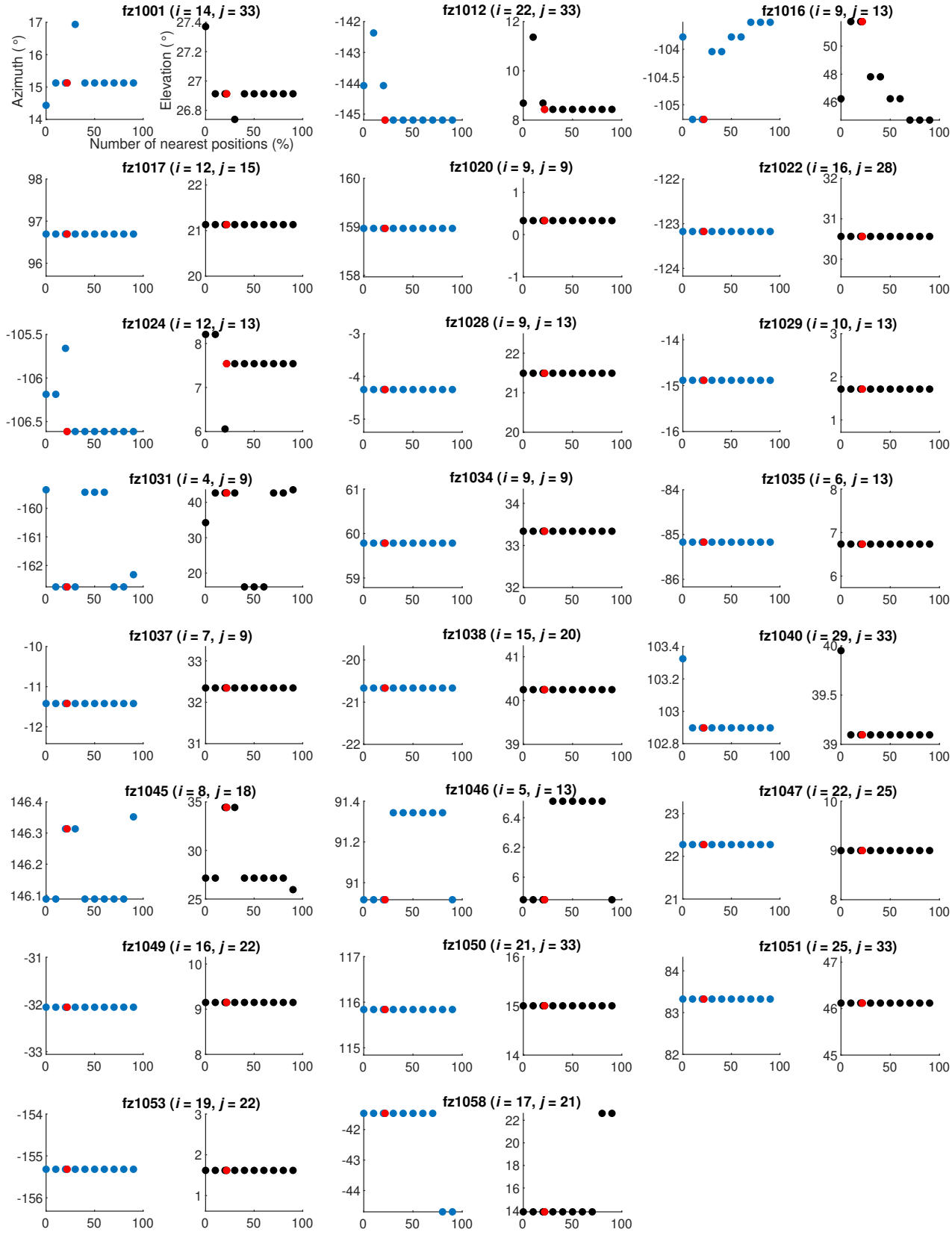

Fig. S6. Impact of the number of neighboring stimulus positions used on the fitting results. The fitting procedure was done as described, but the relative number of neighboring stimulus positions used to calculate the weighting factor ω was varied. Azimuth (blue data points) and elevation (black data points) of the solar position resulting from the fit are plotted as a function of this number. Zero percent means that the nearest-neighbor distance was ignored for the fit calculation. Red dots indicate the number used in this study (22 %). The title of each plot pair indicates data set ID, *i* indicates the number of stimulus positions that elicited a statistically significant AoP-response of the neuron and were used for fit calculation and *j* indicates the total number of positions from which the animal was stimulated during the recording session. Please note individual y-scaling for each plot.

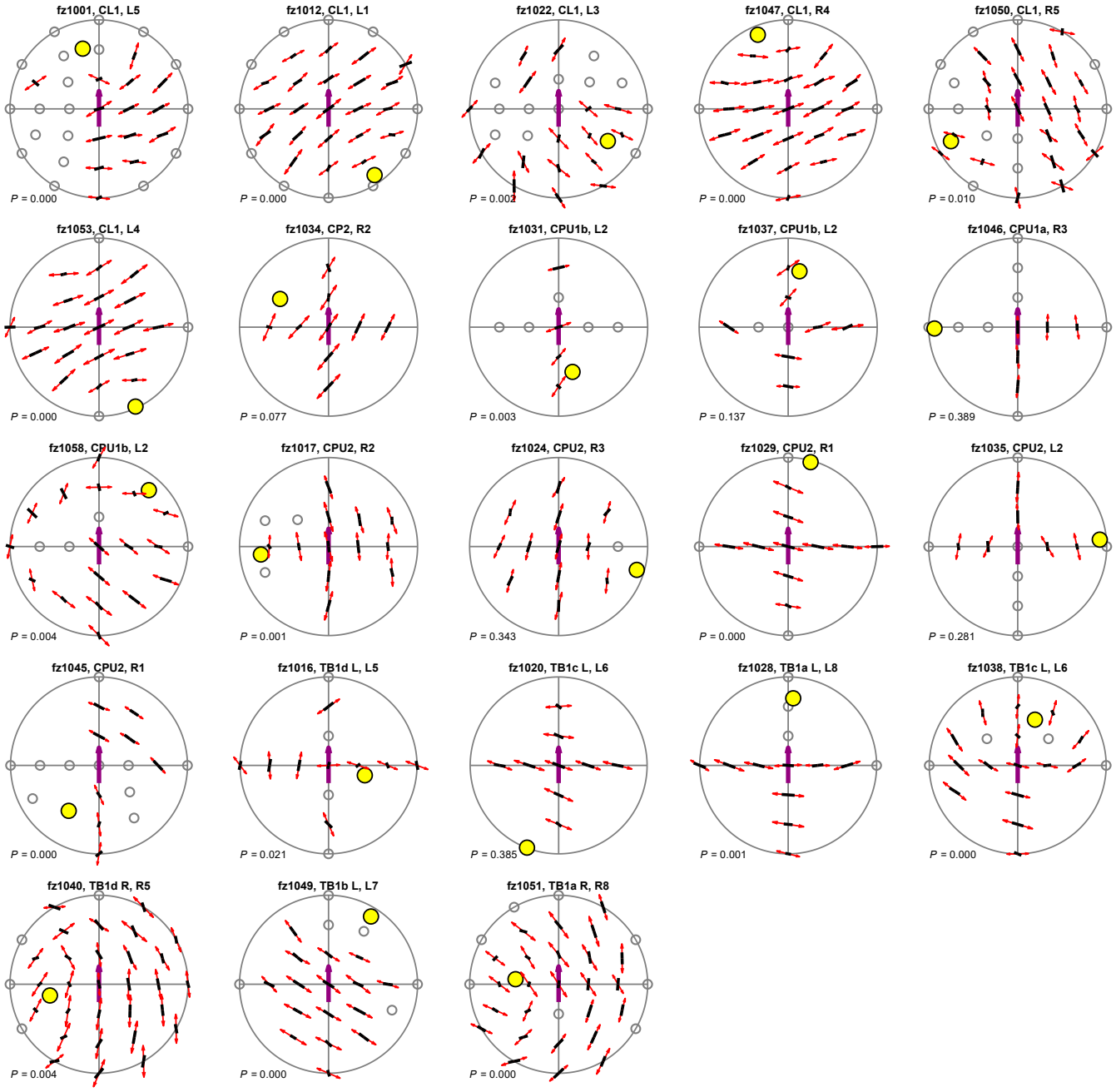

Fig. S7. Response patterns with best-matching polarization pattern; cf. Fig. 4. Flattened spherical top view (cf. Fig. S1 right). The response pattern of preferred AoPs is plotted as red double arrows, the best-matching polarization pattern is plotted as black bars. The yellow marker indicates the solar position that was used to generate this polarization pattern. For each plot the data set ID, the neuron type, the arborization column in the PB and the *P* value of the bootstrap test are indicated.

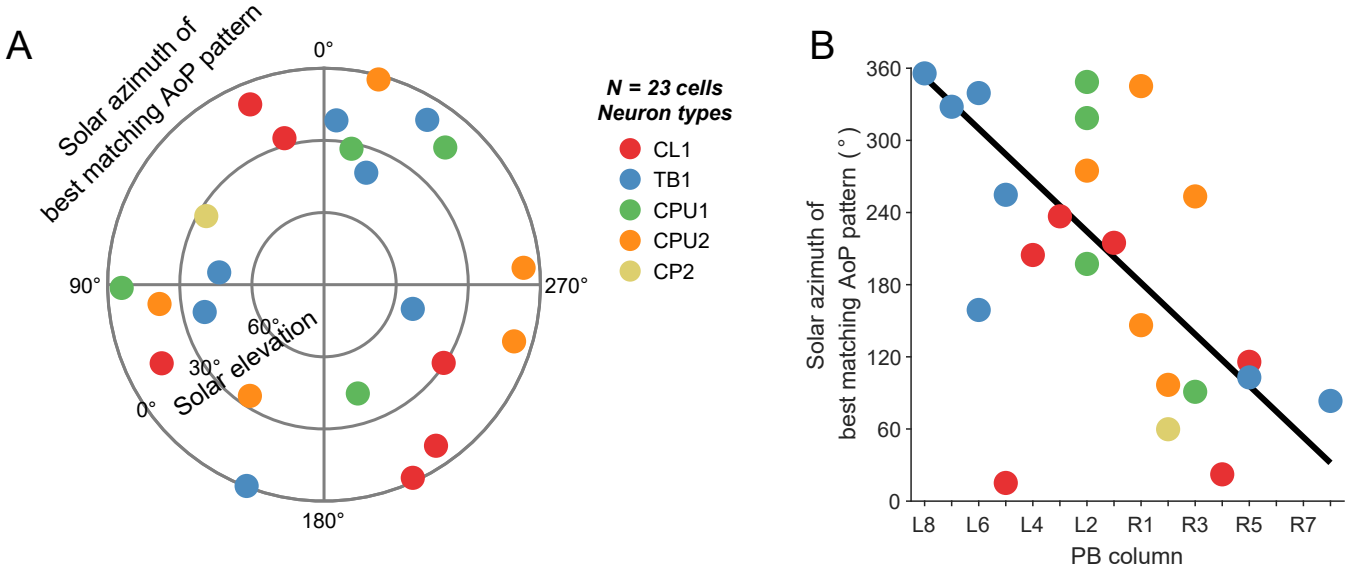

Fig. S8. Compass tuning distribution and tuning regression based on all recorded neurons. This figure shows the same plot types and analysis as Fig. 5, but based on all recorded neurons including those for which the bootstrap test was not significant. The statistics metrics are qualitatively independent of the data-population-size reduction due to bootstrap-testing (cf. Fig. 5). *(A)* Polar plot of the solar positions that correspond to the best-matching polarization patterns of single cells. Azimuth data are uniformly distributed (Rayleigh test, $P = 0.88$, $N = 23$). *(B)* Regression diagram of the column of arborization in the PB vs. the azimuthal component of the solar coordinates in *A*. The azimuth depends on PB column (circular-linear correlation, $\rho = -0.51$, $P = 0.017$). The estimated regression covers nearly 360° along the PB ($f(x) = -21.4x + 352.5$; column L8 corresponds to $x = 0$).

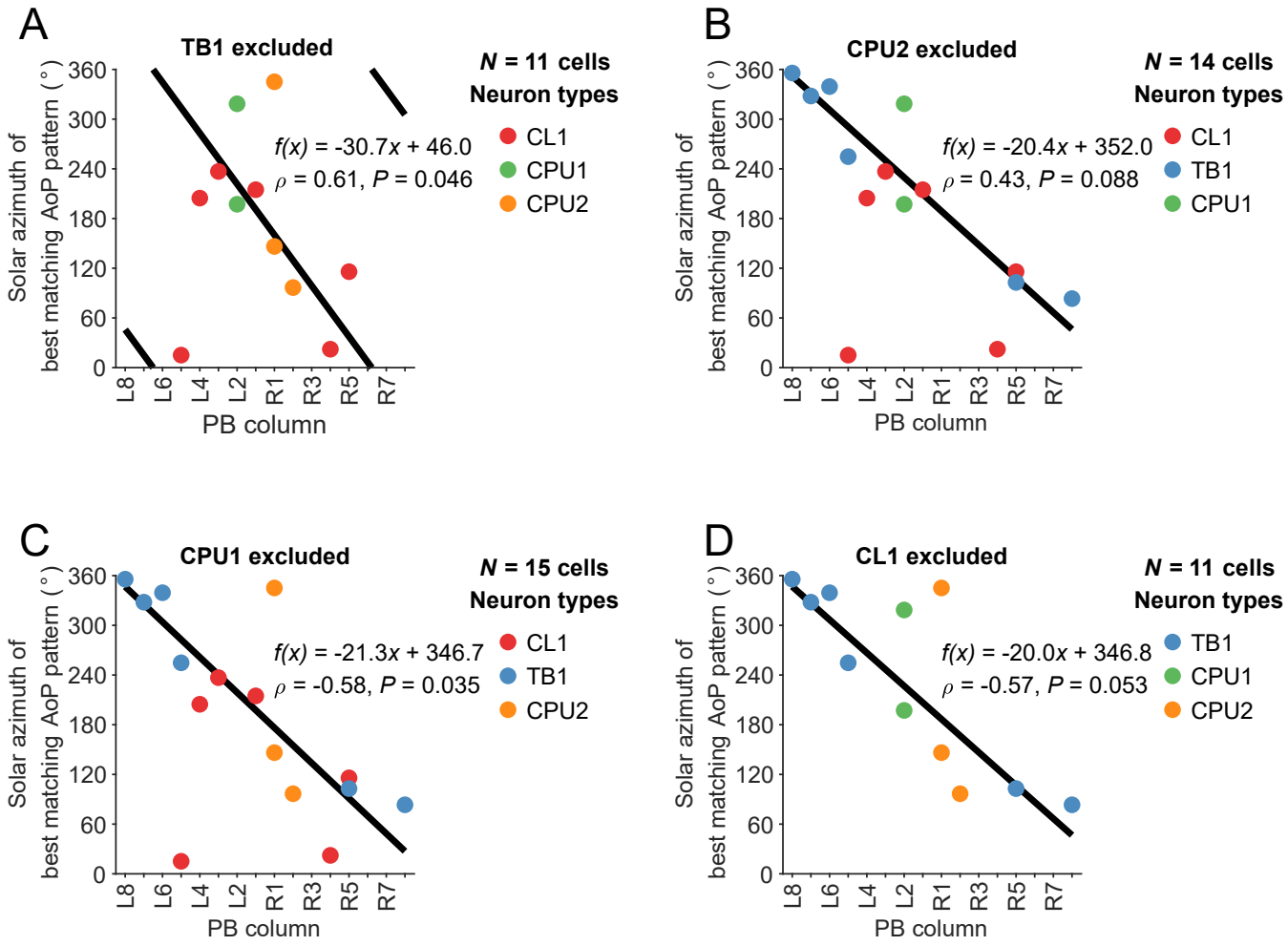

Fig. S9. Robustness analysis of the compass regression. The regression parameters are qualitatively robust against the exclusion of any single cell type. The same circular-linear regression analysis (3) as shown in Fig. 5*B*, but each done with one of the four cell types excluded. *(A)* Excluding TB1 neurons increased the steepness of the regression slope and leveraged the regression intercept accordingly, but did not affect the sign of the slope and the correlation statistics. *(B–D)* Exclusion of any other single cell type did not have a qualitative effect on the regression slope and offset and on the correlation coefficient ρ.

References

1. M Bech, U Homberg, K Pfeiffer, Receptive fields of locust brain neurons are matched to polarization patterns of the sky. *Curr. Biol.* **24**, 2124–2129 (2014).
2. P Berens, CircStat: A MATLAB toolbox for circular statistics. *J. Stat. Softw.* **31**, 1–21 (2009).
3. R Kempter, C Leibold, G Buzsaki, K Diba, R Schmidt, Quantifying circular-linear associations: hippocampal phase precession. *J. Neurosci. Methods* **207**, 113–24 (2012).