Supplementary Information

Details about experimental data acquisition and analysis

All procedures complied with the European Communities Council Directive 2010/63/EU and the German Law for Protection of Animals, and were approved by local authorities, following appropriate ethics review.

Animals

Recordings were performed in 4 adult male Ntsr1-Cre mice (3 hemizygous Tg, 1 negative control, median age at first recording session: 24.4 weeks; B6.FVB(Cg)-Tg(Ntsr1-cre)GN220Gsat/Mmcd; MMRRC, #030648-UCD) and 2 (1 male, 1 female) PV-Cre mice (median age: 17.9 weeks; B6.129P2-Pvalb$^{tm1(cre)Arbr}$/J; Jackson Laboratory, #017320).

Surgery

The surgical procedures are described in detail in [1]. In brief: mice were administered an analgesic (Metamizole, 200 mg/kg, sc, MSD Animal Health, Brussels, Belgium) and put under isoflurane anesthesia (5% in oxygen at start, then lowered to 0.5%–2% in oxygen, CP-Pharma, Burgdorf, Germany), the depth of which was constantly monitored. After shaving and disinfecting the scalp, a skin incision was performed and the skull cleaned of any remaining tissue. Upon positioning the head in a skull-flat position, a custom lightweight aluminium head bar with an opening over dLGN and V1 was placed on the skull and fixated using dental cement. For V1 recordings and optogenetic stimulation unrelated to this study in PV-Cre mice, a small craniotomy above V1 was performed and $\sim 0.2\mu$L of pAAV9/1.EF1a.DIO.hChR2(H134R)-eYFP.WPRE.hGH (Addgene, #20298-AAV9) dyed with fast-green (Sigma-Aldrich, St. Louis, USA) was injected through the entire depth of the cortex. In the Ntsr1-Cre mice used for additional V1 and dLGN recordings, a similar craniotomy was performed and $\sim 0.35\mu$L of stGtACR2 (rAAV2/1-pAAV-hSyn1-SIO-stGtACR2-FusionRed, Addgene, #105677) were injected in the infragranular layers of cortex for experiments with suppression of corticothalamic feedback unrelated to the current study. Post-injection, the opening was filled with Kwik-Cast (WPI Germany, Berlin, Germany). Long-term analgesic (Meloxicam, 2 mg/kg, sc, Böhringer Ingelheim, Ingelheim, Germany) was administered and continued to be administered for 3 consecutive days. After at least 1 week of recovery, animals were gradually habituated to the experimental setup, by first handling them and then simulating the experimental procedure. To allow for virus expression, neural recordings started no sooner than 3 weeks after injection. On the day prior to the first day of recording, mice were fully anesthetized using the same procedures as for the initial surgery, and a craniotomy (ca. 1.5 mm^2) was performed over dLGN and/or V1, and re-sealed with Kwik-Cast. As long as the animals did not show signs of discomfort, the long-term analgesic Metacam was administered only once at the end of surgery, to avoid any confounding effect on experimental results. Recordings were performed daily and continued for as long as the quality of the electrophysiological signals remained high.

Experimental setup

Our experimental configuration for *in-vivo* recordings is described in detail in [1]. In brief: mice were head-fixed and could run freely on an air-suspended styrofoam ball while stimuli were presented to the right visual field on a gamma-corrected LCD screen. Extracellular neural signals were recorded with 32-channel silicon probes (Neuronexus, A1x32Edge-5mm-20-177-A32, Ann Arbor, USA) for the 4 Ntsr1-Cre mice, a 32-channel silicon probe for one PV-Cre mouse (A1x32-Edge-5mm-20-177-A32 and A1x32Edge-5mm-20-177-A32), and a 64-channel silicon probe (A1x64-Poly2-6mm-23s-160-A64) for the other PV-Cre mouse. Ball movements were registered at 90 Hz by two optical mice connected to an Arduino-type microcontroller. Eye movements were monitored under infrared light illumination.

For photostimulation of V1 PV+ inhibitory interneurons, an optic fiber (910 μm diameter, Thorlabs, Newton, USA) was coupled to a light-emitting diode (LED, center wavelength 470 nm, M470F1, Thorlabs, Newton, USA) and positioned with a micromanipulator less than 1 mm above the exposed surface of V1. A black

metal foil surrounding the tip of the head bar holder prevented the photostimulation light from reaching the animal's eyes.

Perfusion and histology

After the final recording session, mice were first administered an analgesic (Metamizole, 200 mg/kg, sc, MSD Animal Health, Brussels, Belgium) and following a 30 min wait period were transcardially perfused under deep anesthesia using a cocktail of Medetomidin (0.5 mL/kg), Midazolam (1 mL/kg), and Fentanyl (1 mL/kg) (ip). Perfusion was first done with Ringer's lactate solution followed by 4% paraformaldehyde (PFA) in 0.2 M sodium phosphate buffer (PBS).

To verify recording site and virus expression, we performed histological analyses. Brains were removed, postfixed in PFA for 24 h, and then rinsed with and stored in PBS at 4°C. Slices (40 μm) were cut using a vibrotome (Leica VT1200 S, Leica, Wetzlar, Germany), mounted on glass slides with Vectashield DAPI (Vector Laboratories, Burlingame, USA), and coverslipped. A fluorescent microscope (BX61 Systems Microscope, Olympus, Tokyo, Japan) was used to inspect slices for the presence of yellow fluorescent protein (eYFP) and DiI. Recorded images were processed using FIJI [2], [3].

Stimulus

We used custom software (EXPO, https://sites.google.com/a/nyu.edu/expo/home) to present visual stimuli on a gamma-calibrated liquid crystal display (LCD) monitor (Samsung SyncMaster 2233RZ; mean luminance 50 cd/m^2, 60 Hz) at 25 cm distance to the animal's right eye (spanning $\sim$ 108x66°, small angle approximation). Mice were presented with three 12 min random sequences of briefly flashed (84 ms), full-screen grating stimuli. The random sequences were drawn from 2304 unique gratings covering 12 orientations ($0, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165°$), 8 contrasts (0, 0.04, 0.10, 0.19, 0.30, 0.46, 0.69, 1), 6 spatial frequencies (0.01, 0.02, 0.06, 0.14, 0.33, 0.80 cyc/°) and 4 spatial phases ($0, 90, 180, 270°$). One sequence consisted of 9216 gratings. Between the sequences, a blank gray screen was displayed for 1 min. For V1 recordings in PV-Cre mice expressing ChR2, light pulses (10 Hz, 1 ms pulses) were delivered from the optical fiber during these periods; analyses of the blank screen responses or photostimulation effects were not included in the current study. Typically, the stimulus sequence was presented once per electrode penetration, except in two cases, where the sequence was run twice during one electrode penetration but data from each run was analyzed separately.

Experimental data pre-processing and spike sorting

Wideband extracellular signals were digitized at 30 kHz (Blackrock microsystems, Blackrock Microsystems Europe GmbH, Hannover DE). To obtain single unit activity from extracellular recordings, the open source, Matlab-based, automated spike sorting toolbox Kilosort [4] was used. Resulting clusters were manually refined using Spyke [5], a Python application that allows the selection of channels and time ranges around clustered spikes for realignment, as well as representation in 3D space using dimension reduction (multichannel PCA, ICA, and/or spike time). Exhaustive pairwise comparisons of similar clusters allowed merging of potentially over-clustered units. All further analyses were performed using an SQL data base and a custom-made analysis pipeline programmed in `python` [6] and managed via `datajoint` [7].

Spike waveshape analysis

From the mean waveform of the maximum-response electrode channel of each single unit, the time between trough and peak (rise time) and the half-width at half-height of the peak were calculated. Exploiting the waveshapes of all V1 units processed using the same pipeline (N = 428 from 10 mice), a k-means algorithm was used to cluster the data into 2 populations.

Optogenetic tagging

To test whether our spike waveshape-based classification method was able to identify PV+ interneurons, we

performed a quantitative analysis of optogenetic tagging. We used the stimulus-associated spike latency test (SALT) method [8] to identify PV+ interneurons via short-latency, light-induced spikes. The algorithm compares the latency distribution of the first spike in a 10 ms window after an optical pulse with baseline distributions before the optical pulse via the Jensen-Shannon divergence. We adapted the algorithm to sample baseline distributions during continuous pulse trains and required identified neurons to have low latency ($L < 4$ ms), high reliability ($R > 0.1$) and small jitter ($J < 2$ ms). Using this method, we found that with the exception of one significantly tagged neuron, all putative PV+ interneurons were contained in the cluster with narrow waveshapes.

Laminar location

We used current source density (CSD) analysis [9] for recordings in area V1 to determine the laminar position of electrode contacts. To obtain the LFP, we first down-sampled the signal to 1 kHz before applying a bandpass filter (3–90 Hz, second-order Butterworth filter). We computed the CSD using the iCSD method [10] implemented in `elephant` [(RRID:SCR_003833) 11]. We assigned the base of layer 4 to the contact that was closest to the earliest CSD polarity inversion. The remaining contacts were assigned to layers based on relative layer thickness reported by [12], assuming a thickness of 1.2 mm for mouse visual cortex. Across all V1 sessions with CSD analysis, the resulting distribution of neurons across layers was: L2/3 - 10/131 (7.6%), L4 - 29/131 (22.1%), L5 - 47/131 (35.9%), L6 - 45/131 (34.4%).

Temporal response kernels via reverse correlation

After eliminating units with overall very low firing rate (<0.1 Hz), the probability of a stimulus preceding a spike by a time δt from -50 to 350 ms was computed for each unique grating stimulus. This was done by binning spike times in 1 ms windows and then counting how often a specific stimulus occurred δt before a spike. After normalizing this histogram by dividing by the total number of spikes, the posterior distribution was calculated according to Bayes' theorem by multiplying with the probability of a spike occurring and dividing by the probability of the stimulus occurring:

$$
\begin{aligned}
P(\text{Spike}|\text{Grating}) &= \frac{P(\text{Grating}|\text{Spike}) \cdot P(\text{Spike})}{P(\text{Grating})} \\
&= \frac{\frac{\text{Grating And Spike Bins}}{\text{Total Spike Bins}} \cdot \frac{\text{Total Spike Bins}}{\text{Total Bins}}}{\frac{\text{Grating Bins}}{\text{Total Bins}}} \\
&= \frac{\text{Grating And Spike Bins}}{\text{Grating Bins}}
\end{aligned}
$$

By dividing by the 1 ms bin duration, the resulting probabilities could be directly converted to firing rates and thus give a temporal response kernel for each unique grating stimulus.

Determining visual responsiveness

To eliminate non-responsive or noise-dominated units, the variance of the temporal kernels across stimuli was calculated and tested for non-randomness using the Wald-Wolfowitz-Test (WWT). The WWT uses the distribution of consecutive ones and zeros in a binary sequence, which should follow a normal distribution in a random sequence, to statistically determine randomness of the sequence. The test can be applied to a non-binary sequence by converting it to a binary sequence via a threshold criterion, typically the mean or median. Before applying the WWT, global trends in the sequence should be removed by either filtering or applying an approximation of the derivative. As the response to grating contrast is a robust indicator of visual responsiveness, the analysis was performed using the aggregate variance across grating contrast, which was computed by averaging the kernels across grating orientation, spatial frequency and spatial phase, before computing the variance across the resulting contrast kernels. To further increase signal-to-noise ratio, the partial variance was squared before computing the differences across time to remove any global trends. On the resulting sequence, the WWT was performed using the median as cutoff criterion. Because the WWT can miss narrow peaks, even if they are high, a second WWT was computed on the absolute values of the sequence. A unit was classified as visually responsive if one of the WWT's was significant and

the test statistic of both WWT's was negative, indicating fewer sign changes than expected by chance.

Determining optimal time point

The time point of optimal response was determined via the peak of the summed aggregate variances across stimuli. First, the partial variances were computed for all four stimulus parameters as described above for grating contrast. The resulting partial variances were then summed and the time point of the first peak exceeding half the modulation depth of the result was selected as the optimal response time point δt_{opt} [peak detection using `scipy`, 13]. For the V1 data set, we removed 7 neurons for which the analysis revealed implausibly short optimal response latencies ($\leq$ 25 ms). We based this minimal latency cutoff on the distribution of optimal time points (5 ms bins) of all recorded V1 neurons, and chose as the cutoff the first bin with at least 3 neurons ($>$ 25 ms). For the dLGN neurons, this minimal latency cutoff removed 3 neurons.

Response profiles

In a 20 ms window around δt_{opt}, responses were averaged over grating spatial phase and frequency, resulting in two-dimensional response matrices covering grating orientation and contrast.

Contrast-invariance

Contrast-invariance of a unit was assessed by applying a singular value decomposition (SVD) to its response matrix, separating the SVD's principle component and residual, and computing the Gamma index of spatial autocorrelation on the residual [14]. The Gamma index is computed by computing a similarity matrix for all data points and then masking the similarity matrix with a contiguity matrix that considers data points that share an edge to be neighbors. The index itself is the sum of all entries in the masked similarity matrix. Patterns are detected by randomly shuffling data points and comparing the original index against the resulting distribution. The strength of the first SVD component was assessed as its power in the SVD: the quotient between the squared first singular value and the sum of squares over all singular values. Residual strength was then computed as one minus first component power. The SVD was calculated using `numpy` [15] and the Gamma index using `pysal` [16]. Neurons with a z-scored Gamma index $g_z > 1.96$ and residual strength $p > 0.05$ were classified as contrast-dependent.

Tuning model

Contrast invariant units were fitted with a two-dimensional tuning model consisting of a hyperbolic ratio function with supersaturation parameter [17] and a wrapped Gaussian [18]:

$$r(c, \psi, \theta) = r_0 + (r_{max} - r_0) \cdot \frac{c^n}{c_{50}^{s \cdot n} + c^{s \cdot n}} \cdot \sum_{m=-\infty}^{m=\infty} \exp\left\{\frac{-(\psi - \theta + \pi m)^2}{2\sigma^2}\right\}, \tag{S1}$$

where c is stimulus contrast, ψ is the stimulus orientation, θ the preferred orientation, r_0 the baseline response and r_{max} the peak response of the neuron. The tuning model was fitted to each neuron's response matrix in a least-squared sense using the `scipy.optimize` library. Quality of fit was assessed via the coefficient of determination (r^2) and only units with $r^2 > 0.4$ were used for further analysis. For characterizing the selectivity of neurons to orientation and contrast, we considered the individual tuning components of the fitted 2D tuning model, after normalizing by the maximal response.

Width parameters of orientation tuning functions, σ in our case, scale non-linearly when tuning is either strong or weak, depending on the specific function. When quantitatively analysing data that contains a broad spectrum of tuning, it is thus advisable to use measures that are not distorted by such non-linear scaling. Accordingly, orientation selectivity [19], [20] was quantified as

$$\text{OSI} = \frac{\sqrt{(\sum R_k \sin(2\theta_k))^2 + (\sum R_k \cos(2\theta_k))^2}}{\sum R_k} \tag{S2}$$

where R_k is the response to the kth direction given by θ_k.

Contrast sensitivity was quantified as contrast at half height of the contrast response function. Since some of our recorded contrast response functions did not saturate, even for full contrast, we preferred this measure as opposed to the parameter c_{50} of the hyperbolic ratio function.

For our sample of recorded neurons, the analysis of contrast response functions revealed neurons in both dLGN and V1 whose response was suppressed by contrast (SbC, Fig. **S2**). Considering our dataset of contrast-invariant dLGN and V1 neurons, which were well-fit by our descriptive, separable model, SbC neurons amounted to 19% of the dLGN population, and 19% of the V1 E population, with no SbC neurons in the V1 I population. The function of SbC neurons is not well understood: 1) it seems likely that V1 SbC neurons might constitute both E and I neurons, being contained within the L6 corticothalamic pyramidal cell population [21], the inhibitory vasoactive intestinal peptide-expressing (VIP) population [22], [23] or a so-far minimally studied class of Sncg inhibitory interneurons [23], 2) stimulus selectivity of SbC neurons is broad and their response latencies are long, and 3) SbC neurons currently are not considered in computational models of V1 [24] such as the SSN. In control analyses, we have repeated our inference procedure with dLGN and V1 E populations additionally containing the SbC neurons (data not shown). These analyses revealed that the order of inferred weigths was largely preserved, even when SbC neurons were included in the populations. Given the uncertainty in terms of V1 cell type associated to SbC neurons, and the conceptual difficulties of incorporating into the SSN one or more potentially distinct populations of SbC neurons, we decided to continue our analyses using only neurons whose responses were enhanced by contrast.

To construct population tuning curves for orientation and contrast, individual orientation and contrast tuning curves were averaged, after aligning individual neurons to their preferred orientation. Population contrast-invariance was assessed by applying the above explained SVD and spatial autocorrelation analysis to the population tuning curves.

Statistics

All statistics were performed using functions from `scipy.stats` and `statsmodels`.

Determining populations' responses from the recorded data for the SSN model fit

The hyperbolic ratio function used to describe and quantify the recorded contrast responses imposes a sigmoidal shape on the contrast response. However, the output of the SSN model itself can explain how S-shaped contrast responses arise from a recurrent network wiring. Therefore, to avoid an additional fitting bias, we did not use the hyperbolic ratio function to represent the recorded populations' responses before we fitted the SSN model to the data. To determine E, I and thalamic population responses to a stimulus of orientation ψ, the function

$$R(c_i, \psi - \theta) = r_0 + (r(c_i) - r_0) \cdot \sum_{m=-\infty}^{\infty} \exp\left\{\frac{-(\psi - \theta + \pi m)^2}{2\sigma^2}\right\} \tag{S3}$$

was fitted to the two-dimensional contrast and orientation responses of individual units to determine the contrast response functions $r(c_i)$ at eight contrast values c_i as well as the width σ of orientation tuning curves (see Swindale [18] for justification of wrapped Gaussian fit). Then the responses of the units were aligned such that their preferred orientations θ coincided with 0°. The E, I, and thalamic population contrast responses at each contrast value c_i in (Fig. 3) were computed as an average $r(c_i)$ in the corresponding population. The population orientation tuning widths were computed as an average σ over the corresponding population.

SSN model with two populations, stability of steady states

The two population SSN model is given by the equations

$$\tau_X \cdot \frac{dr_X(t,C)}{dt} + r_X(t,C) = \Big(J_{XE} \cdot r_E(t,C) - J_{XI} \cdot r_I(t,C) + T_{\text{dLGN}}(C) \cdot g_X\Big)_+^n, \quad X \in \{E,I\}. \tag{S4}$$

The steady states $r_E(C)$ and $r_I(C)$ defined by the equations ($dr_X/dt = 0$)

$$r_X(C) = \Big(J_{XE} \cdot r_E(C) - J_{XI} \cdot r_I(C) + T_{\text{dLGN}}(C) \cdot g_X\Big)_+^n, \quad X \in \{E,I\}. \tag{S5}$$

are stable exactly when the inequalities

$$J_{EE} r_E^{1-1/n} - (n \cdot \det J \cdot r_E^{1-1/n} + J_{II}) r_I^{1-1/n} < 1/n \tag{S6}$$

and

$$\tau_E + \tau_I + \tau_E J_{II} n r_I^{1-1/n} - \tau_I J_{EE} n r_E^{1-1/n} > 0 \tag{S7}$$

are fulfilled [25]. To guarantee stability of the fitted firing rates, we incorporated the inequality in Eq. S6 in the parameter inference algorithm. We note that the second inequality in Eq. S7 can always be fulfilled if we choose sufficiently large τ_E and/or small τ_I.

Determining the two-population SSN parameters from contrast responses

The SSN model with the initially unknown parameters J_{XY}, g_X, and n was required to generate stable steady states $r_X(C)$ (Eq. S5 - Eq. S7), which closely approximated the average recorded cortical and thalamic contrast responses (Eq. S3). For each fixed n and eight contrasts C, Eq. 1 represented an over-determined system of 16 linear equations with six unknown connectivity weights J_{XY}, g_X, which always has a unique solution. We called this solution valid, if additionally, the constants J_{XY}, g_X were positive and lead to a stable steady state of the SSN model. We note that the weights computed directly from the average V1 and thalamic contrast responses did not lead to any valid solutions for the exponents n ranging from 1.1 to 5. Therefore, we randomly generated triplets of V1 and thalamic contrast response curves within ±sem error bar areas of the contrast responses and computed corresponding sets of the connectivity weights J_{XY}, g_X as solutions of the over-determined linear system in Eq. 1 for each triplet. Overall, the fraction of valid connectivity weights was less than 0.1% for all n, and was a monotonically increasing function of n with few valid fits found for n close to 1 (Fig. S5A). Since the initial SSN parameters J_{XY}, g_X, and n were computed for random response triplets and not for the average contrast responses, we optimized them to closely approximate the average responses by minimizing the score function

$$\text{Score}_{\text{fit}}(J,g,n) = \frac{1}{F(n)} \sum_{i=1}^{8} \frac{(r_E^{\text{fit}}(C_i) - r_E^{\text{av}}(C_i))^2}{\sigma_E^2(C_i)} + \frac{(r_I^{\text{fit}}(C_i) - r_I^{\text{av}}(C_i))^2}{\sigma_I^2(C_i)} + \frac{(T_{\text{dLGN}}^{\text{fit}}(C_i) - T_{\text{dLGN}}^{\text{av}}(C_i))^2}{\sigma_{\text{dLGN}}^2(C_i)}. \tag{S8}$$

Based on the score function (Eq. S8), the contrast responses with smaller standard deviation σ_X were approximated with higher precision than those with larger standard deviation. We note that as expected, lower firing rates had lower variability in our recordings. We divided the difference between the fit and the recorded average by the fraction of valid fits F as a function of n to reinforce the exponents n leading to a larger fraction of initial valid fits. Each parameter set J_{XY}, g_X, n in the final distribution of 10^3 fits was a parameter set with the best score out of 10^4 optimized randomly generated valid initial fits.

The external input weight g_I exceeds g_E according to experimental connectivity measurements in Ji et al. [26]

We determined the relation between the parameters g_E and g_I based on the connectivity measurements published in Ji et al. [26]. The V1 E neurons received direct thalamic input with the probabilities 15/19 in

layer 2/3, 19/19 in layer 4, 8/8 in layer 5, and 7/9 in layer 6 [26]. The V1 I neurons received direct thalamic input with the probabilities 14/17 in layer 2/3, 15/15 in layer 4, 15/15 in layer 5, and 9/11 in layer 6 [26]. The adjusted peak amplitudes of postsynaptic potentials amounted to $190 \pm 78pA$ for E and $475 \pm 178pA$ for I V1 populations in layer 2/3, $430 \pm 97pA$ for E and $1111 \pm 260pA$ for I cortical populations in layer 4, $190 \pm 73pA$ for E and $596 \pm 178pA$ for I cortical populations in layer 5, and $160 \pm 49pA$ for E and $412 \pm 167pA$ for I cortical populations in layer 6 [26]. In total, the experimentally measured g_I was higher than g_E in all layers 2/3, 4, 5 and 6 (Fig. 4A).

The connectivity weight J_{EE} exceeds the external input weight g_E based on activity recordings in Lien et al. [27]

Lien et al. [27] reported that the upper bound for the contribution of thalamic inputs compared to the total postsynaptic charge of the E cortical neurons was 36 + 2% for full screen, 100% contrast drifting gratings, [see also 28], [29]. We used the upper bound of 38% to estimate an experimentally plausible region for the relation g_E/J_{EE} in Fig. 4B. Here, we assumed that the relative contribution of the thalamic input $g_E \cdot T_{\text{dLGN}}$ to the E population with respect to the total input to E population $J_{EE} \cdot r_E - J_{EI} \cdot r_I + g_E \cdot T_{\text{dLGN}}$ was smaller than 38%. Using the estimate

$$\frac{g_E \cdot T_{\text{dLGN}}}{J_{EE} \cdot r_E + g_E \cdot T_{\text{dLGN}}} < \frac{g_E \cdot T_{\text{dLGN}}}{J_{EE} \cdot r_E - J_{EI} \cdot r_I + g_E \cdot T_{\text{dLGN}}} < 0.38,$$

we computed the approximate upper bound for the relation g_E/J_{EE}

$$\frac{g_E}{J_{EE}} < \frac{0.38}{1 - 0.38} \cdot \frac{r_E}{T_{\text{dLGN}}} \approx 0.55.$$

Here we used the firing rates r_E and T_{dLGN} recorded in our experiments for 100% of contrast.

Connectivity weights J_{XY} computed from the experimental connectivity measurements reviewed in Table S1

The weights of the network connectivity matrix J_{XY} were computed as a product of connection probability (CP), strength of postsynaptic potential (PSP), and the fraction of neurons in the source population with respect to the total number of neurons included in the network model. The data was extracted from experimental sources introduced in the first rows of Table S1 for the layers 2/3, 4, 5, and 6. We note that only two experimental reports contained complete information on both connectivity measures for all four V1 connections in layers 2/3 and 5 [30], [31], and only one source on connectivity measures in layers 4 and 6 [30]. We assumed that our network contained 89% of excitatory and 11% of PV+ neurons, based on the following calculation: The V1 network consists of approximately 80% of excitatory and 20% of inhibitory neurons. A survey of inhibitory subpopulations in V1 [32] reported that PV+ neurons constitute 37, 49, 53, and 42% of inhibitory neurons in layers 2/3, 4, 5, and 6, respectively. 10 out of 131 recorded neurons belonged to the layer 2/3, 29 to layer 4, 47 to layer 5, and 45 to layer 6. We computed the fraction of PV+ neurons in inhibitory population as the weighted percentage of the recorded neurons $(10 \cdot 0.37 + 29 \cdot 0.49 + 47 \cdot 0.53 + 45 \cdot 0.42)/131 \approx 0.47$. Thus, the percentage of PV+ neurons in our cortical network is $0.47 \cdot 20/(80 + 0.47 \cdot 20) \cdot 100\% \approx 11\%$.

The external input weight g_I exceeds g_E in the SSN model if the firing rate r_I increases faster than r_E for low contrasts

We express the thalamic input T_{dLGN} using both equations in Eq. 1

$$T_{\text{dLGN}}(C) = \frac{r_E(C)^{\frac{1}{n}} - J_{EE} \cdot r_E(C) + J_{EI} \cdot r_I(C)}{g_E} = \frac{r_I(C)^{\frac{1}{n}} - J_{IE} \cdot r_E(C) + J_{II} \cdot r_I(C)}{g_I}. \quad \text{(S9)}$$

Eq. S9 implies

$$r_E(C)^{\frac{1}{n}} + (g_E J_{IE}/g_I - J_{EE}) \cdot r_E(C) = g_E/g_I \cdot r_I(C)^{\frac{1}{n}} + (g_E J_{II}/g_I - J_{EI}) \cdot r_I(C). \tag{S10}$$

For low contrasts and small firing rates, the terms r_X are much smaller than $r_X^{\frac{1}{n}}$ and have a negligible impact on Eq. S10 (Fig. S6A). Since $r_I^{\frac{1}{n}} > r_E^{\frac{1}{n}}$, the inequality $g_E < g_I$ has to be satisfied to fulfill Eq. S10 for small contrasts.

The ISN condition for SSN model implies that J_{EI} exceeds the external input weight g_E in the SSN model

Experiments show that the V1 circuit in mouse operates as an inhibition stabilized network [33]. The condition [34]

$$J_{EE} > \frac{1}{n} r_E^{\frac{1}{n}-1} \tag{S11}$$

guaranties that the network given by the SSN model is inhibition stabilized. The 100% of the inferred connectivity weights satisfied the ISN condition Eq. S11 starting from the smallest contrast value of 4%, Fig. S5C. To derive the relation between the weights J_{EI} and g_E, we compute the derivative with respect to the contrast of the first SSN steady state equation in Eq. 1

$$(\frac{1}{n} r_E^{\frac{1}{n}-1} - J_{EE}) \cdot r_E' = -J_{EI} \cdot r_I' + g_E \cdot T_{\text{dLGN}}'. \tag{S12}$$

We observe in our recorded data that for small and intermediate values of the contrast the dLGN contrast response grows faster than the I firing rate r_I meaning that here T_{dLGN}' exceeds r_I' (Fig. S6B). Thus, using the ISN condition Eq. S11 and the SSN steady state equation in Eq. S12 we obtain for these contrast values

$$0 > (\frac{1}{n} r_E^{\frac{1}{n}-1} - J_{EE}) \cdot r_E' > (g_E - J_{EI}) \cdot r_I'. \tag{S13}$$

Since both E and I firing rates of cortical populations increase monotonically with contrast, their derivatives r_E' and r_I' are positive. Thus, the Eq. S13 implies that the connectivity weight J_{EI} exceeds g_E in the SSN model when the SSN satisfies ISN condition and there exist and interval of contrasts, for which the dLGN firing rate grows faster than the I population firing rate as in our activity recordings (Fig. S6B).

The shapes of contrast responses in V1 and dLGN imply relations $J_{EI} < J_{EE}$, $g_E < J_{EE}$, $J_{II} < J_{IE}$ and $g_I < J_{IE}$

Let $f[0,1] \to [0,\infty)$ with $f(0) = 0$ denote the contrast response function. The recorded contrast response functions are typically nonlinear, resembling the logistic function in their shape. They deviate from the line connecting the points 0 and $f(1)$ given by $Cf(1)$ for every contrast C. The function L_f defined by

$$L_f(C) = f(C)/C - f(1) \tag{S14}$$

quantifies how far the contrast response $f(C)$ at a contrast value C deviates from the line which connects the points 0 and $f(1)$. If $L_f(C)$ is positive (negative) at C, the recorded contrast response $f(C)$ lies above (below) the line connecting 0 and $f(1)$.

We use the SSN steady state equations Eq. S9 to obtain

$$\begin{aligned} -\frac{L_{T_{\text{dLGN}}}(C)}{L_{r_E}(C)} g_E + \frac{L_{r_I}(C)}{L_{r_E}(C)} J_{EI} + \frac{L_{r_E^{1/n}}(C)}{L_{r_E}(C)} &= J_{EE}, \\ -\frac{L_{T_{\text{dLGN}}}(C)}{L_{r_E}(C)} g_I + \frac{L_{r_I}(C)}{L_{r_E}(C)} J_{II} + \frac{L_{r_I^{1/n}}(C)}{L_{r_E}(C)} &= J_{IE}. \end{aligned} \tag{S15}$$

If for some C

$$L_{T_{\text{dLGN}}}(C) > 0, \quad L_{r_E}(C), L_{r_I}(C) < 0, \quad L_{T_{\text{dLGN}}}(C) > |L_{r_I}(C)| > |L_{r_E}(C)|, \tag{S16}$$

and in the case of positive $L_{r_E^{1/n}}(C)$ and $L_{r_I^{1/n}}(C)$

$$\min\{g_E, J_{EI}, g_I, J_{II}\} \geq 1, \quad \min\{L_{T_{\text{dLGN}}}(C), |L_{r_I}(C)|\} > \max\{L_{r_I^{1/n}}(C), L_{r_I^{1/n}}(C)\} \tag{S17}$$

we obtain $J_{EI} < J_{EE}$, $g_E < J_{EE}$, $J_{II} < J_{IE}$ and $g_I < J_{IE}$.

The conditions in Eq. S16 and Eq. S17 hold for the contrast values 30% and 46%, while 99.8% of all inferred connectivity weights satisfy Eq. S17.

Contrast invariance constrains connectivity and input profiles

We denote $\phi \equiv \psi - \theta$, $\phi' \equiv \psi - \theta'$ and show that the property of contrast invariance (Eq. 3)

$$R_X(\phi, C) = r_X(C)\widetilde{r}_X(\phi), \quad X \in \{E, I\} \tag{S18}$$

combined with steady state equations Eq. 2 leads to equations in Eq. 4, which relate orientation tuning curves $\widetilde{r}_X$ with connectivity and input profiles W_{XY} and L_X.

The steady-state equations of the extended SSN model in Eq. 2 are given by

$$R_X(\phi, C) = \Big(\int_{-\pi/2}^{\pi/2} W_{XE}(\phi-\phi')R_E(\phi', C)d\phi' - \int_{-\pi/2}^{\pi/2} W_{XI}(\phi-\phi')R_I(\phi', C)d\phi' + T_{\text{dLGN}}(C)L_X(\phi)\Big)_+^n. \tag{S19}$$

We insert the contrast invariant representation of steady states (Eq. S18) into (Eq. S19) and divide Eq. S19 by $\widetilde{r}_X(\phi)$ to obtain

$$r_X(C) = \Big(J_{XE}(\phi)r_E(C) - J_{XI}(\phi)r_I(C) + T_{\text{dLGN}}(C)g_X(\phi)\Big)_+^n, \tag{S20}$$

where $J_{XY}(\phi) = \int_{-\pi/2}^{\pi/2} W_{XY}(\phi - \phi')\widetilde{r}_Y(\phi')d\phi' / (\widetilde{r}_X(\phi))^{1/n}$ and $g_X(\phi) = L_X(\phi)/(\widetilde{r}_X(\phi))^{1/n}$.

Now we show that g_X are constants independent of ϕ, then we show that J_{XY} are constants provided the contrast response functions are not exactly linearly dependent, i.e. r_E and r_I do not satisfy $r_E(C) = a \cdot r_I(C)$ for all contrasts C with some constant a. We prove this statement by contradiction, i.e. we assume there is at least one parameter J_{XY} or g_X such that $J_{XY}(\phi) \neq J_{XY}(\phi')$ or $g_X(\phi) \neq g_X(\phi')$ for some $\phi \neq \phi'$, and derive a contradiction.

First, we show that g_X are independent of ϕ. We substitute $S = T_{\text{dLGN}}(C)$ into Eq. S20

$$\widehat{r}_X(S) = \Big(J_{XE}(\phi)\widehat{r}_E(S) - J_{XI}(\phi)\widehat{r}_I(S) + Sg_X(\phi)\Big)_+^n. \tag{S21}$$

Here, $\widehat{r}_X(S) = r_X(C) = r_X(T_{\text{dLGN}}^{-1}(S))$. Since $\widehat{r}_X$ are non-negative, the content of the bracket on the right side of the equations is positive and we can remove the sign $+$. Next, we apply the exponent $1/n$ to both sides of Eq. S21 to obtain

$$(\widehat{r}_X(S))^{1/n} = J_{XE}(\phi)\widehat{r}_E(S) - J_{XI}(\phi)\widehat{r}_I(S) + Sg_X(\phi). \tag{S22}$$

Now we denote $\widehat{J}_{XY}(\phi) = J_{XY}(\phi) - J_{XY}(\phi')$ and $\widehat{g}_X(\phi) = g_X(\phi) - g_X(\phi')$ and subtract from Eq. S22 the same equation with ϕ substituted by ϕ' to obtain

$$0 = \widehat{J}_{XE}(\phi)\widehat{r}_E(S) - \widehat{J}_{XI}(\phi)\widehat{r}_I(S) + S\widehat{g}_X(\phi). \tag{S23}$$

Next, we compute a derivative of Eq. S23 with respect to S and set $S = 0$. We obtain

$$0 = \widehat{J}_{XE}(\phi)\widehat{r}'_E(S) - \widehat{J}_{XI}(\phi)\widehat{r}'_I(S) + \widehat{g}_X(\phi). \tag{S24}$$

We note that $\widehat{r}'_X(0) = 0$ always holds for the zero steady state of Eq. S21 corresponding to $S = 0$ input. Then we obtain $\widehat{g}_X(\phi) = 0$ from Eq. S24, which by definition implies $g_X(\phi) = g_X(\phi')$ for all ϕ. We have shown that g_E and g_I are constants independent of ϕ.

Now we show when J_{XY} are independent of ϕ. From $\widehat{g}_E(\phi) = \widehat{g}_I(\phi) = 0$, we obtain that Eq. S23 is equivalent to

$$\widehat{r}_E(S)/\widehat{r}_I(S) = \widehat{J}_{EI}(\phi)/\widehat{J}_{EE}(\phi). \tag{S25}$$

Since the left side of Eq. S25 depends only on S and the right side only on ϕ, both sides are equal to the same constant which we denote by a. In particular, this last observation implies exact linear dependence of the contrast response functions: $r_E(C) = a \cdot r_I(C)$.

We have shown that the property of contrast-invariance Eq. 3 restricts the shape of the input functions L_E and L_I and the interaction profiles W_{EE}, W_{EI}, W_{IE} and W_{II} accordingly to the following relations

$$L_E(\phi) = g_E \cdot (\widetilde{r}_E(\phi))^{1/n}, \quad L_I(\phi) = g_I \cdot (\widetilde{r}_I(\phi))^{1/n}, \tag{S26}$$

$$\begin{aligned}
&\int_{-\pi/2}^{\pi/2} W_{EE}(\phi-\phi')\widetilde{r}_E(\phi')d\phi' = J_{EE} \cdot (\widetilde{r}_E(\phi))^{1/n}, \quad \int_{-\pi/2}^{\pi/2} W_{EI}(\phi-\phi')\widetilde{r}_I(\phi')d\phi' = J_{EI} \cdot (\widetilde{r}_E(\phi))^{1/n}, \\
&\int_{-\pi/2}^{\pi/2} W_{IE}(\phi-\phi')\widetilde{r}_E(\phi')d\phi' = J_{IE} \cdot (\widetilde{r}_I(\phi))^{1/n}, \quad \int_{-\pi/2}^{\pi/2} W_{II}(\phi-\phi')\widetilde{r}_I(\phi')d\theta' = J_{II} \cdot (\widetilde{r}_I(\phi))^{1/n},
\end{aligned} \tag{S27}$$

where the constants g_X and J_{XY} depend on the shape of contrast responses r_E and r_I that are steady states of the two population SSN model.

Determining connectivity and input profiles of the extended SSN model

To determine the input and connectivity profiles L_X and W_{XY} from Eq. S26 and Eq. S27, we used the wrapped Gaussian approximation of orientation tuning curves. The wrapped Gaussian function is given by

$$G_{(\sigma)}(\phi) = \sum_{m=-\infty}^{\infty} \exp\left\{-\frac{(\phi+\pi m)^2}{2\sigma^2}\right\}.$$

The widths of the orientation tuning curves $\widetilde{r}_E$ and $\widetilde{r}_I$ were $\sigma_E \approx 0.31\pi \approx 56^\circ$ and $\sigma_I \approx 0.34\pi \approx 62^\circ$ (Fig. 5B), the orientation tuning curves are represented by

$$\widetilde{r}_E(\phi) = G_{(\sigma_E)}(\phi)/\max_{\phi} G_{(\sigma_E)}, \quad \widetilde{r}_I(\phi) = G_{(\sigma_I)}(\phi)/\max_{\phi} G_{(\sigma_I)}. \tag{S28}$$

To derive the input profiles L_E and L_I using Eq. S26, we fitted normalized wrapped Gaussian functions to the power-law transformations of orientation tuning curves $(\widetilde{r}_E)^{1/n}$ and $(\widetilde{r}_I)^{1/n}$, where n were the power-law exponents inferred from the recorded contrast responses. We obtained that the mean widths of the curves

$(\widetilde{r}_E)^{1/n}$ and $(\widetilde{r}_I)^{1/n}$ were $\sigma_E^{\text{inp}} \approx 0.36\pi \approx 65^\circ$ and $\sigma_I^{\text{inp}} \approx 0.39\pi \approx 70^\circ$, respectively, and the input profiles L_E and L_I (Fig. 5C) were represented by

$$L_E(\phi) = g_E \cdot G_{(\sigma_E^{\text{inp}})}(\phi)/\max_\phi G_{(\sigma_E^{\text{inp}})}, \quad L_I(\phi) = g_I \cdot G_{(\sigma_I^{\text{inp}})}(\phi)/\max G_{(\sigma_I^{\text{inp}})}. \tag{S29}$$

Our next goal was to determine the connectivity profiles W_{XY} using Eq. S27. To this end, we used the formula for the convolution of two wrapped Gaussian functions [35]

$$\frac{\sigma_2}{\sqrt{2\pi}\sigma_1(\sigma_2^2-\sigma_1^2)^{1/2}} \int_{-\pi/2}^{\pi/2} G_{((\sigma_2^2-\sigma_1^2)^{1/2})}(\phi-\phi') G_{(\sigma_1)}(\phi') d\phi' = G_{(\sigma_2)}(\phi). \tag{S30}$$

Next, we combined Eq. S27 and Eq. S30 to obtain the wrapped Gaussian representation of W_{XY}. Using

$$(\widetilde{r}_X(\phi))^{1/n} = G_{(\sigma_X^{\text{inp}})}(\phi)/\max_\phi G_{(\sigma_X^{\text{inp}})}, \quad \widetilde{r}_Y(\phi) = G_{(\sigma_Y)}(\phi)/\max_\phi G_{(\sigma_Y)},$$

Eq. S27, and Eq. S30, we obtained

$$W_{XY}(\phi) = \frac{J_{XY} \cdot \sigma_X^{\text{inp}}}{\sqrt{2\pi} \cdot \sigma_Y \cdot ((\sigma_X^{\text{inp}})^2 - \sigma_Y^2)^{1/2}} \cdot \frac{\max_\phi G_{(\sigma_Y)}}{\max_\phi G_{(\sigma_X^{\text{inp}})}} \cdot G_{(((\sigma_X^{\text{inp}})^2-\sigma_Y^2)^{1/2})}(\phi) \tag{S31}$$

and

$$\sigma_{XY} = ((\sigma_X^{\text{inp}})^2 - \sigma_Y^2)^{1/2}. \tag{S32}$$

For our recorded data, we obtained $\sigma_{EE} = 33^\circ$, $\sigma_{EI} = 19^\circ$, $\sigma_{IE} = 42^\circ$, and $\sigma_{II} = 33^\circ$.

Ascending order between the widths of connectivity profiles

The widths of connectivity profiles follow the order $\sigma_{EI} < \sigma_{EE} < \sigma_{II} < \sigma_{IE}$ in each inferred parameter set. However, this result would also follow for a specific order between only σ_E, σ_I and σ_E^{inp}. Here we show that the above derivations constrain the possible order of connectivity widths: independently of the exact values of σ_E and σ_I, the assumptions

$$\sigma_E < \sigma_I, \quad \sigma_I < \sigma_E^{\text{inp}}, \quad n > 1 \tag{S33}$$

always imply the relations

$$\sigma_{EI} < \sigma_{EE} < \sigma_{IE} \quad \sigma_{EI} < \sigma_{II} < \sigma_{IE}.$$

Indeed, since $\sigma_E < \sigma_I$, we always have $\sigma_E^{\text{inp}} < \sigma_I^{\text{inp}}$ for $n > 1$. Based on this inequality and Eq. S33, we obtain $\sigma_E^2 < \sigma_I^2 < (\sigma_E^{\text{inp}})^2 < (\sigma_I^{\text{inp}})^2$. Next, we obtain

$$(\sigma_E^{\text{inp}})^2 - \sigma_I^2 < (\sigma_E^{\text{inp}})^2 - \sigma_E^2 < (\sigma_I^{\text{inp}})^2 - \sigma_E^2$$

and

$$(\sigma_E^{\text{inp}})^2 - \sigma_I^2 < (\sigma_I^{\text{inp}})^2 - \sigma_I^2 < (\sigma_I^{\text{inp}})^2 - \sigma_E^2.$$

Together with Eq. S32, these chains of inequalities are equivalent to the relations $\sigma_{EI} < \sigma_{EE} < \sigma_{IE}$ and $\sigma_{EI} < \sigma_{II} < \sigma_{IE}$. In particular, inequalities in Eq. S33 always imply that σ_{EI} is the smallest and σ_{IE} the largest connectivity profile width, while σ_{EE} and σ_{II} are constrained between σ_{EI} and σ_{IE}.

The widths of the connectivity profiles need to follow the rank $\sigma_{EI} < \sigma_{EE} \approx \sigma_{II} < \sigma_{IE}$ to balance out the widths of the inputs to the populations in the recurrent network and achieve that the ultimate widths of populations' activity profiles coincide with the observed ones. Precisely, the connectivity profile modifies the width of the orientation-dependent activity profile R_Y of the population Y, together they shape the input from the population Y to the population X. The total input to the population X, in the recurrent network

represented by the SSN model, is the sum of orientation-dependent inputs from E, I, and dLGN populations, Eq. 2

$$\int W_{XE} \cdot R_E - \int W_{XI} \cdot R_I + T_{dLGN} \cdot L_X.$$

This input is then transformed by the neuronal nonlinearity (power-law with the exponent $n \approx 2$) and results in the orientation-dependent activity profile R_X of the population X. Hereby, the width of each orientation-dependent term ($\int W_{XE}\tilde{r}_E$, $\int W_{XI}\tilde{r}_I$ and L_X) in the sum needs to be approximately equal to the width of the activity profile $\tilde{r}_X^{1/n}$. We show this in Eq. 4 using the contrast-invariance of orientation-dependent response.

The parameter σ of the wrapped Gaussian fit describes the width of orientation-dependent activity and connectivity profiles, Eq. **S3**. The wrapped Gaussian fit conveniently represents orientation-dependent activity and connectivity profiles because an analytical formula relates the widths σ of all the profiles in Eq. 4. Precisely, the relation $\int W_{XY} \cdot \tilde{r}_Y = J_{XY}\tilde{r}_X^{1/n}$ implies that $\sigma_{XY}^2 + \sigma_Y^2 = \sigma_{X,1/n}^2$, whereby σ_{XY}, σ_Y, and $\sigma_{X,1/n}$ are the widths of the profiles W_{XY}, $\tilde{r}_Y$, and $\tilde{r}_{X,1/n}$. In particular, $\sigma_{EE}^2 + \sigma_E^2 = \sigma_{E,1/n}^2$ and $\sigma_{EI}^2 + \sigma_I^2 = \sigma_{E,1/n}^2$. Both expressions are equal to the same number $\sigma_{E,1/n}^2$. However, σ_E is smaller than σ_I, meaning that σ_{EI} has to be smaller than σ_{EE}. Analogously, $\sigma_{IE}^2 + \sigma_E^2 = \sigma_{I,1/n}^2$ and $\sigma_{II}^2 + \sigma_I^2 = \sigma_{I,1/n}^2$. Both expressions are equal to the same number $\sigma_{I,1/n}$ which is larger than $\sigma_{E,1/n}$. Thus, together with $\sigma_E < \sigma_I$, we obtain that σ_{II} is smaller than σ_{IE}, but both are larger than σ_{EI} and σ_{EE}, respectively, because $\sigma_{I,1/n}$ is larger than $\sigma_{E,1/n}$.

References

[1] M. A. Spacek, G. Born, D. Crombie, Y. Bauer, X. Liu, S. Katzner, and L. Busse, "Robust effects of corticothalamic feedback during naturalistic visual stimulation," *bioRxiv*, Jun. 2021. doi: 10.1101/776237.

[2] C. T. Rueden, J. Schindelin, M. C. Hiner, B. E. DeZonia, A. E. Walter, E. T. Arena, and K. W. Eliceiri, "Imagej2: Imagej for the next generation of scientific image data," *BMC Bioinformatics*, vol. 18, no. 1, p. 529, 2017.

[3] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, *et al.*, "Fiji: An open-source platform for biological-image analysis," *Nature Methods*, vol. 9, no. 7, pp. 676–682, 2012.

[4] M. Pachitariu, N. A. Steinmetz, S. N. Kadir, M. Carandini, and K. D. Harris, "Fast and accurate spike sorting of high-channel count probes with KiloSort," *Advances in Neural Information Processing Systems*, vol. 29, no. Nips, pp. 4448–4456, 2016, issn: 10495258.

[5] M. A. Spacek, T. Blanche, and N. Swindale, "Python for large-scale electrophysiology," *Frontiers in Neuroinformatics*, vol. 2, p. 9, 2009, issn: 1662-5196. doi: 10.3389/neuro.11.009.2008.

[6] G. Van Rossum and F. L. Drake, *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009, isbn: 1441412697.

[7] D. Yatsenko, E. Y. Walker, and A. S. Tolias, "Datajoint: A simpler relational data model," *arXiv preprint arXiv:1807.11104*, 2018.

[8] D. Kvitsiani, S. Ranade, B. Hangya, H. Taniguchi, J. Z. Huang, and A. Kepecs, "Distinct behavioural and network correlates of two interneuron types in prefrontal cortex," *Nature*, vol. 498, no. 7454, pp. 363–366, 2013, issn: 00280836. doi: 10.1038/nature12176.

[9] U. Mitzdorf, "Current source-density method and application in cat cerebral cortex: Investigation of evoked potentials and EEG phenomena," *Physiological Reviews*, vol. 65, no. 1, pp. 37–100, 1985.

[10] K. H. Pettersen, A. Devor, I. Ulbert, A. M. Dale, and G. T. Einevoll, "Current-source density estimation based on inversion of electrostatic forward solution: Effects of finite extent of neuronal activity and conductivity discontinuities," *Journal of Neuroscience Methods*, vol. 154, no. 1-2, pp. 116–133, 2006, issn: 01650270. doi: 10.1016/j.jneumeth.2005.12.005.

[11] M. Denker, A. Yegenoglu, and S. Grün, "Collaborative HPC-enabled workflows on the HBP Collaboratory using the Elephant framework," in *Neuroinformatics 2018*, 2018, P19. doi: 10.12751/incf.ni2018.0019.

[12] D. Heumann, G. Leuba, and T. Rabinowicz, "Postnatal development of the mouse cerebral neocortex. II. Quantitative cytoarchitectonics of visual and auditory areas," *Journal Fur Hirnforschung*, vol. 18, no. 6, pp. 483–500, 1977, issn: 0021-8359.

[13] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, *et al.*, "Scipy 1.0: Fundamental algorithms for scientific computing in python," *Nature Methods*, vol. 17, no. 3, pp. 261–272, 2020.

[14] L. J. Hubert, R. G. Golledge, and C. M. Costanzo, "Generalized procedures for evaluating spatial autocorrelation," *Geographical Analysis*, vol. 13, no. 3, pp. 224–233, 1981.

[15] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020. doi: 10.1038/s41586-020-2649-2.

[16] S. J. Rey and L. Anselin, "PySAL: A Python Library of Spatial Analytical Methods," *The Review of Regional Studies*, vol. 37, no. 1, pp. 5–27, 2007.

[17] J. W. Peirce, "The potential importance of saturating and supersaturating contrast response functions in visual cortex," *Journal of Vision*, vol. 7, no. 6, pp. 1–10, 2007, issn: 15347362. doi: 10.1167/7.6.13.

[18] N. V. Swindale, "Orientation tuning curves: Empirical description and estimation of parameters," *Biological Cybernetics*, vol. 78, no. 1, pp. 45–56, 1998, issn: 03401200. doi: 10.1007/s004220050411.

[19] T. Bonhoeffer, D.-S. Kim, D. Malonek, D. Shoham, and A. Grinvald, "Optical Imaging of the Layout of Functional Domains in Area 17 and Across the Area 17/18 Border in Cat Visual Cortex," *European Journal of Neuroscience*, vol. 7, no. 9, pp. 1973–1988, 1995.

[20] S. R. Olsen, D. S. Bortone, H. Adesnik, and M. Scanziani, "Gain control by layer six in cortical circuits of vision," *Nature*, vol. 483, no. 7387, pp. 47–54, 2012, issn: 00280836. doi: 10.1038/nature10835.

[21] S. Augustinaite and B. Kuhn, "Complementary Ca2+ Activity of Sensory Activated and Suppressed Layer 6 Corticothalamic Neurons Reflects Behavioral State," *Current Biology*, vol. 30, no. 20, pp. 3945–3960, Aug. 2020, issn: 0960-9822. doi: 10.1016/j.cub.2020.07.069.

[22] D. J. Millman, G. K. Ocker, S. Caldejon, I. Kato, J. D. Larkin, E. K. Lee, J. Luviano, C. Nayan, T. V. Nguyen, K. North, S. Seid, C. White, J. Lecoq, C. Reid, M. A. Buice, and S. E. de Vries, "VIP interneurons in mouse primary visual cortex selectively enhance responses to weak but specific stimuli," *eLife*, vol. 9, M. Vinck and K. M. Wassum, Eds., e55130, Oct. 2020, issn: 2050-084X. doi: 10.7554/eLife.55130.

[23] S. Bugeon, J. Duffield, M. Dipoppa, A. Ritoux, I. Prankerd, D. Nicoloutsopoulos, D. Orme, M. Shinn, H. Peng, H. Forrest, A. Viduolyte, C. B. Reddy, Y. Isogai, M. Carandini, and K. D. Harris, "A transcriptomic axis predicts state modulation of cortical interneurons," *Nature*, 2022, issn: 1476-4687. doi: 10.1038/s41586-022-04915-7.

[24] R. H. Masland and P. R. Martin, "The unsolved mystery of vision," *Current Biology*, vol. 17, no. 15, R577–R582, Aug. 2007, issn: 0960-9822. doi: 10.1016/j.cub.2007.05.040.

[25] N. Kraynyukova and T. Tchumatchenko, "Stabilized supralinear network can give rise to bistable, oscillatory, and persistent activity," *PNAS*, vol. 115, no. 13, pp. 3464–3469, 2018.

[26] X.-y. Ji, B. Zingg, L. Mesik, Z. Xiao, L. I. Zhang, and H. W. Tao, "Thalamocortical innervation pattern in mouse auditory and visual cortex: Laminar and cell-type specificity," *Cerebral Cortex*, vol. 26, no. 6, pp. 2612–2625, 2016.

[27] A. D. Lien and M. Scanziani, "Tuned thalamic excitation is amplified by visual cortical circuits," *Nature Neuroscience*, vol. 16, no. 9, pp. 1315–1323, 2013.

[28] K. Reinhold, A. D. Lien, and M. Scanziani, "Distinct recurrent versus afferent dynamics in cortical visual processing," *Nature Neuroscience*, vol. 18, no. 12, pp. 1789–1797, Dec. 2015, issn: 1546-1726. doi: 10.1038/nn.4153.

[29] Y. T. Li, L. A. Ibrahim, B. H. Liu, L. I. Zhang, and H. W. Tao, "Linear transformation of thalamocortical input by intracortical excitation," *Nature Neuroscience*, vol. 16, no. 9, pp. 1324–1330, 2013, issn: 10976256. doi: 10.1038/nn.3494.

[30] Allen Institute for Brain Science, *Synaptic physiology coarse matrix dataset*, Available from: https://brain-map.org/explore/connectivity/synaptic-physiology, 2019.

[31] X. Jiang, S. Shen, C. R. Cadwell, P. Berens, F. Sinz, A. S. Ecker, S. Patel, and A. S. Tolias, "Principles of connectivity among morphologically defined cell types in adult neocortex," *Science*, vol. 350, no. 6264, aac9462, Nov. 2015, issn: 0036-8075, 1095-9203. doi: 10.1126/science.aac9462.

[32] Y. Gonchar, Q. Wang, and A. H. Burkhalter, "Multiple distinct subtypes of gabaergic neurons in mouse visual cortex identified by triple immunostaining," *Frontiers in neuroanatomy*, vol. 2, p. 3, 2008.

[33] A. Sanzeni, B. Akitake, H. C. Goldbach, C. E. Leedy, N. Brunel, and M. H. Histed, "Inhibition stabilization is a widespread property of cortical networks," *eLife*, vol. 9, T. O'Leary, J. Huguenard, and H. Adesnik, Eds., e54875, Jun. 2020, issn: 2050-084X. doi: 10.7554/eLife.54875.

[34] K. D. Miller and A. Palmigiano, "Generalized paradoxical effects in excitatory/inhibitory networks," *bioRxiv*, 2020.

[35] K. Mardia and P. Jupp, *Directional Statistics*, ser. Wiley Series in Probability and Statistics. Wiley, 2009, isbn: 9780470317815.

[36] S. B. Hofer, H. Ko, B. Pichler, J. Vogelstein, H. Ros, H. Zeng, E. Lein, N. A. Lesica, and T. D. Mrsic-Flogel, "Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex," *Nature Neuroscience*, vol. 14, no. 8, pp. 1045–52, 2011. doi: 10.1038/nn.2876.

[37] H. Ko, S. B. Hofer, B. Pichler, K. A. Buchanan, P. J. Sjöström, and T. D. Mrsic-Flogel, "Functional specificity of local synaptic connections in neocortical networks," *Nature*, vol. 473, no. 7345, pp. 87–91, 2011, issn: 00280836. doi: 10.1038/nature09880.

[38] L. Cossell, M. F. Iacaruso, D. R. Muir, R. Houlton, E. N. Sader, H. Ko, S. B. Hofer, and T. D. Mrsic-Flogel, "Functional organization of excitatory synaptic strength in primary visual cortex," *Nature*, vol. 518, no. 7539, pp. 399–403, 2015, issn: 14764687. doi: 10.1038/nature14182.

[39] S. C. Seeman, L. Campagnola, P. A. Davoudian, A. Hoggarth, T. A. Hage, A. Bosma-Moody, C. A. Baker, J. H. Lee, S. Mihalas, C. Teeter, *et al.*, "Sparse recurrent excitatory connectivity in the microcircuit of the adult mouse and human cortex," *Elife*, vol. 7, e37349, 2018.

[40] M. M. Karnani, J. Jackson, I. Ayzenshtat, J. Tucciarone, K. Manoocheri, W. G. Snider, and R. Yuste, "Cooperative subnetworks of molecularly similar interneurons in mouse neocortex," *Neuron*, vol. 90, no. 1, pp. 86–100, 2016.

[41] J. A. Mazer, W. E. Vinje, J. McDermott, P. H. Schiller, and J. L. Gallant, "Spatial frequency and orientation tuning dynamics in area V1," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 3, pp. 1645–1650, 2002, issn: 00278424. doi: 10.1073/pnas.022638499.

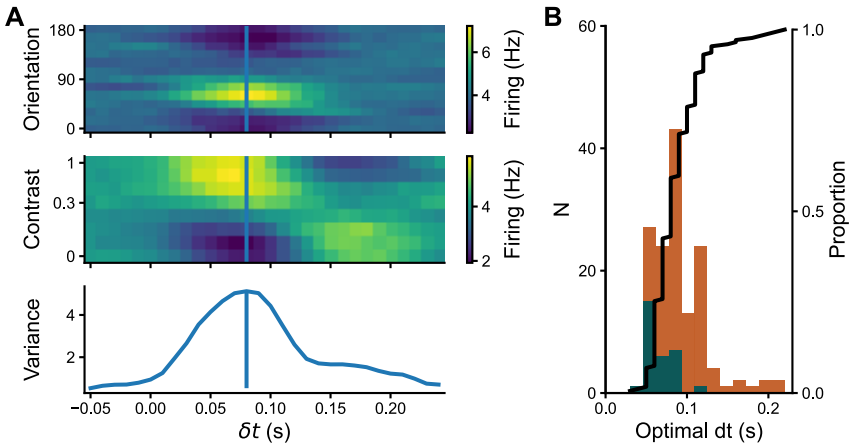

Fig. S1. Reverse correlation and optimal response time point (related to Fig. 2). **A**, Reverse correlation analysis for an example V1 neuron. Reverse correlation computes the firing rate at a time point δt relative to stimulus occurrence, yielding temporal kernels for each stimulus combination (*Middle*: orientation; *bottom*: contrast; average across all other stimulus dimensions for visualization). The optimal response time was calculated by using the sum of the aggregated variances in firing rate across stimulus conditions (*top*, see Supplementary Information) and selecting its peak as the latency of optimal response (*vertical line*) [41]. **B**, Distribution of optimal response times (after application of a latency cutoff of > 25 ms, which removed 7 neurons with implausibly short latencies) (*teal*: inhibitory, *orange*: excitatory, *black*: cumulative distribution). Inhibitory neurons had a significantly lower optimal response time (69 ± 3 ms) than excitatory neurons (89 ± 3 ms; mean $\pm$ sem; two-tailed Welch's t-Test: t=5.18, $p < .001$; 174 neurons.

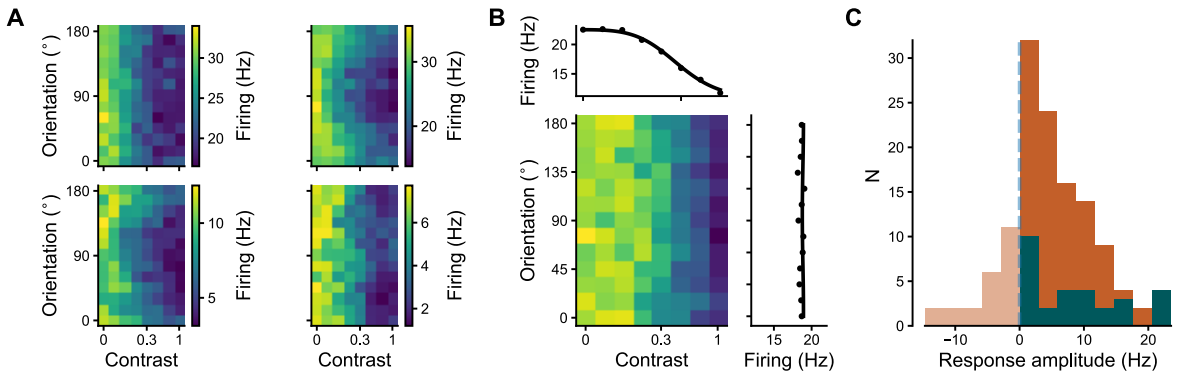

Fig. S2. Suppressed-by-contrast neurons (SbC) (related to Fig. 2). **A**, Four example SbC neurons, characterized by their stronger response to low contrast and an increasing suppression with higher contrasts. **B**, Two-dimensional tuning model fitted to an SbC neuron. **C**, Distribution of response amplitudes, defined as the difference between responses to 100% and 0% contrast. SbC neurons, defined by their negative response amplitude, were found within the broad-spiking, putative excitatory neurons (24/125, *transparent*). Since SbC neurons might correspond to VIP interneurons [22], [23] or a so-far minimally studied class of Sncg interneurons [23], SbC neurons were excluded from further analysis (see also Results).

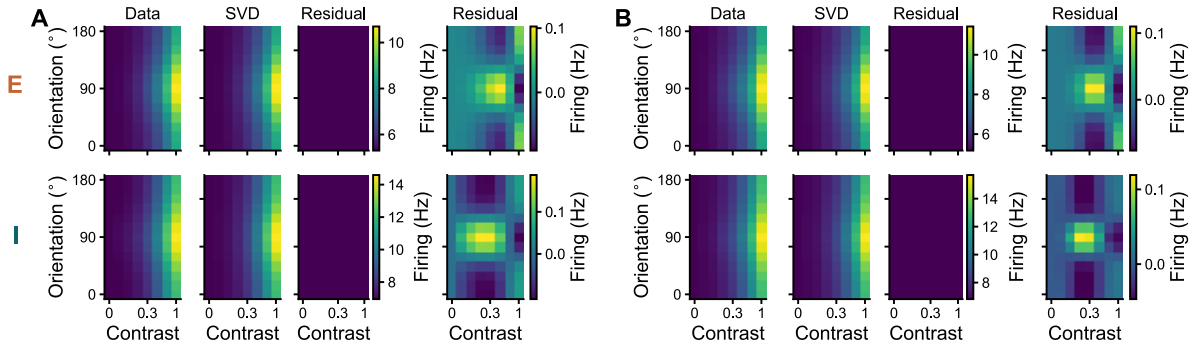

Fig. S3. Contrast-invariance of V1 population response (related to Fig. 2 –Fig. 5. *Top:* excitatory, *bottom* inhibitory. **A**, Population response for two-dimensional tuning model using hyperbolic ratio function and wrapped Gaussian. Residuals are shown once on the same scale as the data and once on a separate scale. The residuals show a significant, but very weak pattern (E: $g_z = 10.41$, residual strength $< 0.1\%$; I: $g_z = 14.31$, residual strength $= 0.1\%$). **B**, Same as (A) for two-dimensional tuning model using model-free contrast response and wrapped Gaussian, as used in Fig. 3 and Fig. 5 (E: $g_z = 11.58$, residual strength $< 0.1\%$; I: $g_z = 15.07$, residual strength $< 0.1\%$). We suspect that the small, albeit significant, spatial pattern in the residual suggests a contribution of supersaturating units, which cannot be captured by a separable model of orientation and contrast. In addition, in our population of recorded neurons, contrast sensitivity is not entirely equally distributed Fig. 2L, which could contribute to the small residual.

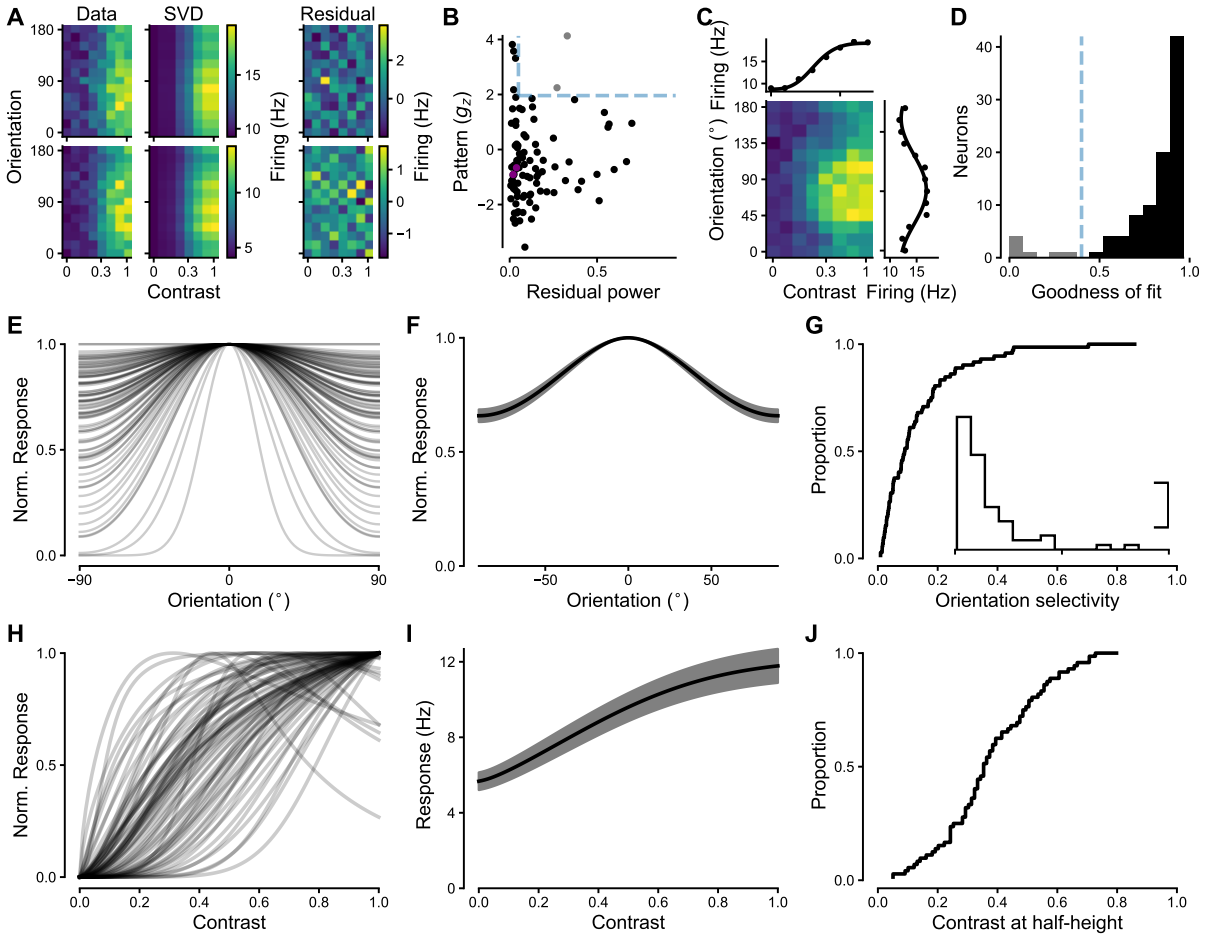

Fig. S4. Responses in dLGN (related to Fig. 2). **A**, Responses of two dLGN example neurons to combinations of orientation and contrast (*left*), first SVD component (*middle*) and SVD residual (*right*). For both examples, the absence of significant spatial patterns indicate that contrast invariance is not substantially violated (*Top*: $g_z = -0.66$, $p = 0.96$; *bottom*: g_z=-0.91, $p = 0.98$). **B**, Violations of contrast invariance were assessed by the power of the SVD residual ($> 5\%$) and significance of spatial autocorrelation ($g_z > 1.96$). *Light dots*: contrast-dependent neurons (2/98 dLGN neurons); *solid dots*: contrast-invariant neurons; *pink dots*: example neurons from (A) (96/98 dLGN neurons). **C**, Two-dimensional tuning fit consisting of a product of a hyperbolic ratio function and a wrapped Gaussian [18] for an example dLGN neuron. **D**, Distribution of fit quality across dLGN neurons. *Dashed line*: fit quality threshold (0.4). *Solid bars*: neurons considered for further analysis (89/96 dLGN neurons; mean $R^2 = 0.79$). **E**, Normalized orientation tuning component of the dLGN neurons with rising contrast response functions (72/89 neurons). **F**, Averaged normalized orientation tuning component. **G**, Cumulative distribution of orientation selectivity [OSI; 19], [20]. Mean OSI of the contrast-invariant dLGN population: 0.14 ± 0.02 (mean ± sem). *Inset:* Density histogram of orientation selectivity. Same x-axis as cumulative distribution, y-scale bar represents 2 neurons per bin of OSI. **H–J**, Same as (E–G) for normalized contrast response component and cumulative distribution of contrast sensitivity (contrast at which the contrast response function reaches half height). Mean contrast at half-height 0.39 ± 0.02, 72 dLGN neurons.

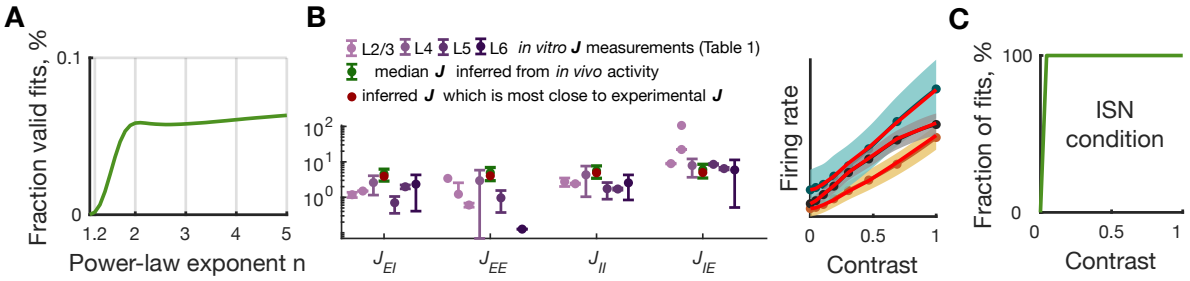

Fig. S5. Inference of connectivity parameters from contrast responses (related to Fig. 3). **A**, The fraction of valid initial connectivity weights is an increasing function of n with few valid fits found for n close to 1. **B**, *Left:* The presented inferred set of connectivity weights (red) is the closest to the weighted average of connectivity constants across layers (purple) extracted from experimental sources Table S1. *Right:* The inferred parameters set (red, left) leads to a close approximation of the average recorded contrast responses. **C**, Starting from the smallest measured contrast value of 4%, the SSN model represented an inhibition stabilized network for all inferred connectivity weights.

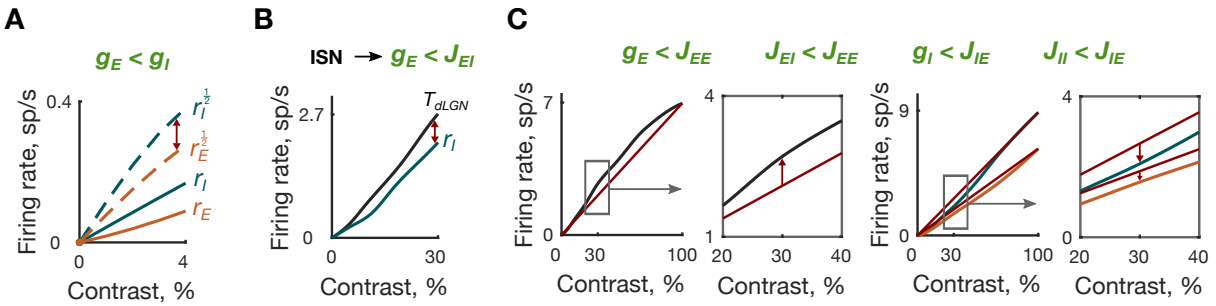

Fig. S6. Shape and magnitude of cortical and thalamic contrast responses and ISN property determine a series of relationships between the V1 connectivity weights (related to Fig. 3 and Fig. 4). **A**, For low contrasts, the firing rates r_X are smaller than $r_X^{\frac{1}{n}}$ and have a negligible impact on Eq. S10. Since $r_E^{\frac{1}{n}} < r_I^{\frac{1}{n}}$, the inequality $g_E < g_I$ must be satisfied to compensate the difference between the firing rates in Eq. S10 for small contrasts. **B**, Eq. S12 implies that the connectivity weight J_{EI} exceeds g_E in the SSN model when the SSN satisfies the ISN condition, and there exists an interval of contrasts, for which the dLGN firing rate grows faster than the I population firing rate r_I, as found in our activity recordings. **C**, There exist an interval of contrasts for which the recorded thalamic contrast response T_{dLGN} lies above (left panels), while the firing rates r_E and r_I lie below the straight lines (red lines) connecting their zero and 100% contrast values (right panels). Additionally, the deviation of the I contrast response from its linear increase exceeds that of the E response. These properties of contrast responses can explain why J_{IE} is the strongest connectivity weight targeting the I population, and J_{EE} is the strongest weight targeting the E population in the inferred connectivity sets (Eq. S15-Eq. S17 in Supplementary Information).

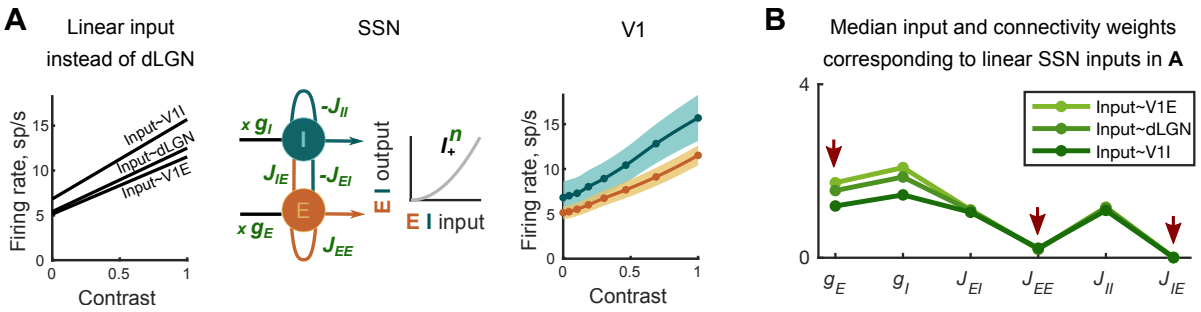

Fig. S7. Inference of connectivity parameters for linear inputs instead of recorded dLGN input (related to to Fig. 3 and Fig. 4). **A**, Linear inputs (*left*), instead of recorded dLGN input introduced in Fig. 3A, Fig. S4, were used for the inference of input and connectivity weights g_X and J_{XY} of the SSN model (*middle*). The linear inputs (*left*) have the same values for the contrast $C = 0$ and $C = 100\%$ as the recorded contrast response functions of the E, dLGN, and I populations, respectively. The connectivity parameters were computed using the same method as in Fig. 3C to generate the thalamo-cortical mapping of linear inputs to the recorded V1 contrast responses (*right*). **B**, The inferred parameters g_X and J_{XY} followed a different order than the order found for the parameters determined from the recorded dLGN contrast response. Inconsistent with direct experimental measurements, J_{EE} was smaller than g_E and J_{IE} was the smallest connectivity weight for all three linear inputs (*red arrows*).

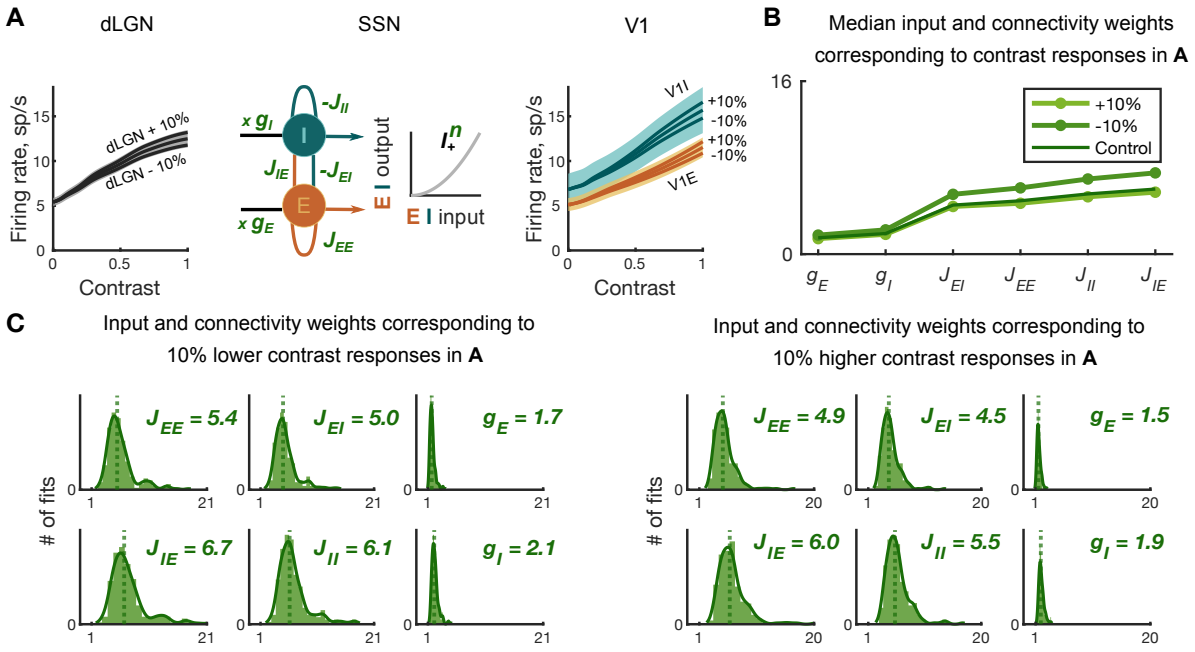

Fig. S8. Inference of connectivity parameters for modified contrast responses (related to Fig. 3). **A**, To compute connectivity and input weights for slight modifications of recorded responses, all three average contrast responses (dLGN, V1E, and V1I) were either increased or decreased by 10% accordingly to the relations (FR-FR(0))*1.1 and (FR-FR(0))*0.9, where FR denotes firing rate. **B**, In both cases most of the inferred weights followed the order found for average contrast responses. **C**, While medians of weights corresponding to decreased contrast responses remained almost the same, the medians of weights computed from increased responses were up to 13% higher.

Layer 2/3							
	Hofer et al. [36]	Ko et al. [37]	Cossell et al. [38]	Seeman et al. [39]	Karnani et al. [40]	Allen Institute for Brain Science [30]	Jiang et al. [31]
CP_{EI}						17/49	27/83
PSP_{EI}						0.31±0.20	0.42±0.03
J_{EI}						1.18±0.76	1.50±0.11
CP_{EE}	45/235	43/222	75/520	13/130		5/80	1/50
PSP_{EE}	0.2		0.45±0.68	0.34±0.32		0.22±0.24	0.34±0.08
J_{EE}	3.41		5.78±8.73	3.03±2.85		1.22±1.34	0.61±0.14
CP_{II}					13/32	36/97	28/78
PSP_{II}					0.58±0.10	0.66±0.74	0.61±0.10
J_{II}					2.59±0.45	2.69±3.02	2.41±0.39
CP_{IE}	36/41					19/50	13/83
PSP_{IE}	1.36					0.27±0.21	1.6±0.23
J_{IE}	106.28					9.13±7.10	22.30±3.21

Layer 4	
	Allen Institute for Brain Science [30]
CP_{EI}	7/34
PSP_{EI}	1.15±0.64
J_{EI}	2.60±1.45
CP_{EE}	17/236
PSP_{EE}	0.46±0.45
J_{EE}	2.95±2.88
CP_{II}	26/55
PSP_{II}	0.83±0.63
J_{II}	4.32±3.28
CP_{IE}	4/33
PSP_{IE}	0.73±0.39
J_{IE}	7.88±4.21

Layer 5		
	Allen Institute for Brain Science [30]	Jiang et al. [31]
CP_{EI}	10/72	8/32
PSP_{EI}	0.46±0.35	0.72±0.09
J_{EI}	0.7±0.53	1.98±0.25
CP_{EE}	13/553	0/12
PSP_{EE}	0.46±0.59	
J_{EE}	0.96±1.23	
CP_{II}	27/126	9/48
PSP_{II}	0.74±0.86	0.83±0.12
J_{II}	1.74±2.03	1.71±0.25
CP_{IE}	7/74	3/36
PSP_{IE}	1.01±0.41	0.87±0.07
J_{IE}	8.50±3.45	6.45±0.52

Layer 6	
	Allen Institute for Brain Science [30]
CP_{EI}	16/69
PSP_{EI}	0.92±0.76
J_{EI}	2.35±1.94
CP_{EE}	3/443
PSP_{EE}	0.21
J_{EE}	0.13
CP_{II}	41/118
PSP_{II}	0.67±0.45
J_{II}	2.56±1.72
CP_{IE}	11/76
PSP_{IE}	0.46±0.42
J_{IE}	5.93±5.41

Table S1: **Connectivity matrix for pyramidal and PV+ neurons in layers 2/3, 4, 5, and 6, summarized from recent *in vitro* studies.** The entries of the connectivity matrix J_{XY} are computed for each layer based on the experimentally measured connection probability (CP) and amplitude of the postsynaptic potential in mV (PSP) using the formula $J_{XY} = \text{CP} * \text{PSP} * N_Y/N * 100\%$, where N_Y/N is the proportion of neurons in the population N_Y. Here we use $N_E/N = 0.89$ and $N_I/N = 0.11$ [32] (see Results).