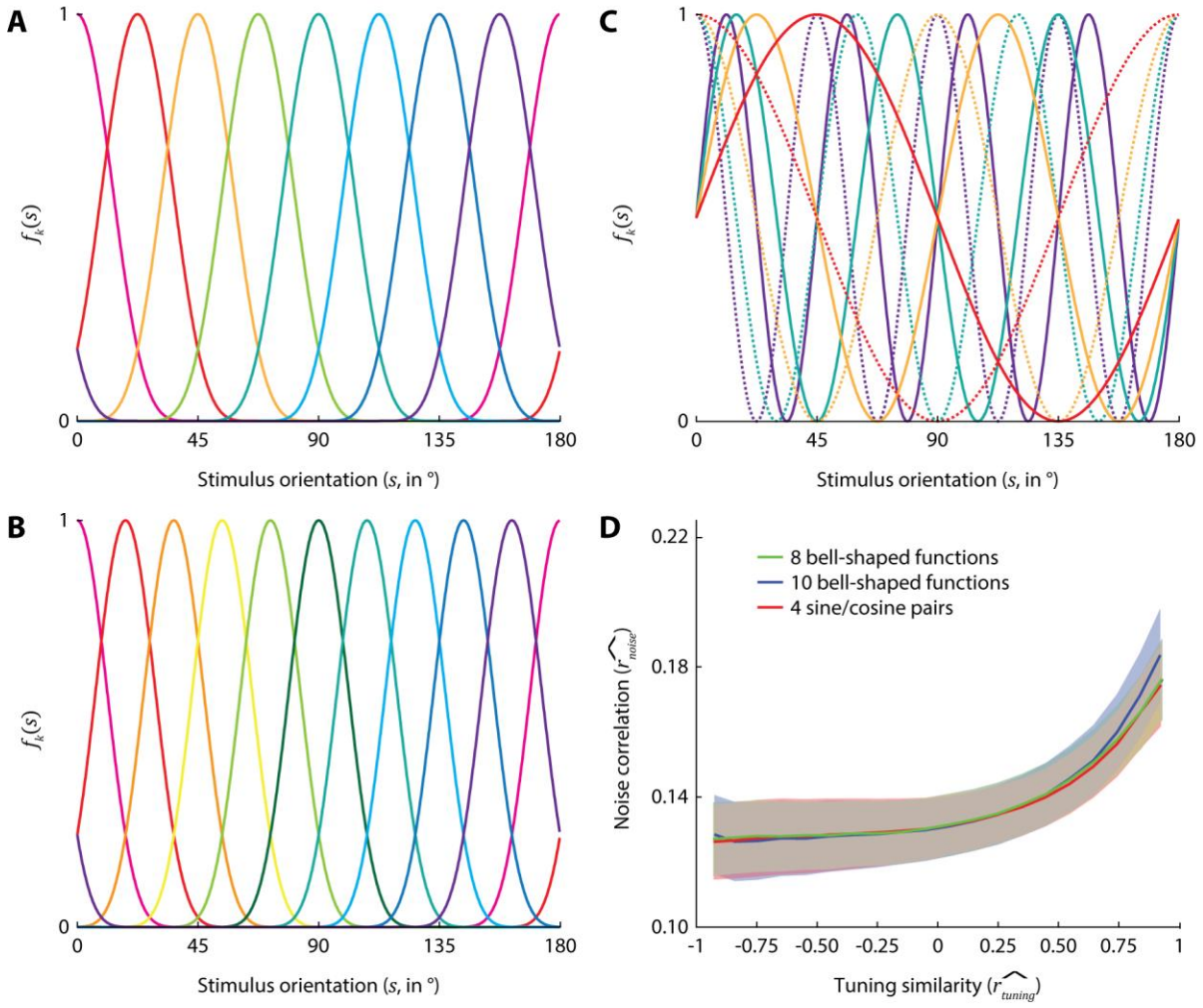
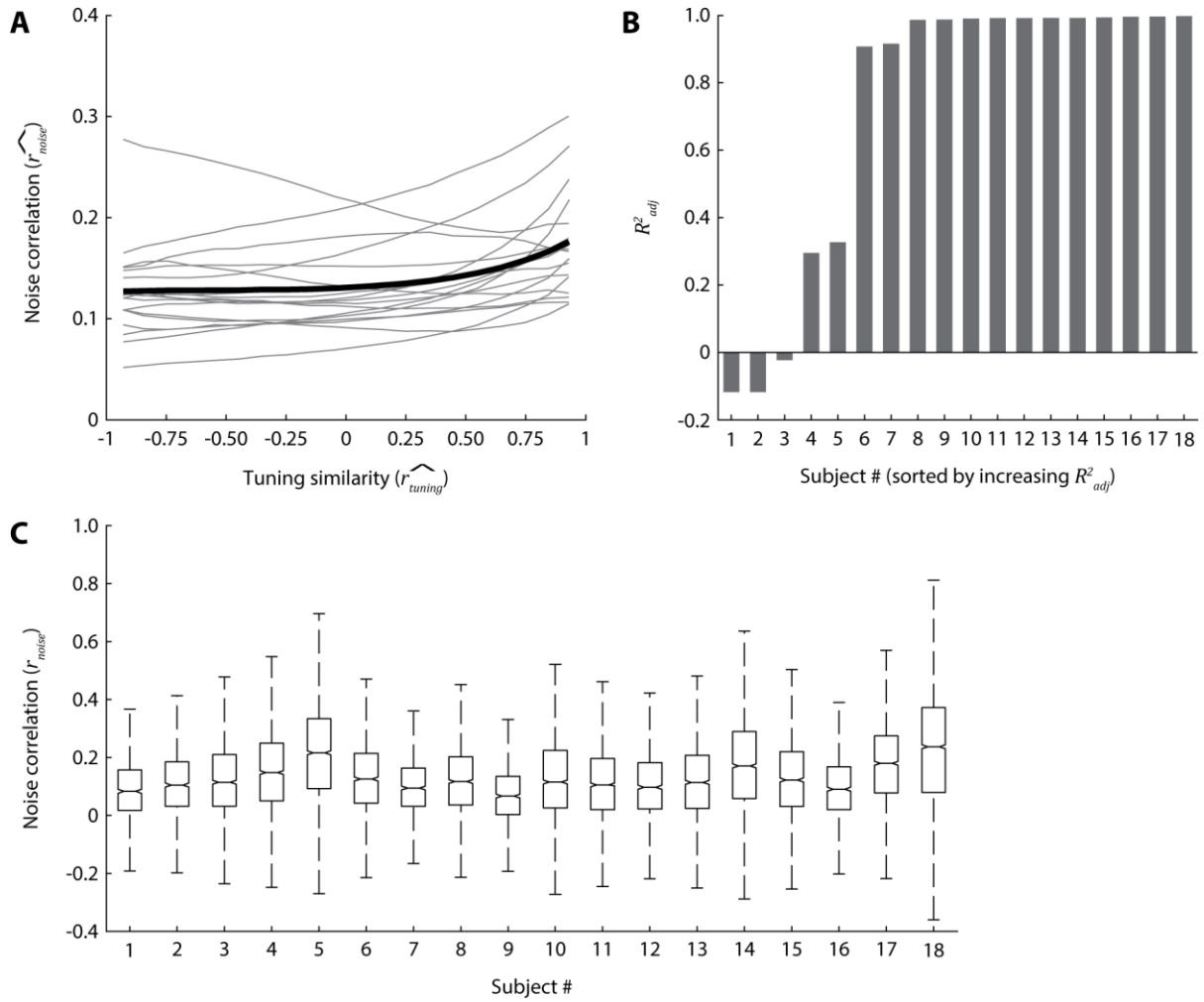


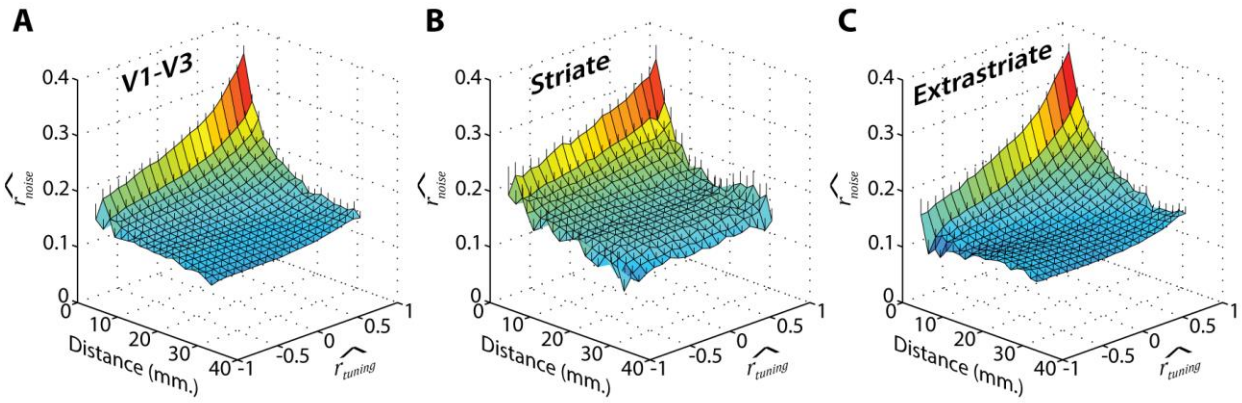
Supplementary figure 1: Estimated tuning similarity and noise correlations are robust to different choices of orientation basis functions used to fit voxel tuning curves.

(**A-C**) Three different basis sets used to fit voxel tuning curves. The value $f_k(s)$ of each (k-th) basis function is plotted against stimulus orientation (s). (**A**) The default basis set of eight bell-shaped functions (zero-rectified cosines, raised to the fifth power) used in the analyses presented in the main text. (**B**) A more expressive basis set of ten bell-shaped functions (zero-rectified cosines, raised to the seventh power). (**C**) An uncorrelated basis of four sines and four cosines, with periods of 180°, 90°, 45° and 25°. Sine/cosine pairs with the same period are shown in the same color, in solid/dotted lines, respectively. This basis spans the same space as the default set shown in **A**, but with orthogonal basis vectors. (**D**) Voxel tuning curves were fit as a linear combination of basis functions in each of the three basis sets show in **A-C**, and tuning similarity and noise correlations between pairs of voxels were calculated as in the main analysis (see **Methods section 2.8.1**). Lines and shaded regions indicate the mean +/- SEM noise correlation across subjects, in 20 equally spaced tuning similarity bins, for the three different basis sets (cf. **Fig. 2A**). Results are virtually identical regardless of which basis set is used to describe voxel tuning, indicating that slight correlations between adjoining functions in our chosen basis set only marginally affect the analyses, and that this basis provides sufficiently fine-grained coverage of orientation space.



Supplementary figure 2: Inter-subject and inter-voxel variability in noise correlations.

(**A**) Average noise correlations ($\widehat{r_{noise}}$) plotted against average tuning similarity ($\widehat{r_{tuning}}$), in 20 equally space tuning similarity bins, for individual subjects (gray lines) and averaged across participants (black line; cf. **Fig. 2A**). (**B**) Goodness-of-fit per subject for the exponential decay function fits to the data in (**A**), quantified by means of the overfitting-adjusted coefficient of determination (R^2_{adj}). Subjects are sorted in order of increasing goodness-of-fit. For the majority subjects (13/18), these data were extremely well fit by an exponential decay with decreasing tuning similarity, with overfitting-adjusted R^2-values in excess of 0.9. Fits were moderately good for another two subjects (R^2_{adj} around 0.3), and below 0 for the remaining three (R^2-values below 0 arise after adjustment for overfitting). (**C**) Box plots showing the distribution of noise correlations across voxel pairs, per subject. Boxes extend from the first to the third quartiles, with notches indicating the medians. Whiskers span the range between the most extreme data points that were not selected as outliers. Outliers were defined as values deviating from the mean by more than 1.5 times the inter-quartile range. For figure clarity, outliers are not shown, due to the large number of data (+/- 2 million voxel pairs per subject). Note that subject order here is not the same as in (**B**).



Supplementary figure 3: Shared noise depends on both orientation tuning and inter-voxel distance.

Given that neighboring voxels in cortex tend to have similar orientation tuning preferences (Freeman et al., 2011; Mannion et al., 2010; Sasaki et al., 2006; Swisher et al., 2010), do noise correlations arise because of noise that is shared between neighboring voxels (e.g., due to the BOLD point spread function (Parkes et al., 2005) or other causes of spatially correlated noise; see e.g. Arcaro et al. (2015), Henriksson et al. (2015), Murphy et al. (2013), and Power et al., (2012)), rather than noise that depends on tuning *per se*? To address this question, voxels pairs were sorted into 20 x 20 bins of similar inter-voxel distance and tuning similarity ($r_{\widehat{tuning}}$). Within each bin, the mean noise correlation, tuning similarity and distance across all pairs of voxels was calculated. This produced, for each observer, a three-dimensional noise correlation surface. Panels (**A**-**C**) show the average surface across subjects for V1-V3 combined, striate (V1) and extrastriate (V2-V3) cortex, respectively. Group-average surfaces were calculated by taking the Fisher-transform of the surface for each subject, averaging across subjects, and then transforming these averages back to the correlation scale. Lines protruding vertically out of the mesh surfaces indicate + 1 SEM.

To quantify the degree to which noise correlations can be explained by spatial distance and/or tuning similarity, noise correlation surfaces obtained for each participant were fitted with each of four models. The models described noise correlations as a function of decreasing tuning similarity (model 1), increasing distance (model 2) or both (models 3 and 4). More specifically, model 1 describes an exponential decay in noise correlations with decreasing tuning similarity (cf. equation (17)):

$$h(m,n) = \alpha \exp\left(-\beta\left(1 - \langle r_{\widehat{tuning}}\rangle_{m,n}\right)\right) + \gamma \quad \text{(S1)}$$

where $h(m,n)$ is the predicted noise correlation for voxel pairs in the (m,n)-th bin (m indexes bins of different tuning similarity and n enumerates bins of different inter-voxel distance), $\langle r_{\widehat{tuning}}\rangle_{m,n}$ is the mean (estimated) tuning similarity between voxel pairs in the (m,n)-th bin, β and α control the rate and starting value of the decay, and γ models a constant baseline correlation among all voxels.

Model 2 describes noise correlations as decaying exponentially with increasing inter-voxel distance:

$$h(m,n) = \kappa \exp\left(-\lambda\langle d\rangle_{m,n}\right) + \gamma \quad \text{(S2)}$$

where $\langle d\rangle_{m,n}$ is the average inter-voxel distance in bin (m,n), and λ and κ determine the rate and starting value of the decay.

Model 3 describes noise correlations as decaying with both increasing distance and decreasing tuning similarity:

$$h(m,n) = \alpha \exp\left(-\beta\left(1 - \langle \widehat{r_{tuning}} \rangle_{m,n}\right)\right) + \kappa \exp\left(-\lambda \langle d \rangle_{m,n}\right) + \gamma \quad \text{(S3)}$$

Model 4 describes the decline in noise correlations as an interaction between an increase in distance and a decrease in tuning similarity:

$$h(m,n) = \alpha \exp\left(-\beta\left(1 - \langle \widehat{r_{tuning}} \rangle_{m,n}\right)\right) \exp\left(-\lambda \langle d \rangle_{m,n}\right) + \gamma \quad \text{(S4)}$$

Each of these exponential decay functions were fit to the data by minimizing the squared error between the Fisher-transformed predictions and the Fisher-transformed bin-average noise correlations, and decay amplitudes (α and κ) and rates (β and λ) were constrained to be non-negative. Goodness-of-fit was assessed by calculating the overfitting-adjusted coefficient of determination (R^2_{adj}) for each model and subject, and then a one-tailed Wilcoxon-signed rank test was used to determine whether the average variance explained across subjects ($\langle R^2_{adj} \rangle$) was reliably greater than zero, for each model. Between models, goodness-of-fit was compared using a two-tailed Wilcoxon-signed rank test.

Although tuning similarity by itself explained significant variance in the noise correlation surfaces (model 1: V1-V3 combined: $\langle R^2_{adj} \rangle$ = 0.22, $p < 10^{-4}$; striate cortex: $\langle R^2_{adj} \rangle$ = 0.19, p = 0.041; extrastriate cortex: $\langle R^2_{adj} \rangle$ = 0.27, $p < 10^{-4}$), these data were best described by an interaction between tuning similarity and distance in extrastriate cortex and V1-V3 combined (model 4: extrastriate cortex: $\langle R^2_{adj} \rangle$ = 0.83, $p < 10^{-5}$; V1-V3 combined: $\langle R^2_{adj} \rangle$ = 0.84, $p < 10^{-5}$; comparisons between model 4 and all other models, in both ROIs: all $p < 0.01$). In striate cortex, a combination of tuning and distance better described the data than tuning alone (model 3: $\langle R^2_{adj} \rangle$ = 0.65, $p < 10^{-5}$; model 4: $\langle R^2_{adj} \rangle$ = 0.64, $p < 10^{-5}$; comparisons of models vs. 3 & 4 vs. model 1, both $p < 10^{-5}$), and this combination was marginally better when compared with a model in which noise correlations decayed with increasing distance alone (model 2 vs. 3, p = 0.074; model 2 vs. 4, p = 0.054). Together, these results indicate that shared fMRI response variability in areas V1-V3 depends on both the distance between voxels and their orientation tuning properties.

Please note that, from a *decoding* perspective, precisely how tuning-dependent correlations arise is of lesser importance: as long as tuning-dependent correlations are present in the data, decoding performance will improve when the generative model takes them into account (as exemplified by the simulations in the main text).

References

Arcaro, M.J., Honey, C.J., Mruczek, R.E.B., Kastner, S., Hasson, U., 2015. Widespread correlation patterns of fMRI signal across visual cortex reflect eccentricity organization. Elife 2015, 1–28. doi:10.7554/eLife.03952

Freeman, J., Brouwer, G.J., Heeger, D.J., Merriam, E.P., 2011. Orientation decoding depends on maps, not columns. J. Neurosci. 31, 4792–804. doi:10.1523/JNEUROSCI.5160-10.2011

Henriksson, L., Khaligh-Razavi, S.M., Kay, K., Kriegeskorte, N., 2015. Visual representations are dominated by intrinsic fluctuations correlated between areas. Neuroimage 114, 275–286. doi:10.1016/j.neuroimage.2015.04.026

Mannion, D.J., McDonald, J.S., Clifford, C.W.G., 2010. Orientation anisotropies in human visual cortex. J. Neurophysiol. 103, 3465–3471. doi:10.1152/jn.00190.2010

Murphy, K., Birn, R.M., Bandettini, P.A., 2013. Resting-state fMRI confounds and cleanup. Neuroimage 80,

349–359. doi:10.1016/j.neuroimage.2013.04.001

Parkes, L.M., Schwarzbach, J. V., Bouts, A.A., Deckers, R.H.R., Pullens, P., Kerskens, C.M., Norris, D.G., 2005. Quantifying the spatial resolution of the gradient echo and spin echo BOLD response at 3 Tesla. Magn. Reson. Med. 54, 1465–1472. doi:10.1002/mrm.20712

Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59, 2142–2154. doi:10.1016/j.neuroimage.2011.10.018

Sasaki, Y., Rajimehr, R., Kim, B.W., Ekstrom, L.B., Vanduffel, W., Tootell, R.B.H., 2006. The Radial Bias: A Different Slant on Visual Orientation Sensitivity in Human and Nonhuman Primates. Neuron 51, 661–670. doi:10.1016/j.neuron.2006.07.021

Swisher, J.D., Gatenby, J.C., Gore, J.C., Wolfe, B. a, Moon, C.-H., Kim, S.-G., Tong, F., 2010. Multiscale pattern analysis of orientation-selective activity in the primary visual cortex. J. Neurosci. 30, 325–30. doi:10.1523/JNEUROSCI.4811-09.2010