bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Pupil size dynamics predict dLGN firing mode over a wide range of timescales

Davide Crombie[1,2], Martin A. Spacek[1], Christian Leibold[1,3,*], and Laura Busse[1,3,*]

[1]Division of Neurobiology, Department Biology II, LMU Munich, Munich, Germany
[2]Graduate School of Systemic Neurosciences, LMU Munich, Munich, Germany
[3]Bernstein Centre for Computational Neuroscience, Munich, Germany
[*]Shared senior authors

Summary

Pupil size is a commonly used proxy for waking brain states such as arousal, and has been related to activity modulations in cortical sensory areas. Here, we asked whether the dorsolateral geniculate nucleus (dLGN), which provides sensory input to the visual cortex, is modulated by pupil-indexed arousal. Observing that the pupil size oscillates at multiple timescales, we developed a method to show that the spiking mode of the dLGN is predicted by pupil size oscillations over several of these timescales. Overall, we found that tonic spikes preferentially occurred during pupil dilation, while bursts occurred during contraction. These preferences could not be explained solely by pupil size *per se* or by the locomotion of the animal, and were also present during periods of stimulus viewing. We conclude that dLGN spiking activity is modulated by pupil-indexed arousal processes on various timescales, influencing the mode in which sensory signals are passed on to the cortex.

Introduction

The way in which sensory systems process information is highly dynamic, and varies in accordance with spontaneous behaviors performed by the animal (Harris & Thiele, 2011; McGinley et al., 2015b; Poulet & Crochet, 2019). Such spontaneous behaviors range from explicit motor actions, such as locomotion (Niell & Stryker, 2010), whisking (Crochet & Petersen, 2006), and orofacial movements (Stringer et al., 2019), to pupil size fluctuations that are thought to index changes in arousal (McGinley et al., 2015a; Reimer et al., 2014; Vinck et al., 2015). Indeed, although pupil size changes with ambient brightness, pupil dilation under constant illumination has been associated with hallmarks of arousal such as improved stimulus encoding (Lin et al., 2019; McGinley et al., 2015a; Reimer et al., 2014; Schwartz et al., 2020) as well as desynchronized low-amplitude, high-frequency cortical activity typical of active behavioral states (McGinley et al., 2015a; Lin et al., 2019; Reimer et al., 2014; Schwartz et al., 2020; Vinck et al., 2015). While pupil-linked arousal is often studied in cortical sensory areas, it is not yet clear what changes may be occurring at the level of the thalamic sensory nuclei that provide the cortex with sensory inputs.

Like the cortex, thalamic neurons also display prominent arousal-related activity patterns, with two state-related firing modes: tonic firing, which is associated with wakefulness; and burst firing, which is prominent during sleep or anesthesia (Steriade & Llinás, 1988; Steriade et al., 1993). Bursts are driven by a hyperpolarization-activated Ca^{2+} current, and therefore occur when thalamic neurons are hyperpolarized and spontaneous firing rates are

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

low (Sherman, 2001). Although bursts are commonly associated with sleep, they can also occur during wakefulness (Guido & Weyand, 1995; Ramcharan et al., 2000), leading to the hypothesis that thalamic nuclei use burst and tonic firing modes to gate the flow of information to, and between, cortical areas (Sherman & Guillery, 1996; Sherman, 2001). This hypothesis is supported by the fact that bursting is not evenly distributed throughout periods of wakefulness, but preferentially occurs during periods of synchronized, low-frequency cortical activity (Swadlow & Gusev, 2001; Bezdudnaya et al., 2006) and behavioral quiescence (Erisken et al., 2014; Niell & Stryker, 2010).

Hypothesizing that pupil-linked arousal might dynamically gate the relay of sensory information in the thalamus across behaviorally relevant timescales within the waking state, we developed a method to test whether pupil size can serve as a multi-scale predictor for burst and tonic spiking in the thalamic dorsolateral geniculate nucleus (dLGN). By decomposing the pupil size signal into a set of oscillatory components, we found that tonic spiking in the dLGN preferentially occurred during pupil dilation, while burst spiking was prevalent during pupil constriction, across a broad range of timescales. These influences on dLGN firing mode were not explained by pupil size *per se*, occurred also in the absence of overt locomotion, and were present even during visual stimulation. Together, our findings reveal that pupil-associated changes in arousal-related neural activity, occurring over timescales ranging from seconds to minutes, are prominent in sensory thalamocortical neurons, constantly changing the flow of sensory information available for further processing throughout the waking state.

Results

Pupil size fluctuates over a scale of seconds to minutes

In order to assess how internal state fluctuations across a wide range of timescales influenced dLGN firing mode, we paired silicon probe recordings with video-based analysis of pupil size dynamics (Figure 1). During these recordings, mice were head-fixed, and free to run on an air-cushioned styrofoam ball under largely iso-luminant visual stimulation conditions, and in the absence of a specific behavioral task (Figure 1A). To monitor spontaneous changes in internal state, we measured locomotion speed and performed close-up video recordings of one eye.

From the video recordings, we extracted a time-varying pupil size signal (PS; Figure 1B, top), which spontaneously fluctuated over the course of the recordings (Figure 1B, middle; in total $N = 60$ continuous PS segments from 19 experiments in 10 mice were collected; median duration: 158 s; Figure S1A). As expected from previous studies (Vinck et al., 2015; Reimer et al., 2014; Erisken et al., 2014; McGinley et al., 2015a), some of these fluctuations were related to overt changes in locomotion (e.g., 40 – 60 s in Figure 1B, bottom), reflected in the linear correlation between the PS and running speed (Figure S1B). In addition, we also observed fast fluctuations reminiscent of the microdilations reported by McGinley et al. (2015a), which seemed independent of locomotion and occurred also during periods of quiescence (e.g., 60 – 90 s in Figure 1B).

We sought to capture these diverse aspects of pupil size dynamics by splitting the PS into its underlying

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

oscillatory components. To this end, we employed the Hilbert-Huang transform (Huang et al., 1998), consisting of empirical mode decomposition (EMD) and Hilbert spectral analysis of the resulting set of signal components, called intrinsic mode functions (IMFs). Applying EMD to the example PS yielded a set of IMFs (Figure 1C), which spanned three orders of magnitude in frequency content (Figure 1C, left column), and varied in relative power (Figure 1C, right column). To quantify the spectral content of the PS, for each IMF, we extracted the instantaneous phase (Figure $1D_1$), frequency (Figure $1D_2$), and amplitude (Figure $1D_3$). Using these quantities, we compiled a marginal power spectrum and found a good correspondence with the traditional Fourier spectrum (median Pearson $r = 0.81$, IQR: $0.71 - 0.93$; Figure S1C). Having verified that the Hilbert-Huang transform preserved the spectral content of the PS, we assigned each IMF a relative power, defined as the mean squared amplitude expressed as a fraction of the sum of the power from all IMFs in the set, and a characteristic frequency, defined as the amplitude-weighted average instantaneous frequency (see values in Figure 1C).

We used the amplitude and frequency information of each IMF to focus on a subset of the IMFs that would be the most informative about ongoing neural activity. First, we used the relationship between frequency and power to eliminate IMFs that likely reflected noise (Figure S1D; Wu & Huang, 2004). Additionally, we focused only on IMFs that completed a minimum of four cycles within the PS segment, to ensure that each phase of the oscillation was sufficiently sampled. The remaining sets of valid IMFs contained 4.6 ± 1.2 IMFs per PS segment, most of which fell into the frequency range of 0.01 to 1 Hz (Figure 1E) and often carried a high proportion of the total PS signal power (Figure 1F, Figure S1E). IMFs in these sets had weak linear (Figure S1F), higher-order (Figure S1G), and phase (Figure S1H) dependencies, particularly when similar in frequency, but otherwise did not form strongly related groups. Therefore, for each PS segment, we obtained a set of largely independent components representing pupil size oscillations occurring over timescales of seconds to minutes.

Spontaneous spiking in the dLGN is tuned to pupil size fluctuations across multiple timescales

Having captured spontaneous pupil size fluctuations with a set of largely independent oscillatory components that spanned a wide range of time scales, we next asked if these pupil size oscillations were accompanied by modulations of arousal-related neural activity in the visual thalamus. To this end, we categorized spikes from each dLGN unit as tonic or burst (Figure 2A; Lu et al., 1992), and performed downstream analyses separately on tonic spikes and burst events, for which all spikes in a burst (mean $\pm$ s.d.: 2.51 ± 0.49 spikes/burst) were treated as a single event. Among our recorded dLGN units, the vast majority (86%, 174/203) displayed bursting events, which occurred at a median rate of 0.12 bursts/s (Figure $S2A_1$) and accounted for $10 \pm 13\%$ of dLGN spikes (Figure $S2A_2$).

To determine if the prevalence of bursting and tonic spiking in the dLGN can be predicted by PS oscillations, we performed a phase tuning analysis relating each PS-IMF to dLGN single-unit spiking activity. For each IMF-unit pair, we collected the instantaneous phase of the IMF at all timepoints where burst events or tonic spikes occurred.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

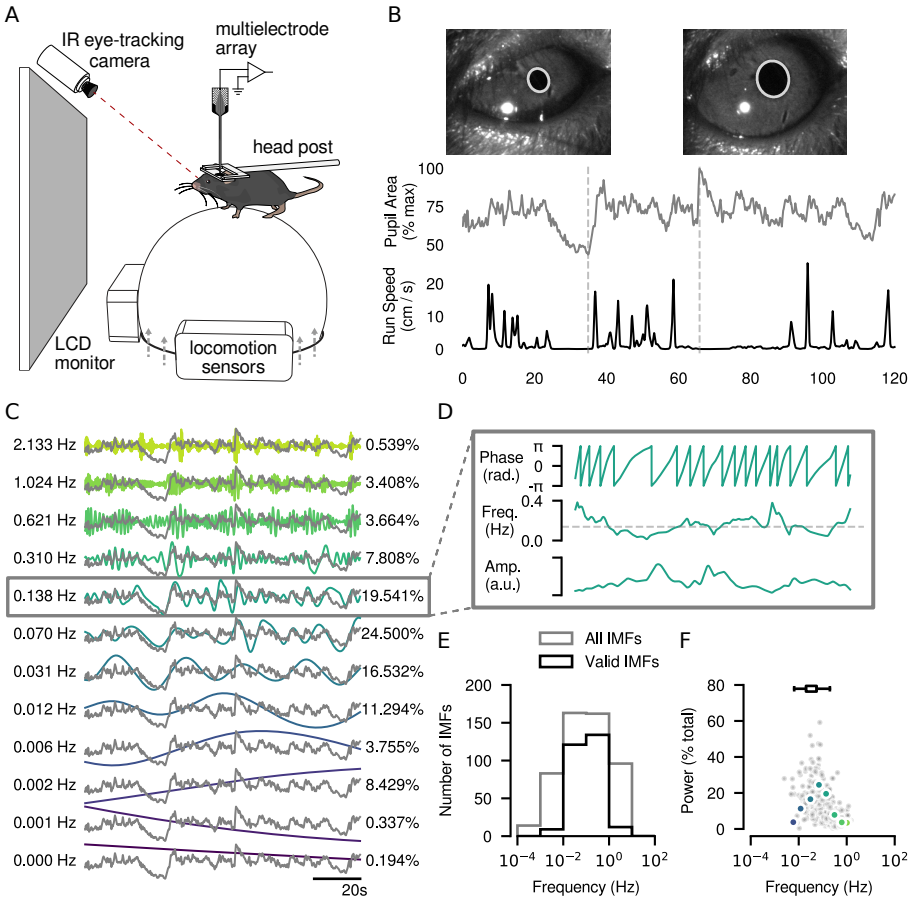


Figure 1 The pupil size signal is represented by a set of component oscillations spanning a wide range of frequencies. **(A)** Schematic of the experimental setup. **(B)** Video-based pupil size tracking and simultaneously recorded running speed for two minutes of an example recording segment. *Top*: Two representative video frames with fitted pupil outline (*light gray*). *Middle*: Pupil area expressed as percentage of the maximum size in the given segment of recording. *Bottom*: Run speed calculated as the Euclidean norm of three perpendicular components of ball velocity (Dombeck et al., 2007). **(C)** Intrinsic mode functions (IMFs) resulting from empirical mode decomposition (EMD) of the pupil size signal (PS) segment from (B), shown along with their characteristic frequencies (*left*) and power (% total, *right*). **(D)** Instantaneous phase (*top*), frequency (*middle*), and amplitude (*bottom*) of the highlighted IMF in (C). The amplitude at each time point was used to weight the instantaneous frequency values to produce a characteristic frequency (*dashed grey line*). **(E)** Distribution of IMF characteristic frequencies from 60 PS segments before (*grey*) and after (*black*) removal of invalid IMFs (see also Figure S1D). **(F)** Scatter plot of characteristic frequency versus power for the valid IMFs from (E). Valid IMFs from (C) are shown with colored dots. Inset whisker plot shows the quartiles of the characteristic frequency distribution when considering only the IMF that carried the highest power for each set.

We used these phases to compile phase histograms, and assess phase tuning. We observed three characteristics of phase tuning, which are illustrated in the example IMF-unit pair shown in Figure 2B: (1) both bursts and tonic spikes are tuned to the phase of the PS IMF component, (2) phase tuning is stronger for bursts than for tonic spikes, and (3) bursts and tonic spikes prefer opposing phases of the IMF oscillation, with bursts occurring predominantly during pupil constriction and tonic spikes during pupil dilation.

Phase tuning of dLGN spiking activity was not only observed in the example unit (Figure 2B) but was also found across the whole set of IMF-unit pairs (Figure 2C). We assessed strength and significance of phase tuning with unbiased methods, accounting for the potential asymmetry of IMF wave shapes (Figure S2B) and short-term patterns of both bursting and tonic spiking (Figure S2C). Specifically, to account for potential inflation of tuning for IMFs with high phase bias (e.g. low frequency IMFs, Figure S2B$_2$), we first transformed IMF phases into ranks,

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

yielding a uniform distribution of phase ranks (Siapas et al., 2005). We then used these phase ranks to compute a bias-corrected, squared resultant vector length (Kutil, 2012) as a measure of tuning strength ($\hat{R}^2$). Additionally, to test the significance of tuning irrespective of spurious influences of short-term ($\sim$ 300 ms) serial dependence in bursting and tonic spiking (Figure S2C), we compared actual tuning strength to a shuffle-distribution of tuning strengths, obtained by randomly permuting 300 ms bins of spiking activity. After accounting for these various sources of biases, we observed significant phase tuning ($p < 0.05$) for tonic spikes in 50% (1360/2727; Figure 2C, *solid blue dots*), and for burst events in 31% (533/1705; Figure 2C, *solid red dots*) of IMF-unit pairs (Figure S2D$_1$). Importantly, the majority of units had significant phase tuning for at least one IMF (tonic spikes: 84%, 529/632; bursts: 68%, 272/399) and many units were in fact significantly tuned to multiple IMFs (Figure S2D$_2$). Significant tuning was not limited to a particular frequency band, but was observed across a broad range of frequencies (Figure S2D$_3$). Together, our results indicate that bursting and tonic spiking activity in the dLGN are modulated over a broad range of timescales, as indexed by oscillatory components of the PS.

Having observed a striking modulation of dLGN spiking activity by the dynamics of the PS, we next quantified whether the strength of modulation differed between bursts and tonic spikes and IMF frequency bands. Summarizing tuning strength as a function of IMF frequency irrespective of phase (Figure 2D), we found that, overall, bursts were more strongly tuned to PS-IMFs than tonic spikes (ANOVA main effect of spike type: $F_1 = 324.21, p = 5.39 \times 10^{-67}$; Figure 2D). In addition, we found a significant effect of frequency (ANOVA main effect of frequency bin: $F_6 = 2.92, p = 7.74 \times 10^{-3}$; spike type-frequency bin interaction: $F_{1,6} = 2.36, p = 0.03$; Figure 2D). Given the dependence of frequency on mean tuning strengths, we next determined which IMF had the strongest tuning for each unit, and asked whether this "best frequency" was consistent across units. Taking the IMF frequency with the strongest tuning for each unit, we find that, although units tend to have their strongest tuning to frequencies around 0.1 Hz, there is a broad range of best frequencies (Figure S2E, *solid bars*). This becomes particularly obvious for longer recordings where a larger range of PS oscillation frequencies could be assessed (Figure S2E, *empty bars*), indicating that conclusions regarding the strongest timescale of modulation might be limited by recording duration of the data included here. Finally, we turned to the striking observation that, despite the variability in tuning phase across frequencies, bursts and tonic spikes consistently preferred opposite phases of the IMF (Figure 2B, C). Although it is evident from the broad segregation of color in Figure 2C that an anti-phase pattern is present in the whole population of IMF-unit pairs, we asked if this preference for opposing phases also holds for bursts and tonic spikes occurring within the same unit. Therefore, for each unit where significant tuning of both spike types was observed, we considered only the PS-IMF to which bursts were most strongly tuned, and compared the tuning phase for bursts and tonic spikes. Although across units the tuning phase varied, a clear separation between the preferred phase of bursts and tonic spikes was still present (Figure 2E). In fact, computing the difference between preferred phases for each unit, we found that overall the preferred phase of each spike type was separated by exactly $\sim$ 180 degrees (Figure 2F; V-test with mean π: $V = 102.22, p = 0$), indicating that bursts and tonic spikes preferentially

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

occur at opposing phases of pupil oscillation. We conclude that dLGN units alternate between a tendency to fire in bursts during pupil contraction and tonic spikes during pupil dilation over multiple timescales.

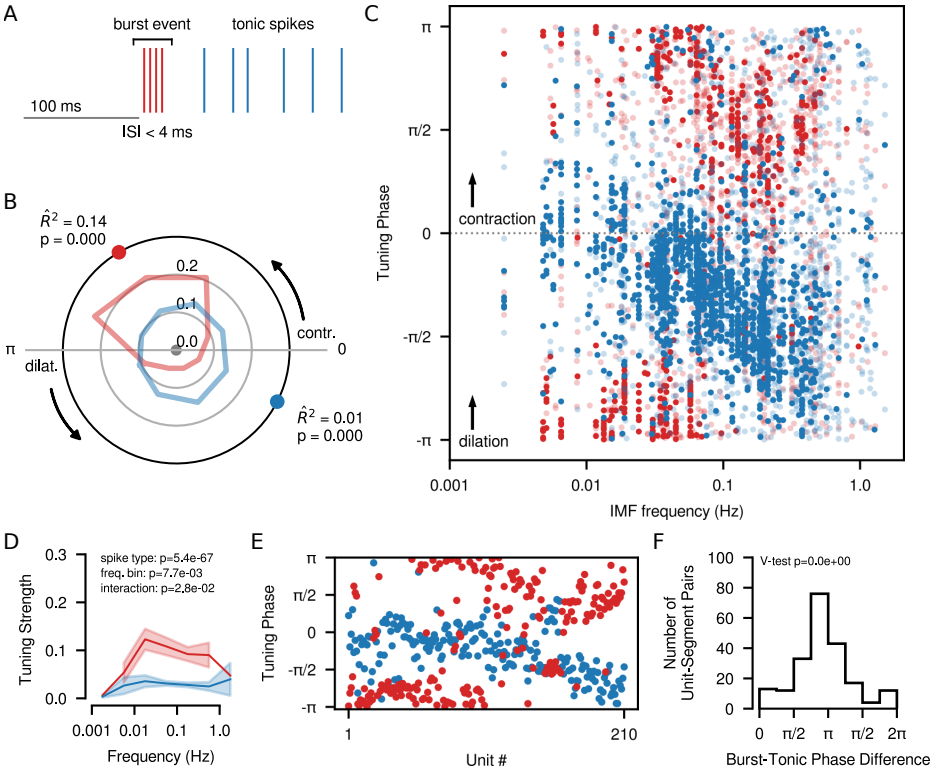


Figure 2 Tonic spikes and bursts in the dLGN are tuned to opposite phases of pupil size oscillations. (A) Bursts (*red*) were defined as $\geq$ 2 spikes occurring after >100 ms of silence and with an inter-spike interval less than 4 ms (Lu et al., 1992), all remaining spikes were classified as tonic spikes (*blue*). All spikes in a burst were treated as belonging to a single burst event. **(B)** Phase distributions for bursts and tonic spikes from an example dLGN unit with respect to the highlighted IMF from Figure 1C and D. Concentric circles indicate the proportion of burst events or tonic spikes occurring in one of 8 phase bins ($\hat{R}^2$: tuning strength; *colored dots*: tuning phase). **(C)** Relationship between tuning phase and characteristic IMF frequency of bursts and tonic spikes for all IMF-unit pairs for which phase tuning was assessed (tonic N = 2727; burst N = 1705). Significant tuning ($p \leq 0.05$; tonic: 1360 / 2727; burst: 533 / 1705) is indicated with *bold dots*. **(D)** Tuning strengths across frequency bins for significantly tuned IMF-unit pairs, *shaded areas* represent 95% confidence intervals estimated as 1.96×SEM (*Inset text*: ANOVA results). **(E)** Preferred phases of bursts and tonic spikes to the IMF with the strongest burst tuning for each unit, sorted by the IMF frequency. Only IMF-unit pairs where both burst and tonic spike tuning were significant were included (N = 210). **(F)** Distribution of differences between preferred phase of burst and tonic spikes for the IMF-unit pairs shown in (E) (*Inset text*: V-test for non-uniformity and mean of π).

The relationship between dLGN spiking activity and pupil dynamics is not driven by pupil size *per se*

So far, we have established that the spiking activity of dLGN units is modulated according to the phase of pupil size oscillations. However, pupil size *per se* is more commonly used as an indicator of arousal state, irrespective of its oscillatory dynamics (*cf.* Reimer et al., 2014; McGinley et al., 2015a). Indeed, it has been suggested that variables capturing pupil size dynamics, such as phase or the temporal derivative of pupil size, may only be explanatory due to their correlation with pupil size (McGinley et al., 2015a).

If this were the case, one would expect that IMFs with high relative power, which contribute more to determining absolute pupil size, would have higher tuning strengths. To test this hypothesis, we examined tuning strength as a function of IMF power Figure 3A, and found that, IMF power was significantly predictive of tuning strength for

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

both bursts ($F_1 = 89.6$, $p = 9.3 \times 10^{-21}$; Figure 3A$_1$) and tonic spikes ($F_1 = 73.6$, $p = 1.6 \times 10^{-17}$; Figure 3A$_2$). IMF power, however, explained only a small fraction of the variability in tuning strengths (bursts: $r^2 = 0.05$; tonic spikes: $r^2 = 0.03$), and its overall influence was small (bursts: slope $= 1.7 \times 10^{-3}$; tonic spikes: slope $= 6.8 \times 10^{-4}$). Such a small, positive relationship might be at least partially explained by the fact that, in general, phase estimation for low power components will be less reliable, and therefore phase tuning measured for these components will be lower. Therefore, the weak contribution of IMF power to determining tuning strengths suggests that pupil size *per se* is not a decisive marker of dLGN firing mode modulation.

Controlling for pupil size is particularly important in the visual system, where the amount of light entering the eye has the potential to constitute a confounding factor when investigating pupil-linked neural activity modulations (discussed in Erisken et al., 2014; Liang et al., 2020; McGinley, 2020; Reimer et al., 2014; Schröder et al., 2020). Hence, to rule out systematic differences in pupil size across IMF phase bins (Figure S3A$_1$ and A$_2$) as the main driver IMF phase tuning, we applied a histogram-matching procedure (Churchland et al., 2010), in which we sub-sampled time periods such that the distribution of pupil sizes were equalized across each of four phase bins by taking the greatest common distribution. Reassuringly, even after equalizing the pupil size distribution across phase bins, we still observed significant IMF phase tuning across the full range of frequencies assessed (Figure S3A$_3$). Furthermore, we found that the tuning strength and phase preference patterns, as described in Figure 2, were preserved: overall, bursts were more strongly tuned than tonic spikes (ANOVA: $F_1 = 163.95$, $p = 8.16 \times 10^{-35}$), with a marginal effect of frequency (ANOVA: $F_5 = 2.81$, $p = 0.02$; interaction: $F_{1,5} = 3.65$, $p = 5.82 \times 10^{-3}$). Finally, despite differences in the mean tuning phases of tonic spikes measured after matching for pupil size compared to when all time periods were considered (Figure 3A$_4$; Watson-Williams test: bursts $F = 1.03$, $p = 0.31$; tonic spikes $F = 7.10$, $p = 0.01$), burst spikes were still tuned to IMF phases corresponding to contraction, while tonic spikes were tuned to dilating phases. This result supports the notion that the main explanatory factor linking pupil size oscillations and dLGN tuning is the phase of the underlying oscillation, rather than the absolute value of pupil size, or the associated changes in retinal irradiation.

Modulation of dLGN spiking activity occurs within locomotion-defined states

It has been previously described that large changes in pupil size often occur around surrounding bouts of locomotion (Vinck et al., 2015). This is also evident in our data, where, in our example segment in Figure 1B, the pupil constricted after a bout of locomotion and dilated preceding a second bout (20 – 40 s); this contraction-dilation cycle was captured well by the highlighted IMF in Figure 1C. Having observed that a number of IMFs were significantly correlated with running speed (37%, 100/270 IMFs; Figure S3B$_1$, *black dots*), we next examined whether this correlation was predictive of phase tuning strength. We found that correlation with running speed was somewhat predictive of IMF phase tuning strength for bursts ($F_2 = 7.85$, $p = 4.06 \times 10^{-4}$; Figure 3B$_1$), and tonic spikes ($F_2 = 11.61$, $p = 9.51 \times 10^{-6}$; Figure 3B$_2$), though the amount of variance explained was minimal in both

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

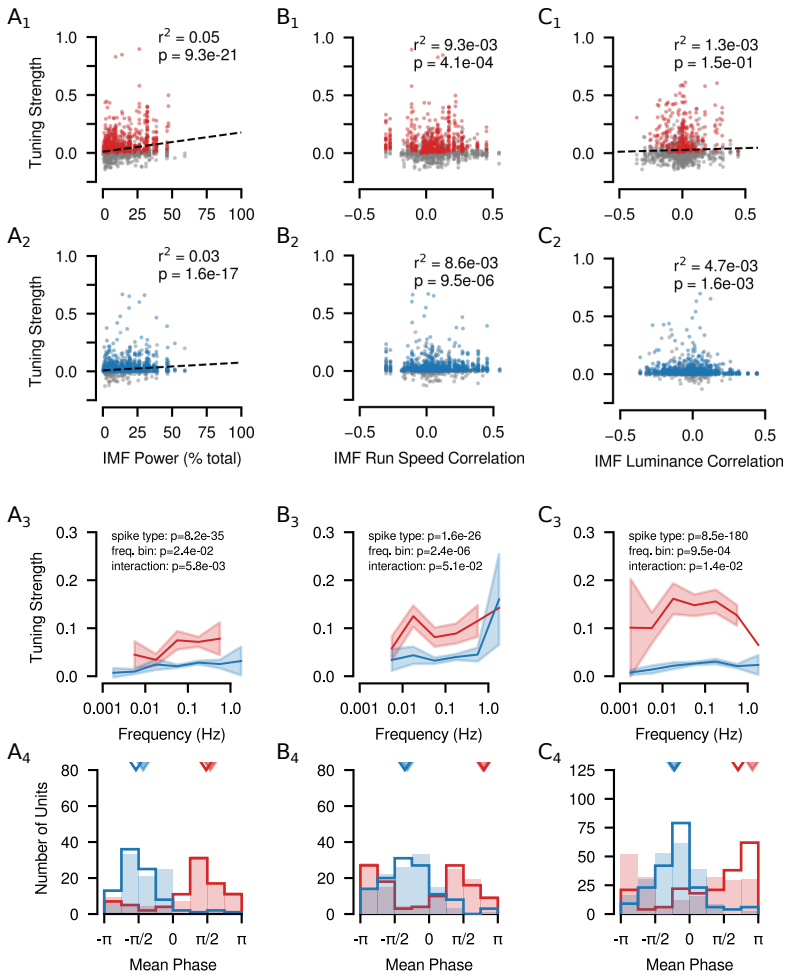


Figure 3 Pupil size oscillations can modulate dLGN activity largely independently of pupil size *per se*, locomotion, and visual stimulation. **(A_1)** Burst tuning strength regressed on IMF power for all IMF-unit pairs. *Colored dots*: significantly tuned IMF-unit pairs. *Dashed line*: regression line (slope = 1.7×10^{-3}), *inset text*: coefficient of determination and p-value of regression. **(A_2)** Same as A_1 for tonic spikes (slope = 6.8×10^{-4}). **(A_3)** Mean tuning strength across frequency bins computed after matching for pupil size (*red*: burst; *blue*: tonic spikes; *shaded areas*: 95% CI, estimated as 1.96×SEM). **(A_4)** Distribution of tuning phases for bursts (*red*) and tonic spikes (*blue*) for the IMF to which each unit has the strongest burst tuning. *Empty bars*: tuning strengths and tuning phases computed after matching for pupil size; *filled bars*: tuning phase distribution for the same IMFs, but computed using all spikes. *Arrow heads*: circular means for each distribution (Watson-Williams test: bursts $F = 4.63$, $p = 0.03$, tonic spikes $F = 11.34$, $p = 0.00$). **(B_1)** Burst tuning strength regressed on IMF correlation with run speed (squared correlation was included as an additional co-variate, model comparison: F = 11.37, p = 7.7×10^{-4}; correlation slope = -0.07; squared correlation slope = 0.20). **(B_2)** Same as B_1 for tonic spikes (squared correlation was included as an additional co-variate, model comparison: F = 23.21, p = 1.5×10^{-6}; correlation slope = -0.02; squared correlation slope = 0.14). **(B_3)** Same as A_3 but for spikes that occurred during quiescence (time periods outside of running bouts, with the exclusion of 5 s immediately preceding and following a running bout). **(B_4)** Same as A_4 but for spikes that occurred during quiescence (Watson-Williams test: bursts $F = 0.05$, $p = 0.82$ and tonic spikes $F = 0.53$, $p = 0.47$). **(C_1)** Same as B_1 but for IMF correlation with stimulus luminance (slope = 0.03). **(C_2)** Same as C_1 but for tonic spikes (squared correlation was included as an additional co-variate, model comparison: F = 3.89, p = 0.05; correlation slope = -0.02; squared correlation slope = 0.07). **(C_3)** Same as A_3 but for spikes that occurred during stimulus viewing. **(C_4)** Same as A_4 but for spikes that occurred during stimulus viewing. Tuning phases given by the *faded bars* are those computed from the data in Figure 2E when no stimulus was present (Watson-Williams test: bursts $F = 17.7$, $p = 0.00$ and tonic spikes $F = 0.56$, $p = 0.45$).

cases (bursts: $r^2 = 9.30 \times 10^{-3}$; tonic spikes: $r^2 = 8.60 \times 10^{-3}$). Furthermore, IMFs with a significant run speed correlation were no more likely to drive significant phase tuning than uncorrelated IMFs (bursts: correlated 28%, 126/456; uncorrelated 33%, 397/1219, $\chi^2 = 3.54$, $p = 0.06$; tonic spikes: correlated 52%, 380/736, uncorrelated: 49%, 950/1943, $\chi^2 = 1.49$, $p = 0.22$). In sum, although we found the animal's locomotion to be reflected in some of the IMFs, this relationship was not decisive in determining which IMFs drove strong modulations in dLGN units.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Thus, pupil size dynamics can provide a marker for dLGN modulation that is independent of overt locomotion.

We further corroborated that pupil size dynamics and overt locomotion are independent by separating our data into periods of locomotion and quiescence, and asked whether the signatures of dLGN IMF tuning were preserved. In order to re-assess IMF phase tuning considering only spikes that occurred during one of the two states, we divided the recordings into periods of locomotion and quiescence while excluding periods of transition between the two states (Figure S3B$_2$). Considering only periods of quiescence, we still observed significant tuning over a broad range of frequencies (Figure S3B$_3$), which showed all the characteristics of our original results: stronger tuning for bursts than tonic spikes (ANOVA main effect of spike type: $F_1 = 119.48$, $p = 1.57 \times 10{-}26$; main effect of frequency bin: $F_5 = 6.89$, $p = 2.40 \times 10^{-6}$; interaction: $F_{1,5} = 2.21$, $p = 0.05$; Figure 3B$_3$), and similar mean tuning phases for bursts and tonic spikes to when all time periods were considered (Watson-Williams test: bursts $F = 0.05$, $p = 0.82$ and tonic spikes $F = 0.53$, $p = 0.47$); Figure 3B$_4$). Although quiescence was the dominant behavioral state during our recordings, and bursts are more scarce during locomotion (Erisken et al., 2014; Niell & Stryker, 2010), the IMF phase tuning was also still present when considering only spikes that occur while the animal is running (Figure S3B$_4$, B$_5$, and B$_6$). In summary, the same general influence of pupil size oscillations on dLGN firing mode was thus observed during quiescence and locomotion, and after eliminating transitions between the two states. Together, these results exclude the possibility that pupil-indexed dLGN modulations were driven by locomotion-related changes in pupil size and arousal.

Spontaneous pupil-indexed dLGN modulations persist during stimulus viewing

Above, we reported stable tuning of dLGN activity to the pupil size signal, which was driven by internal processes and occurred in the absence of a particular visual stimulus. However, it is unclear whether this coupling of internally driven pupil size dynamics and dLGN firing mode would persist during visual stimulation. Two factors might contribute to a de-coupling. Firstly, stimulus-driven dLGN activity, such as bursts triggered by sequences of a prolonged inhibitory stimulus, followed by an excitatory stimulus (Lesica & Stanley, 2004; Lesica et al., 2006), may dominate over spontaneous activity fluctuations. Secondly, the pupil signal itself during stimulus viewing might reflect not only internal processes, but might also be driven by luminance changes (Reimer et al., 2014) or other salient stimulus features. Thus, to test whether visual stimulation might disrupt the coupling between pupil size and dLGN spiking, we replicated our analyses relating PS dynamics to dLGN spiking for experiments in which we repeatedly presented 5 s naturalistic movie clips ($N = 184$ PS segments from 30 experiments in 6 mice).

We first assessed to which degree the PS dynamics reflected features of the visual stimulus. As expected, for some recordings with presentation of a naturalistic movie stimulus, pupil size systematically varied with the stimulus, as evident from the vertical structure in the example recording shown in Figure S3C$_1$. We also noted, however, that large pupil size changes occurred between trials. To quantify the relationship between stimulus-driven and

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

stimulus-independent changes in pupil size, we compared the pupil size variance over the trial mean to the mean variance over corresponding time points across trials (signal-to-noise ratio, SNR, as defined in Baden et al., 2016), where a value of 1 would indicate that all trials were identical and hence the PS was entirely driven by changes in the visual stimulus. In our example, SNR amounted to 0.28 Figure S3C$_1$, which was representative of the recorded data, where we found that at least twice as much variance occurred between trials for all recordings (mean $= 0.15 \pm 0.01$, $N = 30$; Figure S3C$_2$). This low SNR indicates that internally-driven pupil size fluctuations were dominant during stimulus viewing.

While our previous analyses showed that pupil size *per se* was dominated by internally driven processes, we next asked to which degree individual PS-IMF components were correlated to changes in overall stimulus brightness. Using a lag between the stimulus brightness and IMFs of 380 ms, determined by taking the mean lag of the peaks in the absolute cross-correlations of the PS with stimulus brightness (Figure S3C$_3$), we found that 26% (234/889) of the IMFs, largely restricted to frequencies > 0.1 Hz, were significantly correlated with stimulus brightness (Figure S3C$_4$, *black dots*). The strength of this relationship to stimulus brightness, importantly, was not predictive of tuning strength for bursts ($F_1 = 2.06$, $p = 0.15$, $r^2 = 1.3 \times 10^{-3}$; Figure 3C$_1$), and tuning strength for tonic spikes was only weakly influenced ($F_1 = 6.43$, $p = 1.64 \times 10^{-3}$, $r^2 = 4.7 \times 10^{-3}$; Figure 3C$_2$). Consistent with previous studies showing that luminance sequences in naturalistic movies can trigger changes in dLGN firing mode (Lesica & Stanley, 2004; Lesica et al., 2006), IMFs with a significant correlation to stimulus brightness were slightly more likely to drive significant burst tuning (correlated: 23% , 82/349; uncorrelated: 18%, 225/1236; $\chi^2 = 4.55$, $p = 0.03$) and significant tonic tuning (correlated: 62%, 304/494; uncorrelated: 44%, 997/2247; $\chi^2 = 47.18$, $p = 6.47 \times 10^{-12}$). Together, these results suggest that during naturalistic stimulus viewing, brightness-driven modulations of pupil size and dLGN spiking co-exist with the internally-driven pupil size modulations we have measured thus far.

Having shown that even during naturalistic stimulus viewing, coupling of internally driven pupil size dynamics and dLGN firing mode persisted, we tested whether the main characteristics of dLGN IMF phase tuning observed under spontaneous conditions were also present during stimulus viewing. Applying our analyses of IMF phase tuning revealed that the general signatures of dLGN modulation were still present even when the dLGN is engaged in naturalistic stimulus processing: (1) We observed significant tuning over a broad range of frequencies for both tonic spikes and bursts (Figure S3C$_5$); (2) burst tuning was overall stronger than tonic spike tuning (Figure 3C$_3$; ANOVA: $F_1 = 977.49$, $p = 8.50 \times 10^{-180}$), and tuning strengths were modulated by IMF frequency (ANOVA: $F_6 = 3.78$, $p = 9.47 \times 10^{-4}$; interaction: $F_{1,6} = 2.66$, $p = 1.43 \times 10^{-2}$); (3) bursts and tonic spikes also remained tuned to contracting and dilating phases of the IMFs, respectively (Figure 3C$_4$), although the mean preferred phase of burst tuning differed slightly from the spontaneous condition (Watson-Williams test: bursts $F = 17.70$, $p = 0.00$ and tonic spikes $F = 0.56$, $p = 0.45$). Given the overall similarity of our findings between spontaneous visual activity and stimulus viewing, our results suggest that dLGN stimulus responses occur on top of arousal-induced modulations in spiking mode, which could have profound consequences for the type of information that is passed

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

on to the cortex.

Discussion

Our results establish that spiking mode in dLGN, the primary visual thalamic nucleus, is predicted by pupil size dynamics occurring over multiple timescales. We extracted components of the pupil size signal to reveal that dLGN spiking activity was tuned to the phase of pupil size oscillations, at timescales ranging from several seconds to several minutes. Throughout these timescales, bursts of spikes occurred preferentially during phases corresponding to pupil contraction and tonic spikes occurred predominantly during pupil dilation. Tuning to specific phases of the pupil size oscillations could not be explained by differences in pupil size *per se* or by transitions between quiescence and locomotion, and occurred not only during periods without visual stimulation, but also during viewing of naturalistic movies. These results contribute to a growing literature indicating that arousal-related neural activity modulations in the waking state occur at early stages in the sensory processing stream.

While raw pupil size has been successfully employed as a marker for internal arousal processes in sensory circuits (Augustinaite & Kuhn, 2020; Liang et al., 2020; Lin et al., 2019; McGinley et al., 2015a; Molnár et al., 2021; Petty et al., 2021; Schröder et al., 2020; Schwartz et al., 2020), other studies, including ours, have instead considered the pupil size signal as a continuous oscillatory process and focused on its dynamics (Reimer et al., 2014, 2016). We further extended this perspective by exploiting a decomposition method that respects the non-stationary, multi-scale nature of the pupil size signal (Huang et al., 1998), and found that several timescales were related to dLGN firing modes. The presence of dLGN activity modulation across several timescales could result from distinct processes with different intrinsic timescales (e.g. Reimer et al., 2016), which might be expected given the diversity of documented cognitive and physiological influences on pupil size (recently reviewed in Joshi & Gold, 2020; Larsen & Waters, 2018; Mathôt, 2018). Although we do not claim that the components of the pupil size signal recovered by EMD represent the time-courses of independent internal processes, moving beyond pupil size *per se*, and instead considering the phase and temporal scale of pupil size oscillations, promises to be fruitful beyond the analysis of sensory processing, for instance, in tracking the various timescales of action of neuromodulators (Devilbiss & Waterhouse, 2011; Parikh et al., 2007; Rajkowski et al., 1994; Reimer et al., 2016; Ruivo et al., 2017), or the effects of baseline and evoked arousal on task performance (Aston-Jones & Cohen, 2005; de Gee et al., 2014, 2020; van Kempen et al., 2019; Nassar et al., 2012; Schriver et al., 2018; McGinley et al., 2015a).

Pupil size fluctuations are underpinned by the activity of a diversity of loci in the brain (Joshi & Gold, 2020; Larsen & Waters, 2018), among which the activity of neuromodulatory nuclei are a likely source for the pupil-linked changes in dLGN firing mode we observed here. Two prominent sources of neuromodulation in the dLGN are norepinepherine and acetylcholine (NE and ACh; McCormick, 1992; Sherman & Guillery, 1996). NE is provided to the thalamus by the locus coeruleus (LC; Asanuma, 1992; Lindvall et al., 1974; Mackay-Sim et al., 1983), while ACh

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

in the thalamus comes from several brainstem nuclei (PPN/LDT and PBN; Hallanger et al., 1987; Sokhadze et al., 2018) and the basal forebrain (BF; Hallanger et al., 1987; Sokhadze et al., 2019; Woolf & Butcher, 1986). Activity of the LC-NE system has famously been linked with pupil dilation (Aston-Jones & Cohen, 2005; Joshi et al., 2016; Reimer et al., 2016), and, consistent with a role for the LC-NE system in the present findings, stimulation of the LC *in vivo* suppresses bursting in the dLGN (Holdefer & Jacobs, 1994) and primary somatosensory thalamus (VPM; Rodenkirch et al., 2019). Although activity of the cholinergic PPN/LDT and PBN has not been directly linked to pupil size fluctuations, the effects of ACh on thalamocortical neurons (TCs) are similar to those we observed during pupil dilation: stimulation of the PPT/LDT increases firing rates in the VPM (Castro-Alamancos & Oldford, 2002), and switches dLGN neurons from burst to tonic mode (Lu et al., 1993). These effects may be mediated by direct effects of ACh and NE on TCs, where *in vitro* studies of thalamic slices revealed that ACh causes a slow depolarization that could promote tonic firing (McCormick, 1992), and that NE modulates thalamic burst firing by enhancing a hyperpolarization-activated cation current (Pape & McCormick, 1989). Moving forward, *in vivo* work will be needed to identify temporal patterns of neuromodulator release in the dLGN in relation to behavioral variables such as pupil size.

In addition to these direct influences, neuromodulation of dLGN thalamocortical cells may also occur indirectly via other elements of the thalamo-cortical circuit. First, NE and ACh may affect TCs via the thalamic reticular nucleus (TRN; Asanuma, 1992; Bickford et al., 1994; Hallanger et al., 1987; Sokhadze et al., 2019; Woolf & Butcher, 1986), which provides inhibition to dLGN. In particular, TRN neurons display an extended hyperpolarization in response to cholinergic input (McCormick, 1992; Sun et al., 2013), which in the TRN includes the BF, whose cortical projections have been linked to pupil dilation (Nelson & Mooney, 2016; Reimer et al., 2016). Second, the activity level of excitatory layer 6 corticothalamic feedback neurons (CTs) seems to increase with both to the presence of ACh (Sundberg et al., 2018) and pupil size (Augustinaite & Kuhn, 2020), and could thus contribute to promoting the dilation-associated tonic firing we observed here. Furthermore, at lower activity levels, CT feedback can have a suppressive effect on TCs (Crandall et al., 2015), meaning that a pupil-linked decrease in CT feedback (Augustinaite & Kuhn, 2020) could promote bursting. On the other hand, related work from our lab has shown that during locomotion the reduction in thalamic bursting and overall increase in firing rates still occurs even when CT feedback is optogenetically suppressed (Spacek et al., 2020). Together, these findings indicate that arousal-related modulations of thalamic activity involve both cortex-dependent and independent mechanisms (Molnár et al., 2021). Third, pupil-linked modulations have also been observed in retinal input to the dLGN (Liang et al., 2020), which might be mediated by pre-synaptic modulation of the retinogeniculate synapse by serotonin (5-HT; Chen & Regehr, 2003). While thalamic relay cells are themselves generally depolarized by 5-HT (Varela & Sherman, 2009), which is consistent with our findings given that activation of the dorsal raphe is linked with pupil dilation (Cazettes et al., 2021), the dominant pre-synaptic effect of 5-HT is suppression (Chen & Regehr, 2003), consistent with the dilation-related suppression reported by Liang et al. (2020). Therefore, how these modulations of retinal input relate to our

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

findings remains unclear, and might differ for specific stimuli, RGC inputs, and locations within the dLGN.

The modulations in burst and tonic spiking we observed indicate that, throughout periods of wakefulness, and even during stimulus viewing, the dLGN is in constant alternation between two distinct information processing modes. One prominent view is that tonic spikes encode information about a stimulus linearly, while stimulus-driven bursts provide an all-or-none "wake-up call" to the cortex (Sherman, 2001), by providing augmented retinogeniculate (Alitto et al., 2019) and geniculocortical (Swadlow & Gusev, 2001) communication, and improved stimulus detectability (Guido & Weyand, 1995). Burst spikes, in comparison with tonic spikes, also have different feature selectivity, tracking low frequency stimulus content (Bezdudnaya et al., 2006; Mukherjee & Kaplan, 1995; Reinagel et al., 1999) with an earlier response phase (Alitto et al., 2005; Grubb & Thompson, 2005; Spacek et al., 2020), integrating input over longer time-scales with biphasic response kernels (Alitto et al., 2005; Lesica & Stanley, 2004; Lesica et al., 2006; Mease et al., 2017; Zeldenrust et al., 2018), and preferring smaller stimuli (Rivadulla et al., 2003) with a stronger suppressive surround (Alitto et al., 2005). While bursts are sometimes thought to be less informative than tonic spikes (Butts et al., 2010; Reinagel et al., 1999; Rodenkirch et al., 2019), it has also been argued that their information coding role cannot be replaced by tonic spikes (Gaudry & Reinagel, 2008). Indeed, subtle intra-burst mechanisms seem to allow bursts to encode more information than previously thought (Mease et al., 2017). Importantly, stimulus-driven bursts (Lesica & Stanley, 2004; Lesica et al., 2006) are not only dependent on features of the stimulus, but rather the same naturalistic stimulus sequence can elicit bursts or tonic spikes depending on the membrane potential of the neuron (Lesica et al., 2006). Therefore a spontaneous alternation between a hyperpolarized state with bursting, and one with tonic spiking, as suggested by the modulation we have presented here, implies that dLGN visual stimulus coding is dynamic with potential consequences for a downstream processing and decoding mechanism.

The results we have presented here, by separating spikes into tonic and burst, are limited to one fundamental characterization of activity patterns, and relate only to a single dimension of behavior. Other authors have observed a diversity in arousal-related modulations both between (Shimaoka et al., 2018; Williamson et al., 2015) and within (Aydın et al., 2018; Liang et al., 2020; Molnár et al., 2021; Schröder et al., 2020) sensory systems. Therefore future work will need to establish if other thalamic nuclei, or neuronal sub-types within the dLGN (Kerschensteiner & Guido, 2017; Piscopo et al., 2013; Rosón et al., 2019), are similarly modulated by pupil size dynamics. Finally, although the pupil size signal is often lauded as an index for internal modulations that occur in the absence of other overt behaviors, considering pupil size in the context of other behaviors such as locomotion (Reimer et al., 2014; Vinck et al., 2015) or eye movements (Schwartz et al., 2020) promises to reveal a more diverse coupling to neural measures. Indeed, expanding the notion of behavioral state to a high-dimensional set of latent behaviors has been successful in predicting neural activity, especially in thalamic neurons (Stringer et al., 2019). We therefore expect that expanding the definition of arousal to include multi-dimensional behaviors and measuring specific stimulus responses will be essential for a more principled understanding of modulations of stimulus processing by behavioral

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

states.

Star Methods

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

Surgical procedure

Experiments were carried out in 7 adult transgenic PV-Cre mice (median age at first recording session: 22.3 weeks; B6;129P2-Pvalb$^{tm1(cre)Arbr}$/J; Jackson Laboratory) and 4 adult Ntsr1-Cre mice (median age: 21.6 weeks; B6.FVB(Cg)-Tg(Ntsr1-cre)GN220Gsat/Mmcd; MMRRC) of either sex. Transgenic mice were used, as these mice were also included in another study (Spacek et al., 2020) which required selective viral expression of ChR2 in area V1. For the present study, all data came from experiments where no optogenetic manipulation was present, with the exception of the data in Figure 3C and Figure S3C. Surgical procedures were carried out as described in (Spacek et al., 2020).

Thirty minutes prior to surgical procedures, mice were injected with an analgesic (Metamizole, 200 mg/kg, sc, MSD Animal Health, Brussels, Belgium). To induce anesthesia, animals were placed in an induction chamber and exposed to isoflurane (5% in oxygen, CP-Pharma, Burgdorf, Germany). After induction of anesthesia, mice were fixated in a stereotaxic frame (Drill & Microinjection Robot, Neurostar, Tuebingen, Germany). At the beginning of the surgical procedure, an additional analgesic was administered (Buprenorphine, 0.1 mg/kg, sc, Bayer, Leverkusen, Germany) and the isoflurane level was lowered (0.5%–2% in oxygen), such that a stable level of anesthesia could be achieved as judged by the absence of a pedal reflex. Throughout the procedure, the eyes were covered with an eye ointment (Bepanthen, Bayer, Leverkusen, Germany) and a closed loop temperature control system (ATC 1000, WPI Germany, Berlin, Germany) ensured that the animal's body temperature was maintained at 37°C. The animal's head was shaved and thoroughly disinfected using iodine solution (Braun, Melsungen, Germany). Before performing a scalp incision along the midline, a local analgesic was delivered (Lidocaine hydrochloride, sc, bela-pharm, Vechta, Germany). The skin covering the skull was partially removed and cleaned from tissue residues with a drop of H_2O_2 (3%, AppliChem, Darmstadt, Germany). Using four reference points (bregma, lambda, and two points 2 mm to the left and to the right of the midline respectively), the animal's head was positioned into a skull-flat configuration. The exposed skull was covered with OptiBond FL primer and adhesive (Kerr dental, Rastatt, Germany) omitting three locations: V1 (AP: −2.8 mm, ML: −2.5 mm), dLGN (AP: −2.3 mm, ML: −2 mm), and a position roughly 1.5 mm anterior and 1 mm to the right of bregma, designated for a miniature reference screw (00-96 X 1/16 stainless steel screws, Bilaney) soldered to a custom-made connector pin. Unrelated to the purpose of this study, 2 μL of the adeno-associated viral vector rAAV9/1.EF1a.DIO.hChR2(H134R)-eYFP.WPRE.hGH

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(Addgene, #20298-AAV9) was dyed with 0.3 μL fast green (Sigma-Aldrich, St. Louis, USA), and after performing a small craniotomy over V1, a total of ∼0.5 μL of this mixture was injected across the entire depth of cortex (0.05 μL injected every 100 μm, starting at 1000 μm and ending at 100 μm below the brain surface), using a glass pipette mounted on a Hamilton syringe (SYR 10 μL 1701 RN no NDL, Hamilton, Bonaduz, Switzerland). A custom-made lightweight stainless steel head bar was positioned over the posterior part of the skull such that the round opening contained in the bar was centered on V1/dLGN and attached with dental cement (Ivoclar Vivadent, Ellwangen, Germany) to the primer/adhesive. The opening was later filled with the silicone elastomer sealant Kwik-Cast (WPI Germany, Berlin, Germany). At the end of the procedure, an antibiotic ointment (Imax, Merz Pharmaceuticals, Frankfurt, Germany) was applied to the edges of the wound and a long-term analgesic (Meloxicam, 2 mg/kg, sc, Böhringer Ingelheim, Ingelheim, Germany) was administered and continued to be administered for 3 consecutive days. For at least 5 days post-surgery, the animal's health status was assessed via a score sheet.

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 described for the initial surgery, and a craniotomy (ca. 1.5 mm^2) was performed over dLGN and V1 and re-sealed with Kwik-Cast (WPI Germany, Berlin, Germany). 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.

Electrophysiological recordings

Mice were head-fixed on a sytrofoam ball and allowed to run freely. Extracellular signals were recorded at 30 kHz (Blackrock Microsystems). For each recording session, the silicon plug sealing the craniotomy was removed. To record from dLGN, a 32 channel linear silicon probe (Neuronexus A1x32Edge-5mm-20-177-A32, Ann Arbor, USA) was lowered to a depth of ∼2700–3700 μL below the brain surface. We judged recording sites to be located in dLGN based on the characteristic progression of RFs from upper to lower visual field along the electrode shank Piscopo et al. (2013), the presence of responses strongly modulated at the temporal frequency of the drifting gratings (F1 response), and the preference of responses to high temporal frequencies Grubb & Thompson (2003); Piscopo et al. (2013). For *post hoc* histological reconstruction of the recording site, the electrode was stained with DiI (Invitrogen, Carlsbad, USA) for one of the final recording sessions.

For the purposes of a different study (Spacek et al., 2020), during recordings involving naturalistic movie stimulation (Figure 3C and Figure S3C), V1 was optogenetically stimulated using 470 nm light on half of the trials, which were randomly interleaved with control trials. For the purposes of this study, only electrophysiological data from the control trials was considered.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

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 latency period were transcardially perfused under deep anesthesia using a cocktail of Medetomidin (Domitor, 0.5 mg/kg, Vetoquinol, Ismaning, Germany), Midazolam (Climasol, 5 mg/kg, Ratiopharm, Ulm, Germany) and Fentanyl (Fentadon, 0.05 mg/kg, Dechra Veterinary Products Deutschland, Aulendorf, Germany). A few animals, which were treated according to a different license, were anesthetized with sodium pentobarbital (Narcoren, 400 mg/kg, ip, Böhringer Ingelheim, Ingelheim, Germany). Perfusion was first done with Ringer's lactate solution followed by 4% paraformaldehyde (PFA) in 0.2 M sodium phosphate buffer (PBS).

Histology

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, Olympus, Tokyo, Japan) was used to inspect slices for the presence of yellow fluorescent protein (eYFP) and DiI. Recorded images were processed using FIJI (Rueden et al., 2017; Schindelin et al., 2012).

Visual stimulation

Visual stimulation was presented using custom written software (EXPO, https://sites.google.com/a/nyu.edu/expo/home) on a liquid crystal display (LCD) monitor (Samsung SyncMaster 2233RZ, 47×29 cm, 1680×1050 resolution at 60 Hz, mean luminance 50 cd/m^2) positioned at a distance of 25 cm from the animal's right eye. The data presented in all figures, with the exception of Figure 3C and Figure S3C, was recorded while animals were viewing either a static gray screen ($N = 3$ experiments) or a sparse noise stimulus ($N = 16$ experiments). The sparse noise stimulus consisted of non-overlapping white and black squares flashed for 200 ms on a square grid spanning 60 deg, while individual squares spanned 5 deg. The data presented in Figure 3C and Figure S3C were recorded while the animals were repeatedly presented with 5 s naturalistic movie clips (as described in Spacek et al., 2020).

Behavioral tracking

During electrophysiological recording, head-fixed mice were allowed to run on an air cushioned Styrofoam ball. Ball movements were recorded by two optical computer mice which interfaced with a microcontroller (Arduino Duemilanove) and sampled ball movements at 90 Hz. Locomotor activity was quantified by computing the Euclidean norm of three perpendicular components of ball velocity (pitch, roll, and yaw; Dombeck et al., 2007), and herein referred to as run speed. Running bouts were defined as periods of time where the ball speed exceeded 1 cm/s for at least 1 s, with a break of no more than 2 s. In order to account for the influence of locomotion-related arousal

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

that precedes and follows the bouts of locomotion (Vinck et al., 2015), quiescence was defined as any period of time outside of a running bout, excluding the 5 s immediately preceding and following a running bout.

In order to track pupil size, the stimulus-viewing eye was illuminated with and infrared LED light, and the eye was filmed with a camera (Guppy AVT camera; frame rate 50 Hz, Allied Vision, Exton, USA) equipped with a zoom lens (Navitar Zoom 6000). Pupil size was extracted from the videos using a custom, semi-automated algorithm. Each video frame was equalized using an adaptive bi-histogram equalization procedure, and then smoothed using median and bilateral filters. The center of the pupil was initially estimated by taking the darkest point in a convolution of the filtered image with a black square. Next, the peaks of the image gradient along lines extending radially from the center point were used to define the pupil contour. Lastly, an ellipse was fit to the contour, and the area of the ellipse was taken as pupil size. Frames in which the eye was obscured, occurring during eye closure or grooming, were detected by applying a threshold to the mean pixel-wise difference between each frame and a reference frame compiled by taking the median of several manually selected frames during which the eye was open. Data points with eye closure, as well as the three immediately adjacent points, were excluded. Adjustable parameters in the above algorithm were set manually for each experiment. After ellipse fitting, data points at which the fitted pupil position or size were outside of a plausible range were removed from consideration. Unreliable segments, occurring due to eye-closure, grooming, or unstable ellipse fitting, were automatically removed according to *a priori* criteria. Cubic splines were used to interpolate over gaps of < 2 s, and the remaining separated segments of continuous data were smoothed with a Gaussian kernel ($\sigma = 60$ ms) and analyzed individually.

Pupil size signal analysis

Pupil size signal (PS) segments were decomposed into intrinsic mode functions (IMFs) by empirical mode decomposition (EMD) using the PyEMD Python package (https://github.com/laszukdawid/PyEMD). EMD was performed with stringent criteria for termination of the sifting process, such as a low change in energy between two siftings (which favors interpretability of the instantaneous phase over accuracy of the instantaneous amplitude), and termination of the decomposition, requiring that all of the power of the original signal be captured by the resulting components. In contrast to the original broad-band PS signal, the IMFs are locally narrow-band and thus amenable to Hilbert spectral analysis (Huang et al., 1998). To minimize edge effects, each IMF was extrapolated from the beginning and end by mirroring the three closest peaks/troughs across the signal boundary, and interpolating between the mirrored extrema with 3rd-order Bernstein polynomials restricted by gradient at the signal edge. IMFs with fewer than 3 extrema were extended by mirroring the first derivatives. The Hilbert transform was applied to the extended IMFs using the SciPy signal library to obtain the complex-valued analytic signal of each IMF, from which the instantaneous phase and amplitude were respectively computed as the angle and length of the signal vector at each timepoint, and instantaneous frequency was given by the time derivative of the unwrapped phase. Only timepoints corresponding to the un-extrapolated signal were used in further analysis. Each IMF was assigned

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

a characteristic frequency by taking the mean frequency across all time points weighted by the amplitude. The relative power of each IMF was assigned by taking the power (squared amplitude) density of each IMF and expressing it as fraction of the sum of the power densities from all IMFs resulting from the decomposition.

In order to assess whether the above Hilbert-Huang transform (HHT) preserved the spectral power density of the PS, the "marginal power spectrum" was computed from the HHT representation of the PS by binning frequency (into identical bins as those obtained from the FFT of the PS segment), and collecting power values in each frequency bin across time and IMFs. This spectral power function was compared to the Fourier SPD by Pearson correlation.

At this point, some of the lower frequency IMFs were eliminated from further consideration if they did not complete at least four oscillations within the PS segment. Certain low power IMFs were also excluded from further consideration if their power fell below the 95% confidence bounds the frequency-power relationship expected from EMD of white noise (Wu & Huang, 2004). The impact of eliminating IMFs was assessed by computing the fraction of the variance of the original PS explained by a reconstruction of the PS from only the remaining, valid IMFs.

The relationships between IMFs in a decomposition were assessed in three ways. Linear dependence was assessed by computing the Pearson correlation between all pairs of IMFs in a set. Higher-order dependence was assessed by performing ICA on the IMF set using the FastICA algorithm implemented in the scikit-learn Python toolbox. The resulting de-mixing matrices were normalized such that the absolute values of all the contributions to an IMF summed to 1, then the contributions to each IC were Z-scored. The normalized de-mixing coefficients were split into peak contributors (the maximum coefficient for each IC) and non-peak contributions. The "independence" of IMFs was assessed by grouping together the de-mixing coefficients from all decompositions, and comparing the mean value of peak contributions to the mean of non-peak contributions. This value, 2.06 s.d., was compared to a distribution of 1000 analogous distance values, which was compiled by simulating the same number of de-mixing matrices by drawing all coefficients for each IC from a single normalized Gaussian distribution. The actual distance between peak and non-peak distributions was larger than all 1000 simulated distances ($p = 0.000$), indicating that IMFs are overall more independent than would be expected if independent sources were randomly mixed. Finally, as EMD can act as a dyadic filter (Wu & Huang, 2004) producing components that may represent harmonics of the same fundamental oscillation, the phase relationships between IMFs in a set were assessed by computing the spectral coherence for each IMF pair. For each IMF, multi-taper estimates of the spectral coherence between all other IMFs in the set was computed, and values were taken only in the frequency band defined by the characteristic frequency of the primary IMF. We note that, although no definitive conclusions should be drawn, the low values of all these metrics support the view that IMFs are independently occurring oscillations, and do not simply reflect different aspects of the same process.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Spike sorting

The spiking activity of isolated single units was initially extracted from extracellular recordings using the Kilosort spike-sorting toolbox (Pachitariu et al., 2016). The resulting spike clusters were subject to manual curation in Spyke (Spacek et al., 2009), where spikes within a cluster were temporally aligned and plotted in a 3D space (multichannel PCA, ICA, and/or spike time). In this space, clusters could be merged to account for drift in spike shape over the recording session (for example, if the first two wave shape PCs changed smoothly as a function of spike time), or further split using a gradient-ascent based clustering algorithm (GAC; Swindale & Spacek, 2014). Clusters containing only spikes with no consistent and clearly discernible voltage deflection were eliminated from further consideration. Finally, cluster auto-correlograms were examined to ensure that a clear refractory period was present; while the presence of a refractory period was not an indicator of a well-isolated unit, the absence of a refractory period was taken as an indicator that the cluster might need to be further split, or contained a high amount of noise. Remaining clusters were compared using empirical distance metrics to ensure that they were well separated.

In total, 246 clusters were sufficiently well isolated and taken as neural units. The analyses of data recorded during presentation of a gray screen or sparse noise stimulus (all figures except Figure 3C and Figure S3C) included 176 units. These units could appear in multiple PS segments, giving a total of 677 unit-PS segment combinations. The analyses of data recorded during naturalistic movie presentation (Figure 3C and Figure S3C) included 70 units contributing to 1077 unit-PS segment combinations. Some recording session involved both a grey screen/ sparse noise experiment and a naturalistic movie experiments, thus some units (54/246) appeared in both sets of analyses.

Spiking metrics

We identified bursts in our isolated dLGN units as two or more spikes occurring after a period of inactivity of at least 100 ms, and separated by an inter-spike interval of no more than 4 ms (Lu et al., 1992). All spikes that were not part of a burst, according to theses criteria, were labelled as tonic spikes. Burst ratio was defined as the proportion of all spikes in the recording period that were part of a burst.

Short timescale serial dependence of burst events and tonic spiking was assessed by binning event counts (10 ms bins) and computing the normalized auto-correlation of each event train in a -1 s to 1 s window. Taking the average auto-correlation over all units, it was determined by visual inspection that both event types had a primary peak of serial dependence lasting approximately 300 ms. Although we also note that the mean burst auto-correlation has a secondary peak, indicating rhythmicity in the ~ 4 Hz range (Nestvogel & McCormick, 2021).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Circular statistics

In order to assess the modulation of spiking activity by arousal states, we developed a phase tuning analysis to relate bursts and tonic spikes to phases of the pupil size signal (PS) components (IMFs). We required that at least 8 spiking events occurred over the course of the PS segment for a unit to be considered in this analysis (tonic spikes: 632/677 unit-PS segment combinations; bursts: 399/677 unit-PS combinations). Considering tonic spikes and burst events from each unit separately, the Hilbert phase of each ongoing IMF was collected at times when bursts (the time of the first spike of the burst) and tonic spikes occurred. Although modulation might be directly assessed by compiling phase histograms and computing the circular mean of these phases, we adopted various bias-corrections resulting in a more stringent assessment of phase tuning.

Firstly, IMFs are not guaranteed to have a linear phase progression, which is beneficial in that IMFs can capture naturalistic asymmetric wave shapes, but presents a difficulty when assessing phase tuning. We quantified the phase bias of the PS-IMFs by constructing a phase bias index (PBI):

$$\mathrm{PBI} = \frac{\max\{P(\psi)\} - \min\{P(\psi)\}}{\max\{P(\psi)\}} \ , \tag{1}$$

in which $P(\psi)$, the probability of observing a certain phase, was computed for 8 phase bins. A PBI of 0 would indicate that an equal amount of time is spent in all phases, whereas a PBI of 1 would indicate that a certain phase bin is never visited. A PBI of 1 never occurred as we excluded IMFs that did not visit all phase bins at least 4 times from consideration; nonetheless PBI was always > 0, and increased when considering only periods of locomotion or quiescence, reflecting the fact that pupil size increases and decreases can be correlated with locomotion. These biases could lead one to falsely infer consistent phase tuning if $\mathrm{argmax}\{P(\psi)\}$ is similar across IMFs, or stronger phase tuning for low frequency IMFs that tended to have higher PBIs. To account for these biases in the IMF phase distributions, we converted phases to circular ranks, yielding a uniform underlying distribution to which bursts and tonic spikes could be related (Siapas et al., 2005). The angle of the circular mean rank was then converted back to a phase in the original distribution, and this value was reported as the tuning phase.

Next, as the circular mean resultant vector length is a biased statistic, we opted for an unbiased measure of phase tuning strength in order to allow comparison of tuning strength across units with different firing rates, between tonic spikes and bursts, and between conditions where spikes were sub-sampled from the recording period. Rather than using mean resultant vector length (R), we used its squared value, which can be easily bias-corrected (Kutil, 2012). We thus quantified tuning strength as

$$\hat{R}^2 = \frac{n}{n-1}(R^2 - \frac{1}{n}) \ , \tag{2}$$

where R is the mean resultant vector length computed using the circular phase ranks, and n is the number of tonic spikes or bursts.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Lastly, we sought to assess the statistical significance of phase tuning with a measure that was insensitive to the short-term serial dependence of spike trains, which effectively reduces the degrees of freedom of the sample. We adopted the null hypothesis that, if there were no relationship between spiking activity and IMF phase, then the tuning strength of an IMF-unit pair would be the same as that measured from a spike train with the same short-term structure, but with no relationship to the IMF. We tested tuning strengths against this null hypothesis by splitting burst and tonic spike trains into segments of 300 ms, which were shuffled in order to destroy the relationship between IMF phase and spiking activity, and then phase tuning to the IMF was assessed as described above. This procedure was repeated 1000 times for each IMF-unit pair, to compile a null set of tuning strengths. The actual tuning strength, assessed using the intact spike train, was then compared to this null set, and assigned a p-value based on how many elements of the null set had a higher tuning strength value.

Correlation analyses

Variables such as locomotion and ambient luminance are known to be reflected in the pupil size signal. In order to assess the contributions of these variables to the phase tuning we observed, we performed correlation analyses relating each IMF to the animal's locomotion speed or the stimulus brightness. The appropriate lag at which to assess the IMF correlation was determined using the correlation of the raw PS with each signal type. In the case of run speed, the mean cross-correlation had a peak at 1.08 s. As a proxy for the ambient luminance changes induced by the naturalistic movie clips, gamma-corrected pixel intensity values were averaged across all pixels and color channels for each movie frame. This value is referred to as stimulus brightness. The PS stimulus brightness cross-correlation had no strong negative peak in the expected time window in the mean cross-correlation (data not shown). Therefore, the peak of the absolute cross-correlation for each PS was taken, and the mean lag of all the peaks was found to be 0.38 s, which falls between two previous reports of the delay between pupil size and brightness changes in mice (Grozdanic et al., 2003; Lucas et al., 2001). The values of the IMF cross-correlations at these lags were then used to assess correlation strength and the significance of the correlations.

Because serial dependence in a signal inflates cross-correlation values and violates the independence assumption required to directly calculate a p-value, we adopted two permutation-based methods in order to assess the statistical significance of IMFs correlations (Harris, 2020). For IMF run speed correlations, each IMF was correlated with 1000 run speed traces from other experiments. For IMF stimulus brightness correlations, each IMF was correlated to 1000 mock brightness time-courses, compiled by randomly concatenating the brightness time-courses of the set of 20 naturalistic movie clips (5 s duration) used in our experiments, randomly interspersed with 1 s of gray screen normally displayed during inter-trial-intervals. In both cases, the actual correlation was compared to these sets of nonsense correlations in order to obtain a p-value for the true correlation of the IMF with the signal in question.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Tuning strength regressions

Linear regressions were performed in order to assess the contributions of various factors to the observed link between PS-IMFs and spiking activity. In all cases, the full model was specified by:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 \tag{3}$$

Where Y is burst or tonic tuning strength, and X is either IMF power, IMF run speed correlation, or IMF stimulus brightness correlation. Tuning strengths for all IMF-unit pairs were included, regardless of tuning significance. Model fitting was done by ordinary least-squares, implemented in the statsmodels Python library (Seabold & Perktold, 2010). For IMF power, the model only included the linear term. For IMF run speed and brightness correlations, an F-test was used to compare the performance of the full model to a restricted model which included only the linear term. Regression results were reported for the full model only if the squared term made a significant contribution, otherwise the results from the restricted model were reported.

Pupil size distribution-matching

The phase tuning observed in our data indicates that arousal-related activity modulations are reflected in aspects of the pupil size signal beyond whether the pupil is dilated or constricted. This observation presents the opportunity to control for changes in retinal irradiance that occur with pupil size changes as a possible confound for the modulations we observed. This would also contribute to an ongoing debate in the literature, whether modulations of neuronal activity are best linked to pupil size *per se* McGinley et al. (2015a) or its ongoing changes Reimer et al. (2014): The phase tuning we observed across multiple frequency bands suggest that the explanatory power of pupil size *per se* might arise because high-power, low-frequency oscillations (< 0.1 Hz) are associated with spiking modulations centered around peaks and troughs, which likely correspond to times when the pupil is fully dilated or constricted.

In order to remove any effects of pupil size *per se* from the measurement of IMF phase tuning, we adopted a histogram matching procedure (Churchland et al., 2010) to equalize the distribution of pupil sizes across the phase bins of each IMF. For each IMF, the recording was split into segments based on four phase bins (early dilation, late dilation, early contraction, late contraction). For each visit to a phase bin, the mean pupil size during the visit was collected, and used to compile a pupil size histogram. Time periods were then randomly sub-sampled from the recording such that all phase bins had a equal number of visits in each decile of pupil size. The number of visits was determined by taking, for each decile, the smallest number of visits across pupil size bins. Tuning strength and phase were then assessed using only spikes occurring during these sub-sampled visits. The reported tuning strengths and phases were the means over 100 repetitions of this sub-sampling procedure. Tuning significance was assessed using this mean tuning strength, and comparing it to a set of 1000 tuning strengths complied by shuffling 300 ms bins of spiking data. On each of these 1000 shuffles, the spiking data came from one of the 100 time range

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

sub-samplings.

Acknowledgements

This research was supported by the German Research Foundation (DFG) RTG2175 "Perception in context and its neural basis" (CL), DFG SFB870 "Assembly and function of neural circuits", TP 19 (118803580) (LB), and by SmartStart program of the Bernstein Network for Computational Neuroscience funded by the Volkswagen Foundation (DC). We would like to thank M. Sotgia for lab management and support with animal handling and histology, Y. Bauer, G. Born, and A. Kotkat for their contributions to the data collections and spike sorting, S. Schörnich for IT support, and B. Grothe for providing excellent research infrastructure.

Author Contributions

Davide Crombie: Conceptualization, Methodology, Software, Formal analysis, Data Curation, Writing, Writing - Review & Editing, Visualization, Funding acquisition. **Martin A. Spacek**: Software, Investigation, Data Curation. **Laura Busse**: Conceptualization, Methodology, Writing, Writing - Review & Editing, Project Administration, Supervision, Funding Acquisition. **Christian Leibold**: Conceptualization, Methodology, Writing, Writing - Review & Editing, Project Administration, Supervision, Funding Acquisition.

Declaration of Interests

The authors declare no competing interests.

References

Alitto, H., Rathbun, D. L., Vandeleest, J. J., Alexander, P. C., & Usrey, W. M. (2019). The augmentation of retinogeniculate communication during thalamic burst mode. Journal of Neuroscience, 39, 5697–5710.

Alitto, H. J., Weyand, T. G., & Usrey, W. M. (2005). Distinct properties of stimulus-evoked bursts in the lateral geniculate nucleus. Journal of Neuroscience, 25, 514–523.

Asanuma, C. (1992). Noradrenergic innervation of the thalamic reticular nucleus: a light and electron microscopic immunohistochemical study in rats. Journal of Comparative Neurology, 319, 299–311.

Aston-Jones, G. & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu. Rev. Neurosci., 28, 403–450.

Augustinaite, S. & Kuhn, B. (2020). Complementary ca2+ activity of sensory activated and suppressed layer 6 corticothalamic neurons reflects behavioral state. Current Biology.

Aydın, Ç., Couto, J., Giugliano, M., Farrow, K., & Bonin, V. (2018). Locomotion modulates specific functional cell types in the mouse visual thalamus. Nature communications, 9, 1–12.

Baden, T., Berens, P., Franke, K., Rosón, M. R., Bethge, M., & Euler, T. (2016). The functional diversity of retinal ganglion cells in the mouse. Nature, 529, 345–350.

Bezdudnaya, T., Cano, M., Bereshpolova, Y., Stoelzel, C. R., Alonso, J.-M., & Swadlow, H. A. (2006). Thalamic burst mode and inattention in the awake lgnd. Neuron, 49, 421–432.

Bickford, M., Günlük, A., Van Horn, S., & Sherman, S. (1994). Gabaergic projection from the basal forebrain to the visual sector of the thalamic reticular nucleus in the cat. Journal of Comparative Neurology, 348, 481–510.

Butts, D. A., Desbordes, G., Weng, C., Jin, J., Alonso, J.-M., & Stanley, G. B. (2010). The episodic nature of spike trains in the early visual pathway. Journal of neurophysiology, 104, 3371–3387.

Castro-Alamancos, M. A. & Oldford, E. (2002). Cortical sensory suppression during arousal is due to the activity-dependent depression of thalamocortical synapses. The Journal of physiology, 541, 319–331.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Cazettes, F., Reato, D., Morais, J. P., Renart, A., & Mainen, Z. F. (2021). Phasic activation of dorsal raphe serotonergic neurons increases pupil size. Current Biology, 31, 192–197.

Chen, C. & Regehr, W. G. (2003). Presynaptic modulation of the retinogeniculate synapse. Journal of Neuroscience, 23, 3130–3135.

Churchland, M. M., Byron, M. Y., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., Newsome, W. T., Clark, A. M., Hosseini, P., Scott, B. B., et al. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nature neuroscience, 13, 369–378.

Crandall, S. R., Cruikshank, S. J., & Connors, B. W. (2015). A corticothalamic switch: controlling the thalamus with dynamic synapses. Neuron, 86, 768–782.

Crochet, S. & Petersen, C. C. (2006). Correlating whisker behavior with membrane potential in barrel cortex of awake mice. Nature neuroscience, 9, 608–610.

de Gee, J. W., Knapen, T., & Donner, T. H. (2014). Decision-related pupil dilation reflects upcoming choice and individual bias. Proceedings of the National Academy of Sciences, 111, E618–E625.

de Gee, J. W., Tsetsos, K., Schwabe, L., Urai, A. E., McCormick, D., McGinley, M. J., & Donner, T. H. (2020). Pupil-linked phasic arousal predicts a reduction of choice bias across species and decision domains. Elife, 9, e54014.

Devilbiss, D. M. & Waterhouse, B. D. (2011). Phasic and tonic patterns of locus coeruleus output differentially modulate sensory network function in the awake rat. Journal of neurophysiology, 105, 69–87.

Dombeck, D. A., Khabbaz, A. N., Collman, F., Adelman, T. L., & Tank, D. W. (2007). Imaging Large-Scale Neural Activity with Cellular Resolution in Awake, Mobile Mice. Neuron, 56, 43–57.

Erisken, S., Vaiceliunaite, A., Jurjut, O., Fiorini, M., Katzner, S., & Busse, L. (2014). Effects of Locomotion Extend throughout the Mouse Early Visual System. Current Biology, 24, 2899–2907.

Gaudry, K. S. & Reinagel, P. (2008). Information measure for analyzing specific spiking patterns and applications to lgn bursts. Network: Computation in Neural Systems, 19, 69–94.

Grozdanic, S., Betts, D. M., Allbaugh, R. A., Sakaguchi, D. S., Kwon, Y. H., Kardon, R. H., & Sonea, I. M. (2003). Characterization of the pupil light reflex, electroretinogram and tonometric parameters in healthy mouse eyes. Current eye research, 26, 371–378.

Grubb, M. S. & Thompson, I. D. (2003). Quantitative characterization of visual response properties in the mouse dorsal lateral geniculate nucleus. Journal of neurophysiology, 90, 3594–3607.

Grubb, M. S. & Thompson, I. D. (2005). Visual response properties of burst and tonic firing in the mouse dorsal lateral geniculate nucleus. Journal of neurophysiology, 93, 3224–3247.

Guido, W. & Weyand, T. (1995). Burst responses in thalamic relay cells of the awake behaving cat. Journal of neurophysiology, 74, 1782–1786.

Hallanger, A. E., Levey, A. I., Lee, H. J., Rye, D. B., & Wainer, B. H. (1987). The origins of cholinergic and other subcortical afferents to the thalamus in the rat. Journal of Comparative Neurology, 262, 105–124.

Harris, K. D. (2020). Nonsense correlations in neuroscience. bioRxiv.

Harris, K. D. & Thiele, A. (2011). Cortical state and attention. Nature reviews neuroscience, 12, 509–523.

Holdefer, R. N. & Jacobs, B. L. (1994). Phasic stimulation of the locus coeruleus: Effects on activity in the lateral geniculate nucleus. Experimental Brain Research, 100, 444–452.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., & Liu, H. H. (1998). The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences, 454, 903–995.

Joshi, S. & Gold, J. I. (2020). Pupil size as a window on neural substrates of cognition. Trends in Cognitive Sciences.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. Neuron, 89, 221–234.

Kerschensteiner, D. & Guido, W. (2017). Organization of the dorsal lateral geniculate nucleus in the mouse. Visual neuroscience, 34.

Kutil, R. (2012). Biased and unbiased estimation of the circular mean resultant length and its variance. Statistics, 46, 549–561.

Larsen, R. S. & Waters, J. (2018). Neuromodulatory correlates of pupil dilation. Frontiers in neural circuits, 12, 21.

Lesica, N. A. & Stanley, G. B. (2004). Encoding of natural scene movies by tonic and burst spikes in the lateral geniculate nucleus. Journal of Neuroscience, 24, 10731–10740.

Lesica, N. A., Weng, C., Jin, J., Yeh, C.-I., Alonso, J.-M., & Stanley, G. B. (2006). Dynamic encoding of natural luminance sequences by lgn bursts. PLoS Biol, 4, e209.

Liang, L., Fratzl, A., El Mansour, O., Reggiani, J. D., Chen, C., & Andermann, M. L. (2020). Selective gating of retinal information by arousal. bioRxiv.

Lin, P.-A., Asinof, S. K., Edwards, N. J., & Isaacson, J. S. (2019). Arousal regulates frequency tuning in primary auditory cortex. Proceedings of the National Academy of Sciences, 116, 25304–25310.

Lindvall, O., Björklund, A., Nobin, A., & Stenevi, U. (1974). The adrenergic innervation of the rat thalamus as revealed by the glyoxylic acid fluorescence method. Journal of comparative Neurology, 154, 317–347.

Lu, S.-M., Guido, W., & Sherman, S. M. (1992). Effects of membrane voltage on receptive field properties of lateral geniculate neurons in the cat: contributions of the low-threshold ca2+ conductance. Journal of Neurophysiology, 68, 2185–2198.

Lu, S.-M., Guido, W., & Sherman, S. M. (1993). The brain-stem parabrachial region controls mode of response to visual stimulation of neurons in the cat's lateral geniculate nucleus. Visual neuroscience, 10, 631–642.

Lucas, R. J., Douglas, R. H., & Foster, R. G. (2001). Characterization of an ocular photopigment capable of driving pupillary constriction in mice. Nature neuroscience, 4, 621–626.

Mackay-Sim, A., Sefton, A. J., & Martin, P. R. (1983). Subcortical projections to lateral geniculate and thalamic reticular nuclei in the hooded rat. Journal of comparative neurology, 213, 24–35.

Mathôt, S. (2018). Pupillometry: Psychology, physiology, and function. Journal of Cognition, 1.

McCormick, D. A. (1992). Neurotransmitter actions in the thalamus and cerebral cortex and their role in neuromodulation of thalamocortical activity. Progress in neurobiology, 39, 337–388.

McGinley, M. J. (2020). Brain states: Sensory modulations all the way down. Current Biology, 30, R1263–R1266.

McGinley, M. J., David, S. V., & McCormick, D. A. (2015a). Cortical membrane potential signature of optimal states for sensory signal detection. Neuron, 87, 179–192.

McGinley, M. J., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C. R., Tolias, A. S., Cardin, J. A., & McCormick, D. A. (2015b). Waking state: rapid variations modulate neural and behavioral responses. Neuron, 87, 1143–1161.

Mease, R. A., Kuner, T., Fairhall, A. L., & Groh, A. (2017). Multiplexed spike coding and adaptation in the thalamus. Cell reports, 19, 1130–1140.

Molnár, B., Sere, P., Bordé, S., Koós, K., Zsigri, N., Horváth, P., & Lőrincz, M. L. (2021). Cell-Type Specific Arousal-Dependent Modulation of Thalamic Activity in the Lateral Geniculate Nucleus. Cerebral Cortex Communications. tgab020.

Mukherjee, P. & Kaplan, E. (1995). Dynamics of neurons in the cat lateral geniculate nucleus: in vivo electrophysiology and computational modeling. Journal of neurophysiology, 74, 1222–1243.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. Nature neuroscience, 15, 1040.

Nelson, A. & Mooney, R. (2016). The basal forebrain and motor cortex provide convergent yet distinct movement-related inputs to the auditory cortex. Neuron, 90, 635–648.

Nestvogel, D. B. & McCormick, D. A. (2021). Visual thalamocortical mechanisms of waking state dependent activity and alpha oscillations. bioRxiv.

Niell, C. M. & Stryker, M. P. (2010). Modulation of visual responses by behavioral state in mouse visual cortex. Neuron, 65, 472–479.

Pachitariu, M., Steinmetz, N. A., Kadir, S. N., Carandini, M., & Harris, K. D. (2016). Fast and accurate spike sorting of high-channel count probes with KiloSort. In Advances in Neural Information Processing Systems 29, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, & R. Garnett, eds. (Curran Associates, Inc.), pp. 4448–4456.

Pape, H.-C. & McCormick, D. A. (1989). Noradrenaline and serotonin selectively modulate thalamic burst firing by enhancing a hyperpolarization-activated cation current. Nature, 340, 715–718.

Parikh, V., Kozak, R., Martinez, V., & Sarter, M. (2007). Prefrontal acetylcholine release controls cue detection on multiple timescales. Neuron, 56, 141–154.

Petty, G. H., Kinnischtzke, A. K., Hong, Y. K., & Bruno, R. M. (2021). Effects of arousal and movement on secondary somatosensory and visual thalamus. bioRxiv.

Piscopo, D. M., El-Danaf, R. N., Huberman, A. D., & Niell, C. M. (2013). Diverse visual features encoded in mouse lateral geniculate nucleus. Journal of Neuroscience, 33, 4642–4656.

Poulet, J. F. & Crochet, S. (2019). The cortical states of wakefulness. Frontiers in systems neuroscience, 12, 64.

Rajkowski, J., Kubiak, P., & Aston-Jones, G. (1994). Locus coeruleus activity in monkey: phasic and tonic changes are associated with altered vigilance. Brain research bulletin, 35, 607–616.

Ramcharan, E. J., Gnadt, J. W., & Sherman, S. M. (2000). Burst and tonic firing in thalamic cells of unanesthetized, behaving monkeys. Visual neuroscience, 17, 55–62.

Reimer, J., Froudarakis, E., Cadwell, C. R., Yatsenko, D., Denfield, G. H., & Tolias, A. S. (2014). Pupil fluctuations track fast switching of cortical states during quiet wakefulness. Neuron, 84, 355–362.

Reimer, J., McGinley, M. J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D. A., & Tolias, A. S. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nature communications, 7, 1–7.

Reinagel, P., Godwin, D., Sherman, S. M., & Koch, C. (1999). Encoding of visual information by lgn bursts. Journal of neurophysiology, 81, 2558–2569.

Rivadulla, C., Martinez, L., Grieve, K. L., & Cudeiro, J. (2003). Receptive field structure of burst and tonic firing in feline lateral geniculate nucleus. The Journal of physiology, 553, 601–610.

Rodenkirch, C., Liu, Y., Schriver, B. J., & Wang, Q. (2019). Locus coeruleus activation enhances thalamic feature selectivity via norepinephrine regulation of intrathalamic circuit dynamics. Nature neuroscience, 22, 120–133.

Rosón, M. R., Bauer, Y., Kotkat, A. H., Berens, P., Euler, T., & Busse, L. (2019). Mouse dlgn receives functional input from a diverse population of retinal ganglion cells with limited convergence. Neuron, 102, 462–476.

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., & Eliceiri, K. W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. BMC Bioinf., 18.

Ruivo, L. M. T.-G., Baker, K. L., Conway, M. W., Kinsley, P. J., Gilmour, G., Phillips, K. G., Isaac, J. T., Lowry, J. P., & Mellor, J. R. (2017). Coordinated acetylcholine release in prefrontal cortex and hippocampus is associated with arousal and reward on distinct timescales. Cell reports, 18, 905–917.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. Nat. Methods, 9, 676–682.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Schriver, B. J., Bagdasarov, S., & Wang, Q. (2018). Pupil-linked arousal modulates behavior in rats performing a whisker deflection direction discrimination task. Journal of Neurophysiology, 120, 1655–1670.

Schröder, S., Steinmetz, N. A., Krumin, M., Pachitariu, M., Rizzi, M., Lagnado, L., Harris, K. D., & Carandini, M. (2020). Arousal modulates retinal output. Neuron.

Schwartz, Z. P., Buran, B. N., & David, S. V. (2020). Pupil-associated states modulate excitability but not stimulus selectivity in primary auditory cortex. Journal of Neurophysiology, 123, 191–208.

Seabold, S. & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. In 9th Python in Science Conference.

Sherman, S. M. (2001). Tonic and burst firing: dual modes of thalamocortical relay. Trends in neurosciences, 24, 122–126.

Sherman, S. M. & Guillery, R. (1996). Functional organization of thalamocortical relays. Journal of neurophysiology, 76, 1367–1395.

Shimaoka, D., Harris, K. D., & Carandini, M. (2018). Effects of arousal on mouse sensory cortex depend on modality. Cell reports, 22, 3160–3167.

Siapas, A. G., Lubenov, E. V., & Wilson, M. A. (2005). Prefrontal phase locking to hippocampal theta oscillations. Neuron, 46, 141–151.

Sokhadze, G., Campbell, P. W., & Guido, W. (2019). Postnatal development of cholinergic input to the thalamic reticular nucleus of the mouse. European Journal of Neuroscience, 49, 978–989.

Sokhadze, G., Seabrook, T. A., & Guido, W. (2018). The absence of retinal input disrupts the development of cholinergic brainstem projections in the mouse dorsal lateral geniculate nucleus. Neural development, 13, 1–15.

Spacek, M. A., Blanche, T. J., & Swindale, N. V. (2009). Python for large-scale electrophysiology. Front. Neuroinform., 2, 9.

Spacek, M. A., Born, G., Crombie, D., Bauer, Y., Liu, X., Katzner, S., & Busse, L. (2020). Robust effects of corticothalamic feedback during naturalistic visual stimulation. bioRxiv.

Steriade, M. & Llinás, R. R. (1988). The functional states of the thalamus and the associated neuronal interplay. Physiological reviews, 68, 649–742.

Steriade, M., McCormick, D. A., & Sejnowski, T. J. (1993). Thalamocortical oscillations in the sleeping and aroused brain. Science, 262, 679–685.

Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C. B., Carandini, M., & Harris, K. D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. Science, 364, eaav7893.

Sun, Y.-G., Pita-Almenar, J. D., Wu, C.-S., Renger, J. J., Uebele, V. N., Lu, H.-C., & Beierlein, M. (2013). Biphasic cholinergic synaptic transmission controls action potential activity in thalamic reticular nucleus neurons. Journal of Neuroscience, 33, 2048–2059.

Sundberg, S. C., Lindström, S. H., Sanchez, G. M., & Granseth, B. (2018). Cre-expressing neurons in visual cortex of ntsr1-cre gn220 mice are corticothalamic and are depolarized by acetylcholine. Journal of Comparative Neurology, 526, 120–132.

Swadlow, H. A. & Gusev, A. G. (2001). The impact of'bursting'thalamic impulses at a neocortical synapse. Nature neuroscience, 4, 402–408.

Swindale, N. V. & Spacek, M. A. (2014). Spike sorting for polytrodes: a divide and conquer approach. Front. Syst. Neurosci., 8, 6.

van Kempen, J., Loughnane, G. M., Newman, D. P., Kelly, S. P., Thiele, A., O'Connell, R. G., & Bellgrove, M. A. (2019). Behavioural and neural signatures of perceptual decision-making are modulated by pupil-linked arousal. Elife, 8, e42541.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Varela, C. & Sherman, S. M. (2009). Differences in response to serotonergic activation between first and higher order thalamic nuclei. Cerebral Cortex, 19, 1776–1786.

Vinck, M., Batista-Brito, R., Knoblich, U., & Cardin, J. A. (2015). Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. Neuron, 86, 740–754.

Williamson, R. S., Hancock, K. E., Shinn-Cunningham, B. G., & Polley, D. B. (2015). Locomotion and task demands differentially modulate thalamic audiovisual processing during active search. Current Biology, 25, 1885–1891.

Woolf, N. J. & Butcher, L. L. (1986). Cholinergic systems in the rat brain: III. projections from the pontomesencephalic tegmentum to the thalamus, tectum, basal ganglia, and basal forebrain. Brain research bulletin, 16, 603–637.

Wu, Z. & Huang, N. E. (2004). A study of the characteristics of white noise using the empirical mode decomposition method. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 460, 1597–1611.

Zeldenrust, F., Chameau, P., & Wadman, W. J. (2018). Spike and burst coding in thalamocortical relay cells. PLOS Computational Biology, 14, 1–36.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Supplemental Information

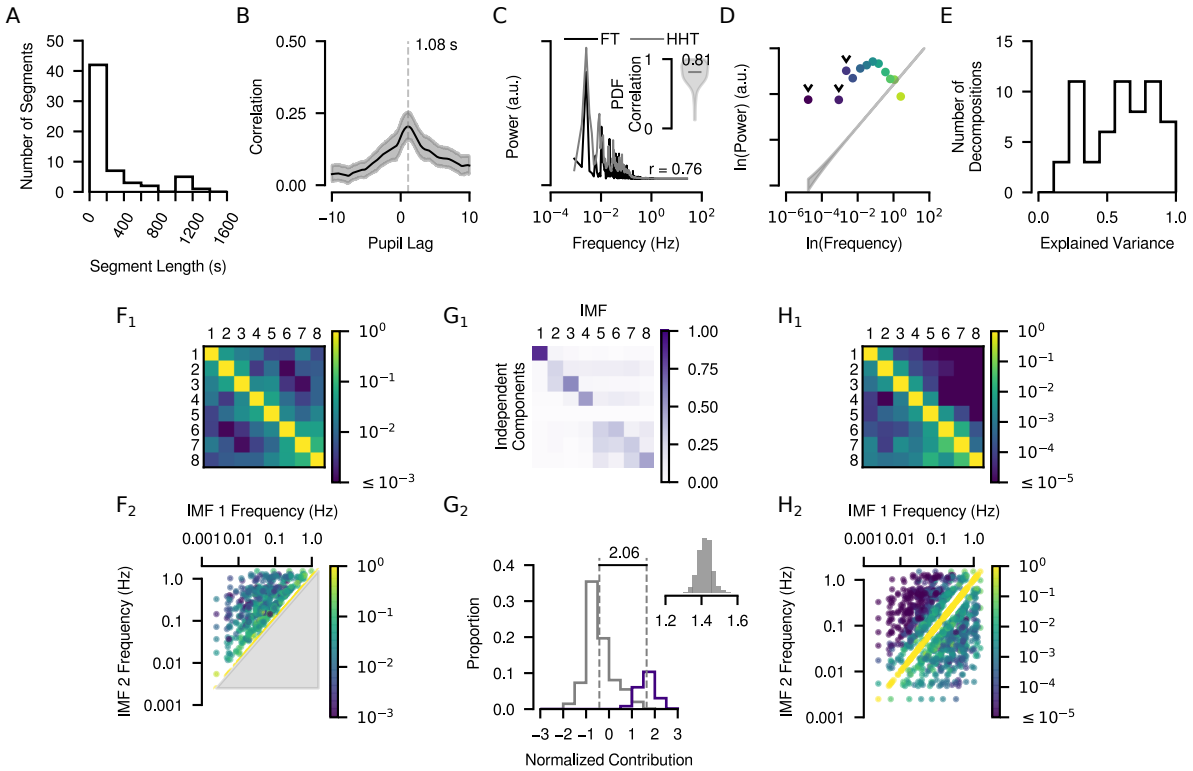


Figure S1 **(A)** Distribution of length of PS segments. N = 60 from 19 sessions and 10 mice. Median length: 158 s. **B** Normalized cross-correlation of the PS and run speed (*grey*: bootstrapped 95% CI), averaged across all PS segments. Peak: 1.08 s. At this lag, 14/59 PS segments had a significant correlation with run speed ($p < 0.05$). **(C)** Comparison of the power density function as measured by the HHT (*grey*) and the Fourier spectrum power density function (*black*) for the decomposition in Figure 1C. Pearson $r = 0.76$. Inset: distribution of PDF correlations for all PS decompositions (median Pearson $r = 0.81$, IQR: 0.71 – 0.93). **(D)** Scatter plot of the frequency-power relationship of IMFs from the example decomposition in Figure 1C. The statistical significance of IMFs can be assessed by comparing their power-frequency relationship to that expected from EMD of white noise (Wu & Huang, 2004). *Grey area*: 95% confidence bounds of the frequency-power relationship expected from white noise. IMFs with low power according to this test (e.g., the highest frequency IMF) were eliminated from further consideration. Additionally, IMFs that did not complete at least four full oscillations within the recording period (*arrow heads*) were eliminated from further consideration. **(E)** Histogram of the fraction of variance in the original PS explained by the remaining, valid IMFs. Median: 0.65, IQR: 0.42 – 0.81. **(F_1)** Pair-wise correlation matrix of the set of valid IMFs from Figure 1C. Overall correlations between IMFs were low, weak correlations were largely limited to adjacent pairs of IMFs that were similar in frequency. **(F_2)** Pair-wise correlation scatter plot for valid IMFs in all sets. Correlation values tended to be stronger close to the diagonal. **(G_1)** De-mixing matrix from ICA blind source separation applied to the IMFs in Figure 1C. Coefficients are normalized such that the contributions to each independent component (rows) sum to 1, and rows are sorted by the location of the peak value across the columns (representing the IMFs). Contributions to the independent components tend to come from adjacent IMFs. **(G_2)** Contributions from each IMF to an independent component were Z-scored, peak contributions (*purple*) were separated from non-peak contributions (*gray*), and both sets of values were compiled into histograms. Difference between the means of the peak and non-peak distributions: 2.06. *Inset*: Distribution of differences between means for peak and non-peak contributions, when all contributions are drawn from the same Gaussian distribution (sampling procedure repeated 1000 times). The difference between the actual peak and non-peak contributions was larger than expected when contributions were drawn randomly from the same distribution ($p = 0.000$), as would be expected when components were completely mixed, indicating that IMFs have a certain degree of independence. **(H_1)** Pair-wise spectral coherence matrix of the IMF set from Figure 1C. For each row, coherence was evaluated at the frequency bin closest to the characteristic frequency of IMF2 in each pair. Coherence values were normalized within each row such that the coherence of an IMF with itself in the appropriate frequency band was one. **(H_2)** Pair-wise coherence scatter plot for all IMF sets, non-zero correlation values tended to be stronger close to the diagonal.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

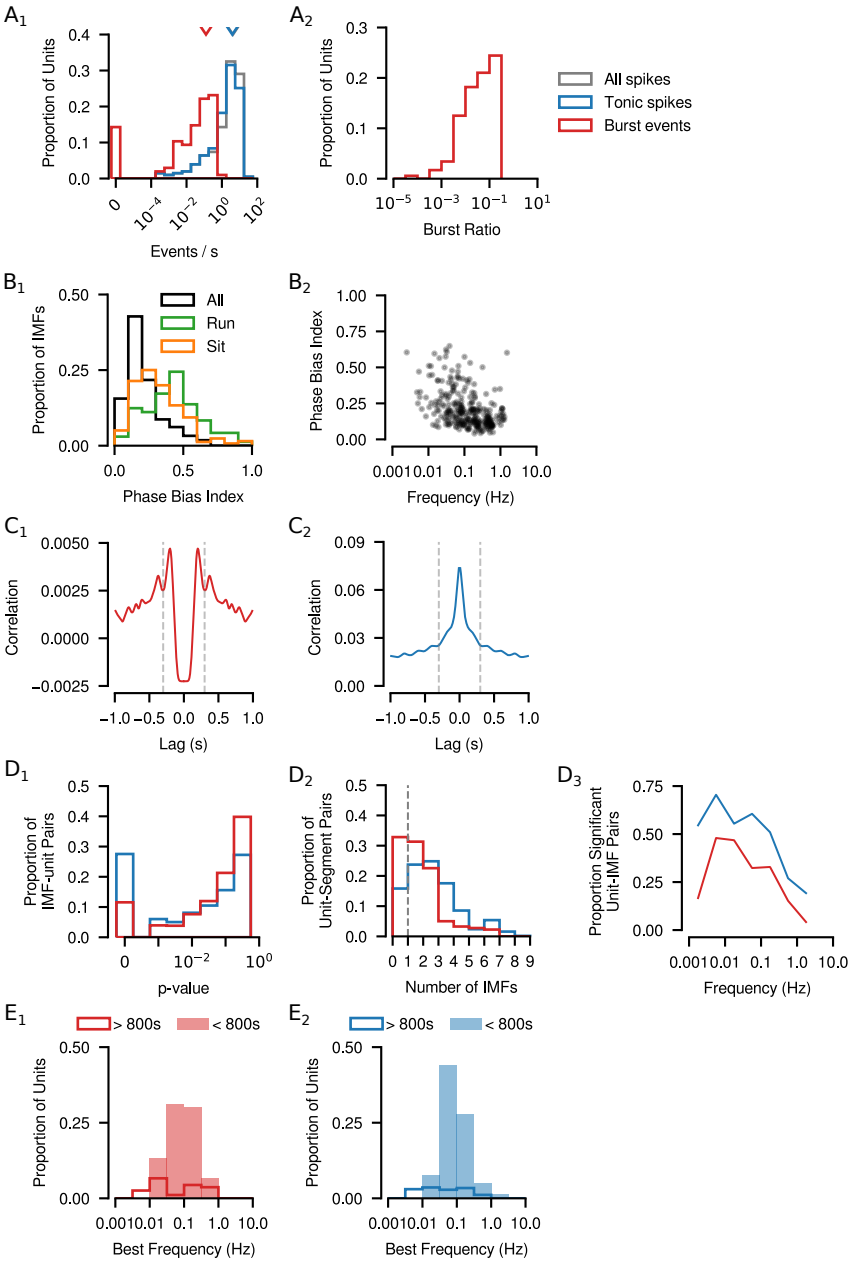
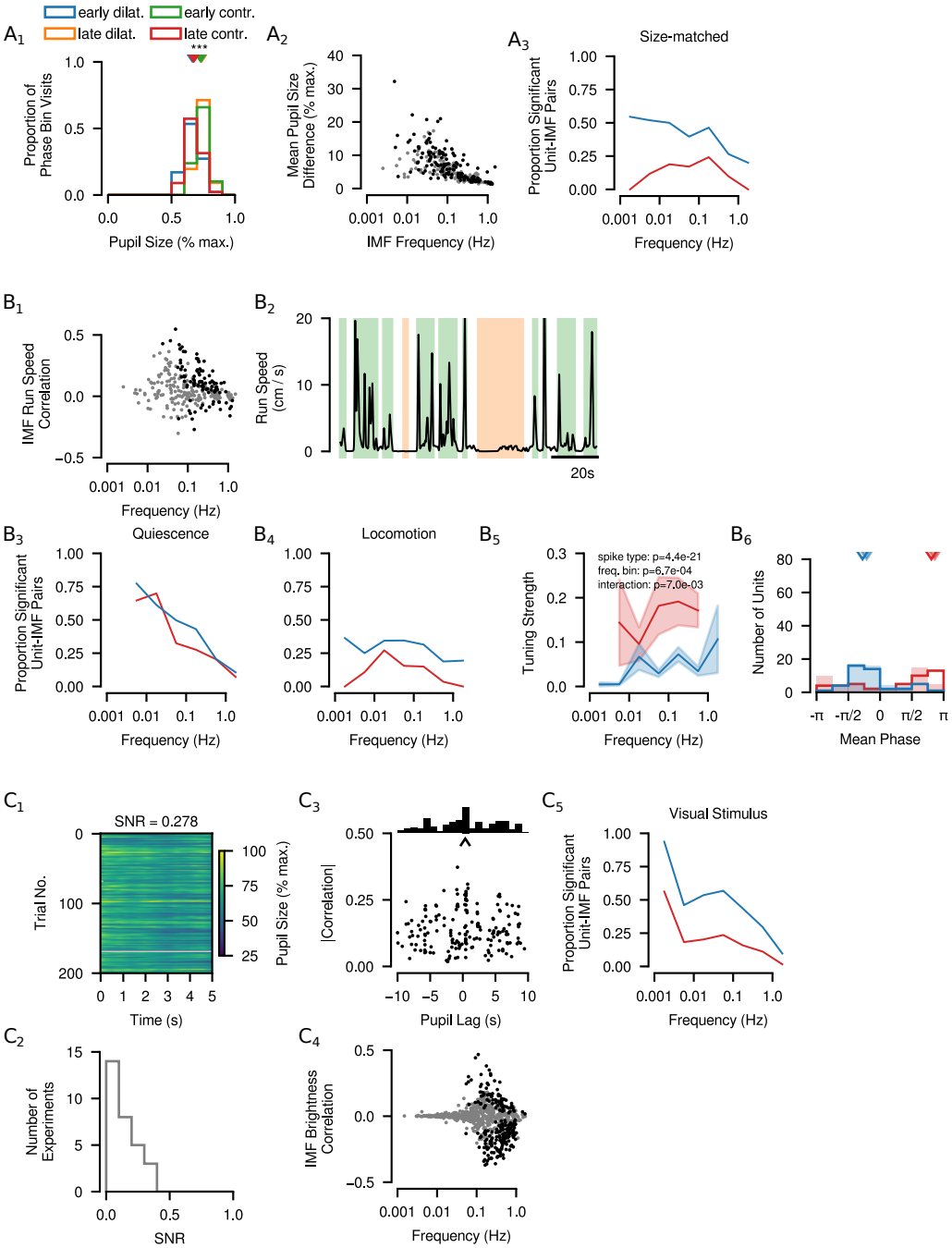


Figure S2 **(A_1)** Distributions of mean rate of spiking events in our sample of dLGN single units. All spikes belonging to a burst were treated as a single unitary event. *gray*: all spikes, median rate: 5.48 / s, ($N = 203$ units); *blue*: tonic spikes, median rate: 4.43 / s ($N = 203$). IRed: burst events, median rate: 0.12 / s ($N = 174$). Bursts were detected in 86 % (174/203) of units. **(A_2)** Distribution of burst ratio (burst pikes / all spikes), median 0.04 % ($N = 174$ units). **(B_1)** Distribution of phase bias index (PBI) across IMFs. Higher values indicate that IMFs spend an un-equal amount of time in each phase. Sampling phases only from periods of locomotion (*green*) or quiescence (*orange*) shifts the PBI distribution to higher values. **(B_2)** Phase bias index as a function of IMF frequency. B(C_1) Mean auto-correlation of burst events, $N = 174$ units. Dashed lines: ± 300 ms with notable short-term structure. **(C_2)** Same as C_1 but for tonic spikes.**(D_1)** Distribution of significance of phase tuning (p-values), determined from spike train permutation test. **(D_2)** Distribution of the number of IMFs to which each dLGN unit is significantly tuned. **(D_3)** Proportion of IMF-unit pairs in each frequency bin that have significant tuning. **(E_1, E_2)** Distribution of IMF frequency to which units are most strongly tuned, split between PS segments that are longer or shorter than 800 s, for bursts (E_1) and tonic spikes (E_2).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2021.04.30.442134; this version posted April 30, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S3 **($\mathbf{A}_1$)** Distribution of mean pupil sizes in each of four IMF phase bins for the example IMF shown in Figure 1D. The difference in mean pupil size between the early dilation (having the smallest mean pupil size, *blue*) and early contraction (having the largest mean pupil size, *green*) phase bins was significant (independent-samples t-test: $t = 7.86$, $p = 3.93 \times 10^{-13}$). **($\mathbf{A}_2$)** Mean differences between the phase bin with the largest pupil sizes and the phase bin with the smallest pupil sizes plotted against IMF frequency (*black dots*: significant difference between pupil size distributions for the two phase bins). **($\mathbf{A}_3$)** Proportion of IMF-unit pairs with significant tuning, after matching pupil size distributions across phase bins, as a function of frequency bin for bursts (*red*) and tonic spikes (*blue*). **($\mathbf{B}_1$)** IMF correlation with run speed plotted against IMF frequency (*black dots*: significant correlation). The significance of the IMF-run speed correlation was assessed by comparing the veridical correlation (value at a lag of 1.08 s, Figure S1B) with the distribution of values obtained by correlating the IMF with run speed traces from different experiments. **($\mathbf{B}_2$)** Run speed for the example segment show in Figure 1B. Highlighted regions represent periods defined as locomotion bouts (*green*) and quiescence (*orange*). **($\mathbf{B}_3$)** Same as A_3 with burst and tonic spike phase tuning assessed using only spikes occurring during periods of quiescence. **($\mathbf{B}_4$)** Same as A_3 with burst and tonic spike phase tuning assessed using only spikes occurring during periods of locomotion. **($\mathbf{B}_5$)** Mean tuning strength in each frequency bin computed using only spikes occurring during locomotion (*shaded areas*: 95% CI estimated as 1.96×SEM). **($\mathbf{B}_6$)** Distribution of tuning phases for bursts and tonic spikes for the IMF to which each unit has the strongest burst tuning. *Empty bars*: tuning strengths and tuning phases computed using only spikes that occurred during locomotion. *Filled bars*: tuning phase distribution for the same IMFs, but computed using spikes from all time periods. *Arrows*: circular means for each distribution (Watson-Williams test: bursts $F = 0.61$, $p = 0.26$ and tonic spikes $F = 0.35$, $p = 0.87$). **($\mathbf{C}_1$)** PS for trials in an experiment where the same 5 s naturalistic video clip was repeatedly presented, excluding trials on which optogenetic stimulation was present. **($\mathbf{C}_2$)** Distribution of signal to noise ratio (SNR) of pupil size for all experiments, in which naturalistic video clips were presented (mean $= 0.15 \pm 0.01$, $N = 30$). **($\mathbf{C}_3$)** Scatter plot of peak PS-stimulus brightness correlation value versus lag. Top: distribution of peak lags within the 20 s window, the mean lag across all cross-correlations was 0.38 s. **($\mathbf{C}_4$)** IMF correlation with run speed at the 0.38 s lag plotted against IMF frequency (*black dots*: significant correlation). The significance of the IMF-stimulus brightness correlation was assessed by comparing the veridical correlation at the 0.38 s lag with the distribution of values obtained by correlating the IMF with brightness traces compiled by randomly concatenating the brightness of the various video clips used in all experiments. **($\mathbf{C}_5$)** Same as A_3 with burst and tonic spike phase tuning assessed during stimulus viewing, including only spikes from trials without optogenetic stimulation.