

Supplemental #1: Calculation of LGN receptive fields and temporal scales

[Accompanying Butts et al. (2007)]

To draw a direct comparison between the relevant stimulus and response time scales, we must first consider how the visual pathway processes stimuli. A neuron's *receptive field* represents how visual stimuli are integrated spatially and temporally, and embodies the location and timing of visual stimuli to which the neuron responds (Reid et al., 1997). As we demonstrate below, filtering by the receptive field determines what aspects of the stimulus are represented by the neuron, and in doing so provides a temporal signal that can be directly compared to the time course of the neuronal response. In this way, "relative precision" is a comparison between the functional time scales of stimulus processing and those in the neuronal response.

Calculation of the neuron's spatiotemporal receptive field

Though historically the term *receptive field* only referred to spatial aspects of visual stimuli that a neuron responds to, the *spatiotemporal receptive field* (STRF) describes both spatial and temporal elements of the stimuli that evoke neuronal spikes (Reid et al., 1997; Chichilnisky, 2001). Because the main text of this paper focuses on the relationship between the temporal dynamics of the stimulus and response, here we provide examples of the temporal components of the receptive field (Figs. S1-1 and S1-2)

We calculate STRFs using standard techniques (Reid et al., 1997; Chichilnisky, 2001); for the data presented in the main text, we recorded spikes over five minutes during the presentation of binary white noise stimuli, consisting of frames of uncorrelated dark or light pixels updated at 60 Hz. In this condition, the STRF is simply the average stimulus that evokes a spike, and Fig. S1-1 demonstrates an example typical of an LGN X-cell. The spatial location of the non-zero portion of the receptive field is shown in Fig. S1-1A. [To avoid introducing noise into our model estimates using this receptive field, we assume the receptive field outside of a three-pixel radius from its center (given by the largest value of the spike-triggered average stimulus), demonstrated by the outer *blue circle* shown.]

Fig. S1-1B shows the temporal kernels at two spatial locations: in the center (*red*), and surround (*blue*). The STRF thus demonstrates the spatial location of stimuli that the neuron responds to (Fig. S1-1A) and their temporal characteristics (Fig. S1-1B). [For demonstration purposes, this example STRF was mapped using significantly more data, though doing such does not affect the results that are presented.]

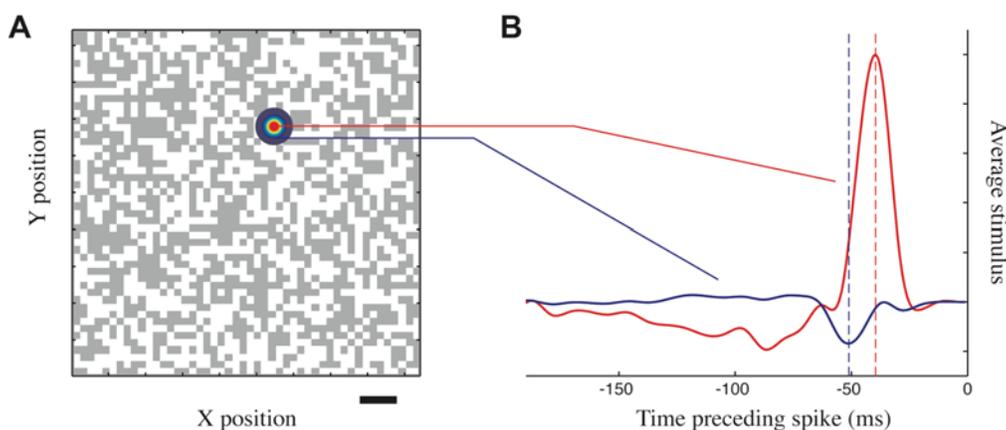


Fig. S1-1: The spatiotemporal receptive field of an LGN X-cell mapped with binary spatiotemporal noise. Red: excitatory, blue: suppressive. **A.** Spatial map of the average stimulus 40 ms preceding spikes, with an example frame of the noise stimulus superimposed. Scale bar is 1°. **B.** Temporal filter at the center of the receptive field (*red*) and 0.4° outside center (*blue*).

Note that we used binary spatiotemporal noise instead of natural movies to map the STRF because natural movies have complex spatiotemporal correlations and require more data to map in this way. Thus, for simplicity, we adjusted the contrast of the noise stimulus to evoke the same overall firing rate; the subtle differences previously described between receptive fields mapped with natural-movie and white-noise stimuli (Sharpee et al., 2006) are not expected to result in any substantial differences in the measures of precision used in this paper, and the RF predictions from this method (e.g., Fig. 1d) provide reasonable fits.

Calculation of the neuron's temporal receptive field in spatially uniform noise (SUN)

In the context of a spatially uniform noise (SUN) visual stimulus, the receptive field only describes how stimuli are processed temporally, and has no spatial component. The receptive field of the same neuron in the context of the SUN stimulus (Fig. S1-2, *solid line*) has a shorter latency and narrower temporal profile compared to the temporal filter of the neuron at the receptive field center in the spatiotemporal context (*dashed red line*). This is likely due to the fact that the suppressive surround is delayed relative to that of the center (Fig. S1-1B, *blue line*) (Enroth-Cugell et al., 1983; Sakai and Naka, 1995; Meister and Berry, 1999). Because spatially uniform stimuli activate them simultaneously, the opposite-signed surround suppresses the later part of the center response, leading to a faster and less prolonged temporal filter (Fig. S1-2, *black solid line*).

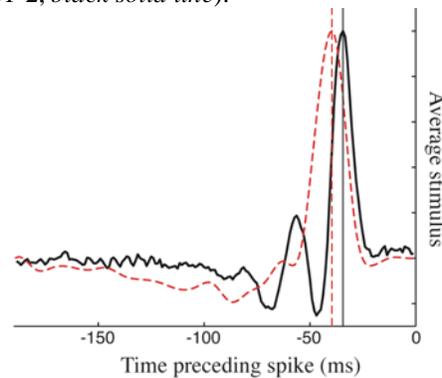


Fig. S1-2: The temporal receptive field of the same X-cell mapped with SUN stimuli (*black*), compared with the temporal filter at the center of the spatiotemporal receptive field (*dashed red*).

Generation of response predictions through a receptive-field-based model

Since the receptive field represents the average stimulus a neuron responds to, it can be used to predict when the neuron will respond: *i.e.*, when the stimulus matches the neuron's receptive field. This can be mathematically represented as the "filtered stimulus" $g(t)$, given by:

$$g(t) = \int d\tau k(\tau) s(t - \tau)$$

where $s(t)$ is the stimulus and $k(\tau)$ is the neuron's receptive field (pictured above with τ representing the time preceding the spike). Processing of the stimulus by the receptive field results in the filtered stimulus, and is shown for the "realistic model" neuron example below (Fig. S1-3).

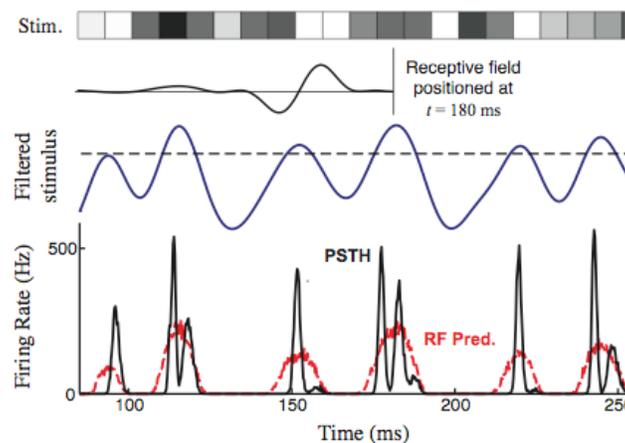


Fig. S1-3: Generation of the RF-based prediction. *Top:* The temporal receptive field matches the stimulus (an OFF to ON transition) at $t = 180$ ms, resulting in a large value of the filtered stimulus at that time. *Middle:* The time course of the filtered stimulus demonstrates how long a relevant stimulus is in the neuron's receptive field. The filtered stimulus can be mapped to a firing rate through a measured non-linearity (threshold: *dashed line*). *Bottom:* The resulting RF prediction (*dashed red line*) inherits the time course of the filtered stimulus, and cannot capture the finer temporal features of the neuron's response (*black line*).

The filtered stimulus can be directly mapped to a firing rate through a measured static non-linearity (Chichilnisky, 2001), resulting in the “RF prediction” of Figure S1-3 (*bottom*). These two stages of processing, the linear processing resulting in the filtered stimulus followed by a mapping to a firing rate by a non-linearity, comprise the canonical “LN (linear-non-linear) model” of visual neurons.

Precision as a comparison of time scales between the filtered stimulus and the response

Importantly, since $g(t)$ designates when stimuli are present that drive the neuron to fire, its time course provides a basis for evaluating whether the neuron’s response is precise relative to the stimuli it encodes. [Note that the RF prediction inherits this time course.] To quantitatively compare the filtered stimulus and response (and predicted response) time scales, we compare their autocorrelation functions, shown for an example LGN neuron in the context of SUN (Fig. S1-4).

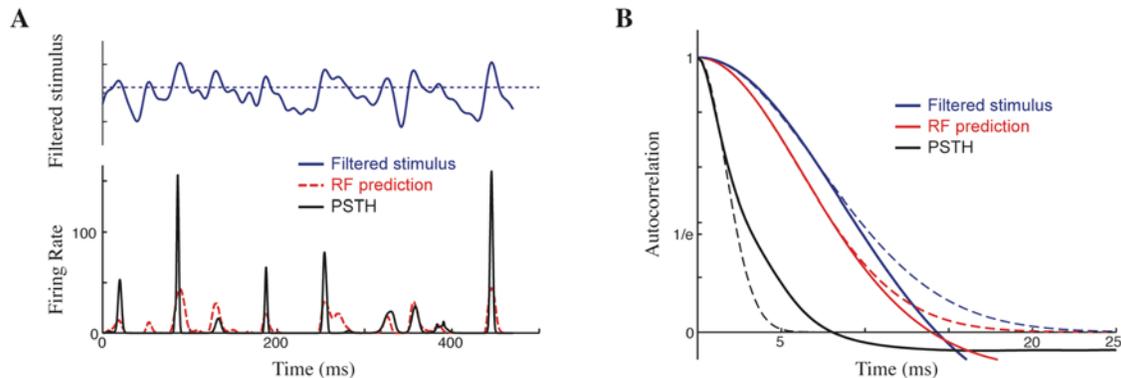


Fig. S1-4: Quantitative comparison of filtered stimulus and response time scales. **A.** A 500 ms time course of the filtered stimulus $g(t)$ (blue), PSTH $\lambda(t)$ (black), and RF prediction (red) for the LGN X-cell during full-field stimuli. **B.** The corresponding autocorrelation functions of each (solid lines), with the corresponding Gaussian fit to their top portion (dashed lines).

The autocorrelation functions clearly demonstrate the remarkable difference between filtered stimulus and PSTH time scales, since the correlation function of the PSTH falls off in less than a third of the time. [It also illustrates the similarity in the time scales of the filtered stimulus and RF prediction, demonstrating the negligible effect of rectification on this measure.]

To associate a single number with the time scale of the autocorrelation function, consider first a completely “precise” neuron. In this case, spikes will occur at the exact same time on every trial, and its PSTH will be comprised of a series of δ -functions. However, in general, trial-to-trial variability leads to temporal “jitter”, which might be approximated (to first order) in the PSTH as a convolution of the series of delta functions with a Gaussian function that has a standard deviation σ_j .

In the limit that each spike is spaced far apart relative to the width of the Gaussian jitter, the resulting autocorrelation function $C(\tau)$ can be computed analytically, and is proportional to:

$$C(\tau) \propto \int dt \lambda(t) \lambda(t - \tau) \approx \int dt e^{-t^2/2\sigma_j^2} e^{-(t-\tau)^2/2\sigma_j^2} \\ \propto e^{-\tau^2/4\sigma_j^2}$$

This itself is a Gaussian with standard deviation $\sigma_j \sqrt{2}$.

In this way, from the autocorrelation function we can infer the average jitter σ_j of the PSTH (note that as we define it, the “time scale” of the response τ_R is twice the jitter, equal to the full width of the PSTH event): $\tau_R = 2\sigma_j$ is the autocorrelation lag at which the Gaussian has a height of $1/e$. For spikes that are closer together, the autocorrelation function will also reflect correlations between spikes, but the shape near the origin (at small correlation lag times) will still be dominated by the isolated spike events, and is thus well-described by a Gaussian.

We therefore fit the top portion of the autocorrelation function with a Gaussian (Fig. S1-4B, dashed lines), and from this fit estimate the response time scale. This provides a direct means for comparing the time scale of the continuous filtered stimulus with that of the neuronal response, and relating these measures to more conventional measures of neural precision such as jitter.

ADDITIONAL METHODS

Experimental Procedures. Please note that the surgical and experimental preparations used for this study have been described in detail previously (Weng et al., 2004).

The visual stimuli. The LGN neurons were studied under both spatially uniform noise (SUN) and natural movie conditions. SUN stimuli consisted of spatially uniform luminances randomly selected from a Gaussian distribution with zero mean (corresponding to the midway point of the full range of monitor luminance) and an RMS contrast of 0.55, presented at 120 Hz.

The naturalistic movie sequence was recorded by members of the laboratory of Peter König (Institute of Neuroinformatics, ETH/UNI Zürich, Switzerland) using a removable lightweight CCD-camera mounted to the head of a freely roaming cat in natural environments such as grassland and forest (Kayser et al., 2003). A 48×48 windowed area of this movie was temporally upsampled from 30 Hz to 60 Hz using commercial software (MotionPerfect, Dynapel Systems, New York) and normalized to have a constant mean and standard deviation of luminance for each frame (contrast held at 0.4 of maximum), as described in detail in (Lesica et al., 2006). For straightforward receptive field mapping using spike-triggered average techniques, a five-minute spatiotemporal binary noise sequence was also presented to each cell with the same spatial resolution and 0.55 contrast (see above). All stimuli were displayed on a CRT display at a resolution of 0.2° per pixel with a monitor refresh rate of 120 Hz.

Response Events and Smoothing of the PSTH. The peri-stimulus time histogram (PSTH) is used to estimate response time scales and single spike information, and when measured from recorded data (main text, Figs. 1a,b) was estimated from 60 or 120 repeated trials. The PSTH in the context of both SUN and natural movie stimuli was easily divided into individual *response events*, where each event boundary was determined by more than 5 ms of zero firing rate in the PSTH.

Though the number of repeated stimulus trials was enough to distinguish PSTH features of the response by eye, calculations of the response time scales (described above) required smoothing at the finer time resolutions. This was performed using Bayesian adaptive regression splines (BARS) (Kass et al., 2003), which captures both fast changes in firing rate as well as smooth regions of the PSTH that have statistically insignificant fluctuations. These methods were compared to several other types of smoothing (including a method invented by the authors), and while it produced a better estimate of the PSTH, did not significantly change the results presented.

REFERENCES

- Chichilnisky EJ (2001) A simple white noise analysis of neuronal light responses. *Network* 12:199-213.
- Enroth-Cugell C, Robson JG, Schweitzer-Tong DE, Watson AB (1983) Spatio-temporal interactions in cat retinal ganglion cells showing linear spatial summation. *J Physiol* 341:279-307.
- Kass RE, Ventura V, Cai C (2003) Statistical smoothing of neuronal data. *Network* 14:5-15.
- Kayser C, Salazar RF, König P (2003) Responses to natural scenes in cat V1. *J Neurophysiol* 90:1910-1920.
- Lesica NA, Weng C, Jin J, Yeh CI, Alonso JM, Stanley GB (2006) Dynamic Encoding of Natural Luminance Sequences by LGN Bursts. *PLoS Biol* 4:e209.
- Meister M, Berry MJ, 2nd (1999) The neural code of the retina. *Neuron* 22:435-450.
- Reid RC, Victor JD, Shapley RM (1997) The use of m-sequences in the analysis of visual neurons: linear receptive field properties. *Vis Neurosci* 14:1015-1027.
- Sakai HM, Naka K (1995) Response dynamics and receptive-field organization of catfish ganglion cells. *J Gen Physiol* 105:795-814.
- Sharpee TO, Sugihara H, Kurgansky AV, Rebrik SP, Stryker MP, Miller KD (2006) Adaptive filtering enhances information transmission in visual cortex. *Nature* 439:936-942.
- Weng C, Yeh CI, Stoelzel CR, Alonso JM (2004) Receptive field size and response latency are correlated within the cat visual thalamus. *J Neurophysiol* 93:3537-3547.

Supplemental #2: The change in temporal scales of the response with stimulus class [Accompanying Butts et al. (2007)]

Temporal scales of the neural response

Our manuscript demonstrates that relevant time scales of the neuronal response are not fixed, but depend on the temporal structure of the input. Figure S2-1 demonstrates a significant change in several measures of the temporal scale of the neural response: the first-spike jitter (Fig. S2-1A), the total jitter (Fig. S2-1B), and the *response time scale* (Fig. S2-1C), focused on in the main text. Comparing these measures here demonstrates that the observed difference in response time scales from spatially uniform noise (SUN) to natural movie stimuli is a general feature not limited to any one measure of temporal scale.

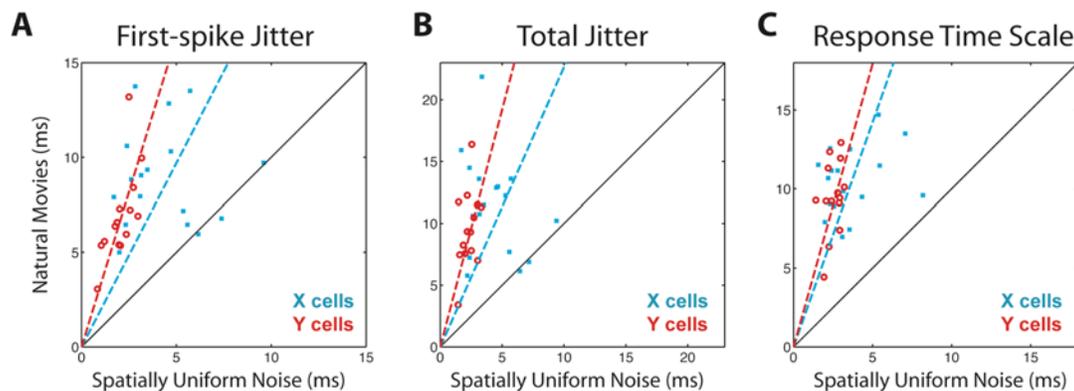


Figure S2-1: Multiple response time scales change dramatically with different types of dynamic stimuli.

A. The change in average first-spike jitter of each neuron in the context of SUN stimuli (*horizontal axis*) and natural movies (*vertical axis*). The first-spike jitter is defined as the standard deviation of the first-spike of a given event across stimulus repeats; the average weights each event by the number of trials where at least one spike occurred in that event. The dashed line represents the average change, which is 1.9-fold for X-cells and 3.3-fold for Y-cells. **B.** The change in total jitter for each cell in the two stimulus types, defined as the standard deviation of all spikes in the event across stimulus repeats, where the average now weights each event by the total number of spikes in the event. Here, the average change (*dashed line*) is 2.3-fold (X-cells) and 3.8-fold (Y-cells). (A) and (B) depend on dividing up the neural response into “events”, which are easily delineated by a contiguous period (> 5 ms) of zero spike-rate in the PSTH. **C.** The change in response time scale across neurons (average 2.8-fold and 3.6-fold for X- and Y-cells, respectively), derived from the autocorrelation time of the response PSTH, as described at length in Supplemental #1.

Event-by-event measures of temporal scale

The dependence of response time scale on the time scales of the stimulus can also be observed on an event-by-event basis in the context of a single stimulus type. For example, the SUN stimulus considered in the main text is constructed from sequences of randomly drawn luminance values, and as a result contains ‘stimulus features’ at many different time scales over the course of a single trial. The neuron responds when the stimulus matches the receptive field (*i.e.*, where the filtered stimulus is above threshold, Fig. S2-2A), and as a result, the properties of these above-threshold excursions vary from event to event over the course of the experiment, and will co-vary with the relevant properties of the associated neuronal response.

We first look at the relationship between the duration above threshold of the filtered stimulus and the duration of the neuronal response (time between first spike in the event and last spike). There were 139 separate response events generated by a typical LGN neuron over a 10 second SUN stimulus presentation (Fig. S2-2B). The duration of the response is less than the time above threshold of the filtered stimulus, and the average ratio between them is 3.0 (*blue dashed line*), similar to the measurement of *relative precision* between stimulus and response time scales in the main text (though there is significant variability from event to event).

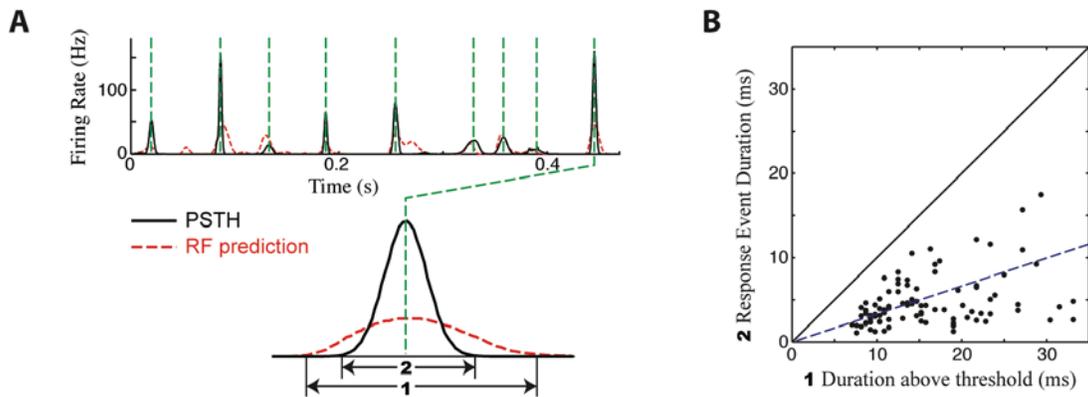


Figure S2-2: Relationship between stimulus and response time scales preserved on an event-by-event level. **A.** The filtered stimulus above threshold (*i.e.*, RF prediction, see Suppl. #1) and PSTH of a typical LGN neuron are shown, demonstrating how above-threshold features of the filtered stimulus (FS, *dashed red*) can be associated with response events (*solid black*), to measure the relationship between properties of each. **B.** In particular, over the 139 events of a 10 second SUN stimulus, the duration above threshold of the FS (*horizontal axis*) is on average 3.0 times longer than the duration of the associated response (*vertical axis*): this ratio is shown as a *blue dashed line*.

The correlation coefficient (CC) is a simple measure of the relationship between filtered stimulus (FS) feature and response event durations. The $CC = 0.36$ between FS and response duration for the neuron considered in this example. Likewise, the correlation between other measures of stimulus time scale is of similar magnitude in this example: $CC = 0.30$ between the FS duration and total response jitter, and $CC = 0.30$ between the FS duration and first-spike jitter.

This can be compared to other aspects of the FS features that may be related to response time scales (Fig. S2-3A). In particular, the amplitude of the filtered stimulus feature is related to how well a given stimulus matches the neuron's receptive field, and the slope at threshold is related to how quickly such a feature is integrated. However, these other properties, though correlated with FS duration, are not nearly as predictive of the temporal properties of the response. Figure S2-3B shows the average correlation coefficient between each of these other parameters and temporal features of the response.

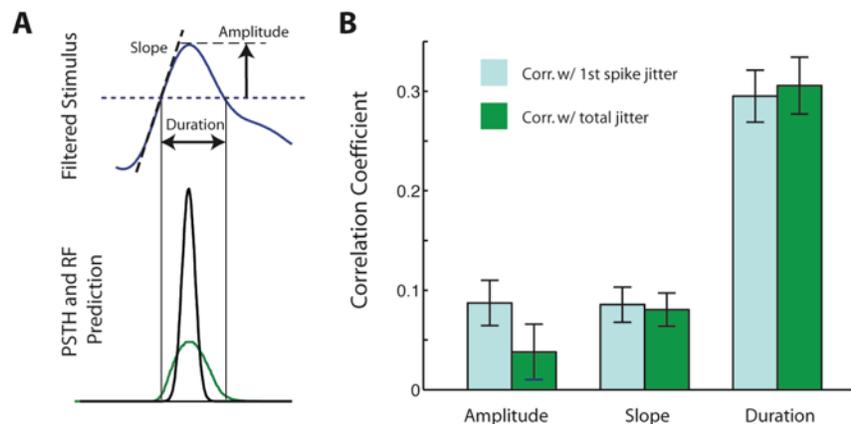


Figure S2-3: Relationship between response time scales and other filtered stimulus properties. **A.** An illustration of other properties of above-threshold features of the filtered stimulus that may be related to response precision. **B.** The average correlation coefficient across LGN neurons between properties of the FS and measures of response time scale ($n = 20$).

Such a result is in contrast to an explanation for precision in the context of an *in vitro* cortical neuron responding to a dynamically varying injected current (Mainen and Sejnowski, 1995). In this condition, the slope at threshold crossing was the largest determinant of neural precision (see Suppl. #3). Clearly, such results do not map to the more abstract functional relationship between the presented stimulus filtered by a receptive field and a direct injected current. [Furthermore, they found neural precision only changed by roughly a factor of two despite a 25-fold change in the time scale of the injected current, as discussed in Suppl. #3.] Rather, precision in neurons in the early visual system likely arises from the suppression of the response following its onset (*e.g.*, Berry and Meister, 1998), and is the subject of much current modeling (see Suppl. #3).

REFERENCES

- Berry MJ, 2nd, Meister M (1998) Refractoriness and neural precision. *J Neurosci* 18:2200-2211.
- Mainen ZF, Sejnowski TJ (1995) Reliability of spike timing in neocortical neurons. *Science* 268:1503-1506.

Supplemental #3: Exploring the source of temporal precision in common phenomenological models

[Accompanying Butts et al., 2007]

The biophysical properties of LGN neurons, as well as the upstream retinal network responsible for their responses to visual stimuli, are quite complex and not entirely understood. Nevertheless, there is a wide range of “phenomenological” visual neuron models designed to capture certain response properties to certain classes of visual stimuli. The most ubiquitous visual system model is the receptive-field-based (or Linear-Non-Linear “LN”) model described in the main text, which characterizes the complex spatiotemporal processing of the retina with a single linear filter, and can mostly capture the overall magnitude and coarse timing of LGN responses, despite not being able to reproduce their fine temporal structure (main text, Fig. 1). More complicated models have been built with this receptive-field-based framework as a foundation to capture higher-order features of the neuronal response: some approximate the contribution of different retinal cell classes in order to understand spatiotemporal stimulus interactions (Korenberg et al., 1989; Sakai and Naka, 1995; Snippe et al., 2000), others focus on capturing how the response changes with different luminance and contrast levels (Shapley and Victor, 1979; Chander and Chichilnisky, 2001; Bonin et al., 2005; Mante et al., 2005), and others approximate aspects of the spike generation process more accurately (Brillinger, 1992; Mainen and Sejnowski, 1995; Berry and Meister, 1998; Keat et al., 2001; Paninski, 2004; Paninski et al., 2004). Each model is able to capture certain aspects of the neural response to certain types of stimuli, but the ability to fit the responses in complex stimulus environments such as natural scenes is still elusive and likely involves a combination of all these different aspects (Meister and Berry, 1999; Carandini et al., 2005).

Threshold-crossing models of precision

Several types of phenomenological models address how temporal aspects of the neuronal response are generated, directly or indirectly. In particular, the spike generation mechanism itself is commonly thought to underlie the temporal precision of neuronal responses (Mainen and Sejnowski, 1995), and is the basis of several “threshold-crossing” models of visual neurons (Keat et al., 2001; Paninski et al., 2004). Such a model generates precisely-timed spikes when an internal variable (often thought of as the membrane potential) crosses a threshold (Fig. S3-1A). Intrinsic noise can change the precise timing of the threshold crossing, resulting in temporal jitter. The magnitude of jitter (and resulting time scale of the neural response) therefore depends on the magnitude of the slope of the potential relative to the magnitude of the noise: a lower slope translates into more temporal jitter for the same noise level (Fig. S3-1A). Such a model could potentially predict the existence of relative precision, since stretching the time scale of the stimulus (while holding the amplitude constant) will proportionally decrease the slope of the filtered stimulus, and increase the temporal jitter in the neuronal response.

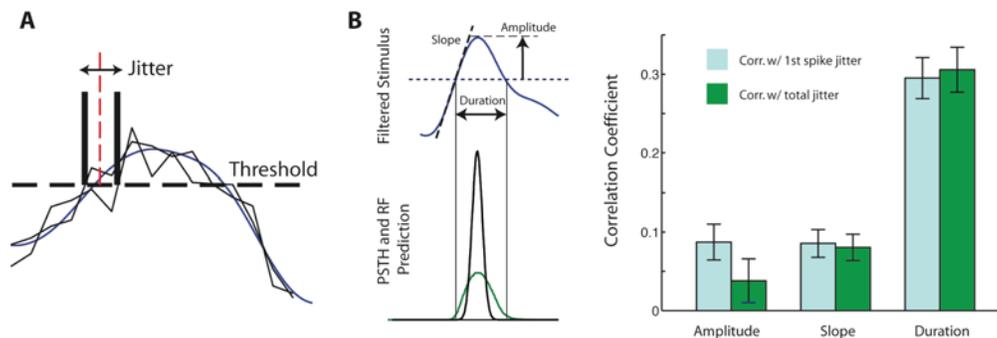


Fig S3-1: Threshold-crossing model of precision. **A.** Illustration of how temporal jitter is generated in a threshold-crossing model. **B.** Correlation between various elements of the filtered stimulus (amplitude, slope, duration) and the response time scales demonstrates that slope at threshold is not a good predictor of the temporal jitter in LGN data relative to filtered stimulus duration, reproducing Supplemental Figure S2-3.

However, as demonstrated in Supplemental #2 (reproduced as Fig. S3-1B), such a relationship between threshold-crossing-slope and response duration is not supported by our observations. On an event-by-event basis, the slope of the filtered stimulus at threshold does not predict either the level of precision of the first spike, nor the temporal structure in the response overall. Furthermore, while Mainen and Sejnowski (1995) found a strong relationship between the slope of the current injected directly into cortical neurons *in vitro* and their precision, even in that study the relationship between slope and precision does not hold up quantitatively across different stimulus time scales: a 25-fold change in the time scale of the injected current (which would result in an 25-fold decrease in the average slope at threshold), only results in at most a 2-fold change in response time scales.

The lack of applicability of the threshold-crossing model for precision to the LGN is also supported by another experimental manipulation that we performed. We presented the natural movie stimulus at two different contrasts: “high contrast” (0.40) described in the main body of the paper, and low contrast (0.15). While these experiments are part of a future study on the effects of contrast on coding in the LGN, we bring up these data in this context because it is an experimental manipulation that also affects the ratio of the slope of the signal to that of the putative “noise” in the system: while the noise is thought to remain roughly constant as a function of contrast (Passaglia and Troy, 2004), the magnitude (and slope) of the filtered stimulus will be reduced 2.7-fold. Thus, we would expect to see a 2.7-fold decrease in precision if the threshold-crossing model of precision were to apply. However, our preliminary results indicate that the level of relative precision of the LGN neuron is approximately maintained in the natural movies for the 2.7-fold decrease in contrast (not shown). Though the response time scale does increase, the time scale of the filtered stimulus also increases, as the receptive field becomes temporally broader (Shapley and Victor, 1978). Of course, our measure of precision breaks down for low enough contrast, as spurious spikes increasingly dominate the spike train. The dependence of the precision on contrast is part of this future study.

Refractory and “suppressive” models of precision

Another phenomenological explanation for precision is neural “refractoriness” (Berry and Meister, 1998), corresponding to a decreased probability of spiking immediately following a neural spike. Neural refractoriness could explain the observed precision if spiking were to suppress subsequent spiking for the duration of the filtered stimulus event (approximately 15 ms in the case of SUN, Fig. S3-2A). Such suppression is significantly longer than the “absolute refractory period” linked to the biophysical properties of action potential generation, and as a result has been referred to as a “relative refractory period” (Berry and Meister, 1998; Reinagel and Reid, 2000), “recovery function” (Brillinger, 1992), or “spike-history term” (Paninski, 2004) and is the subject of current modeling efforts, including those by the authors of this manuscript (Butts et al., 2007).

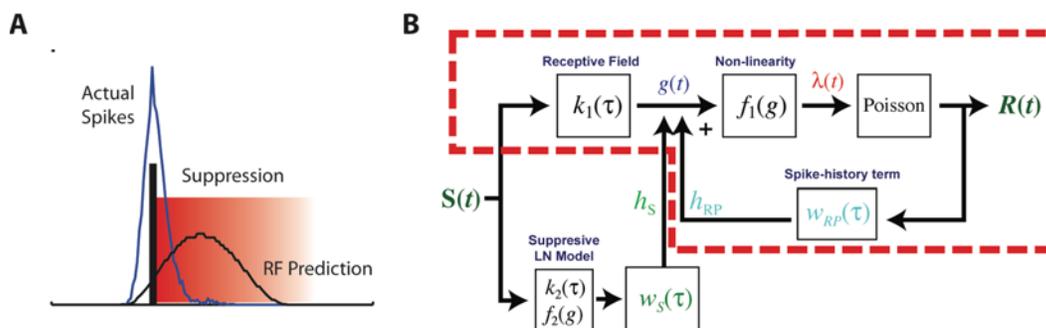


Fig S3-2: Refractory-based models of precision. **A.** Illustration of how refractoriness leads to neural precision. **B.** Proposed “phenomenological” model to explain precision in multiple stimulus contexts. The dashed red box demonstrates a simpler model with a spike-history term that can explain some elements of neural precision in the context of SUN (Paninski, 2004), but cannot quantitatively reproduce precision in natural movies.

Unfortunately, similar to most previous studies of neural precision, modern models of RGCs and LGN neurons have only been validated for a single class of stimuli – the spatially uniform noise (SUN)

considered in the main text – and none claim to be valid for other classes. For example, the “realistic neuron model” considered in the main text was specifically designed to reproduce the trial-to-trial variability of RGC and LGN neurons in the context of SUN (Keat et al., 2001), but has not been extended to capture more complicated interactions needed to predict responses to spatiotemporal (not just temporal) stimuli (Korenberg et al., 1989; Sakai and Naka, 1995; Meister and Berry, 1999; Bonin et al., 2005). Likewise, these models that are designed to capture these more complicated interactions have not been extended to spiking models and likewise cannot incorporate spike-dependent effects.

In preliminary (unpublished) analyses of hybrid models that may approximate crucial elements of retinal processing and spike generation (Butts et al., 2007), we have found that the “relative refractory” explanation for precision in the context of SUN breaks down in spatiotemporal stimuli where spiking is driven by more complicated spatiotemporal interactions. We find that it is necessary to include feed-forward suppression – not tied to neuronal spikes – in order to explain precision in the context of spatiotemporal stimuli (see Fig. S3-3B). Note that such suppression can result in an apparent “relative refractory period” in the context of SUN stimuli (but not natural movies), highlighting the importance of studying multiple stimulus contexts in describing neural precision.

In this way, the presence of relative precision that we report in the main text offers a significant experimental constraint for current state-of-the-art models, which have typically only been tested in isolated stimulus contexts.

Implications for different types of stimulus

The mechanism underlying precision will likely have implications for other aspects of stimulus processing. Our modeling work suggests that precision arises from suppressive mechanisms in the retina that have particular spatiotemporal tuning to the stimulus as well (with spatial tuning similar to the neuron’s surround) – consistent with a recent study of the spatial components of contrast gain control (Bonin et al., 2005). If this were the case, one would expect that the degree (and even presence) of relative precision could be modulated by stimuli that differentially stimulate the excitatory versus suppressive pathways. The two stimulus classes used in this study (SUN and natural movies) both involved coherent stimulation of these surround mechanisms. However, the presence of the relative precision that we report is likely conditional on such stimulation, highlighting the importance of studying aspects of neural coding in contexts that are relevant to natural processing.

REFERENCES

- Berry MJ, 2nd, Meister M (1998) Refractoriness and neural precision. *J Neurosci* 18:2200-2211.
- Bonin V, Mante V, Carandini M (2005) The suppressive field of neurons in lateral geniculate nucleus. *J Neurosci* 25:10844-10856.
- Brillinger DR (1992) Nerve cell spike train data analysis: a progression of technique. *J American Statistical Assoc* 87:260-271.
- Butts DA, Weng C, Jin J, Yeh CI, Lesica NA, Stanley GB, Alonso JM, Paninski L (2007) Direct measurement of "suppression" in the LGN in the context of natural stimuli and its implications for visual coding. *Cosyne*.
- Carandini M, Demb JB, Mante V, Tolhurst DJ, Dan Y, Olshausen BA, Gallant JL, Rust NC (2005) Do we know what the early visual system does? *J Neurosci* 25:10577-10597.
- Chander D, Chichilnisky EJ (2001) Adaptation to temporal contrast in primate and salamander retina. *J Neurosci* 21:9904-9916.
- Keat J, Reinagel P, Reid RC, Meister M (2001) Predicting every spike: a model for the responses of visual neurons. *Neuron* 30:803-817.
- Korenberg MJ, Sakai HM, Naka K (1989) Dissection of the neuron network in the catfish inner retina. III. Interpretation of spike kernels. *J Neurophysiol* 61:1110-1120.
- Mainen ZF, Sejnowski TJ (1995) Reliability of spike timing in neocortical neurons. *Science* 268:1503-1506.

- Mante V, Frazor RA, Bonin V, Geisler WS, Carandini M (2005) Independence of luminance and contrast in natural scenes and in the early visual system. *Nat Neurosci* 8:1690-1697.
- Meister M, Berry MJ, 2nd (1999) The neural code of the retina. *Neuron* 22:435-450.
- Paninski L (2004) Maximum likelihood estimation of cascade point-process neural encoding models. *Network* 15:243-262.
- Paninski L, Pillow JW, Simoncelli EP (2004) Maximum Likelihood Estimation of a Stochastic Integrate-and-Fire Neural Encoding Model. *Neural Comput* 16:2533-2561.
- Passaglia CL, Troy JB (2004) Impact of noise on retinal coding of visual signals. *J Neurophysiol* 92:1023-1033.
- Reinagel P, Reid RC (2000) Temporal coding of visual information in the thalamus. *J Neurosci* 20:5392-5400.
- Sakai HM, Naka K (1995) Response dynamics and receptive-field organization of catfish ganglion cells. *J Gen Physiol* 105:795-814.
- Shapley R, Victor JD (1979) The contrast gain control of the cat retina. *Vision Res* 19:431-434.
- Shapley RM, Victor JD (1978) The effect of contrast on the transfer properties of cat retinal ganglion cells. *J Physiol* 285:275-298.
- Snippe HP, Poot L, van Hateren JH (2000) A temporal model for early vision that explains detection thresholds for light pulses on flickering backgrounds. *Vis Neurosci* 17:449-462.

Supplemental #4: The relationship between jitter and the frequency content of the stimulus

[Accompanying Butts et al. (2007)]

In the main text of this paper, we show that the temporal precision of LGN responses can be understood in terms of the frequency content of the stimulus being represented. This supplemental describes how the frequency content of the stimulus that is represented by the neuron is influenced by (1) the temporal filtering of the neuron's receptive field and (2) the trial-to-trial variability in the spike timing ("jitter").

The temporal filtering of the receptive field

The temporal components of a neuron's receptive field shape what frequencies in the stimulus can be represented in the neuron's response. In the frequency domain, the filtered stimulus $G(f)$ is simply a multiplication of the stimulus $S(f)$ with the Fourier transform of the STRF $K(f)$ (also referred to as the transfer function): $G(f) = K(f) S(f)$. Since we are interested in the power in the filtered stimulus $|G(f)|^2 = |K(f)|^2 |S(f)|^2$, below we show the square magnitude of these three functions for both the temporal (Fig. S1-2) and spatiotemporal (Fig. S1-1) receptive fields shown in Supplemental #1.

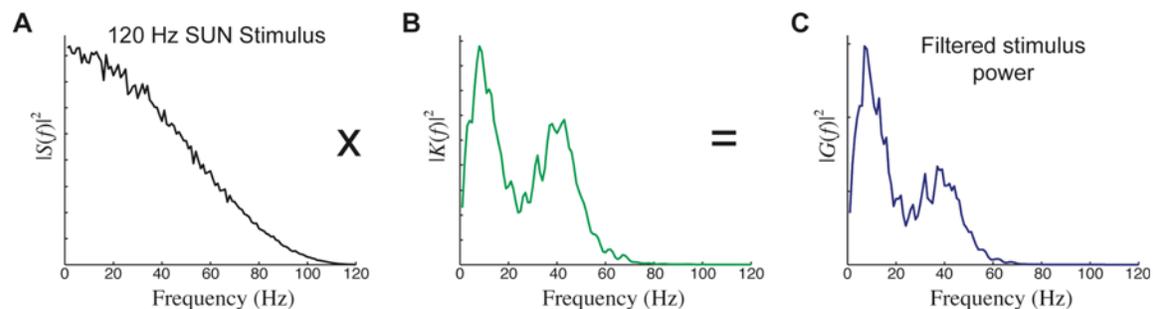


Fig. S4-1: The effect of filtering by the temporal receptive field. **A.** Power in the SUN stimulus described in the main text. **B.** Square-magnitude of the transfer function of the filter shown in Fig. S1-2. **C.** Resulting power in the filtered stimulus. Note that the filtered stimulus power is largely shaped by the transfer function.

The power spectrum of the natural movie stimulus that we present is typical for natural vision, with most power at low frequencies that gradually falls off with increasing frequency (Fig. S4-2A). Given this broad shape, the band-pass properties of the receptive field itself (Fig. S4-2B) have the largest effect on the distribution of filtered stimulus power (Fig. S4-2C), since they attenuate both low (< 5 Hz) and high frequencies (> 20 Hz). Though our movies do not capture natural eye movements, their effects would have to be very significant to alter this basic result. In fact, fixational eye movements have been modelled as random drift that generates a power spectrum proportional to $1/f^2$ up to 40 Hz (Rucci et al., 2000), which arguably would ensure a filtered stimulus power spectrum like that shown below, even when viewing a static scene.

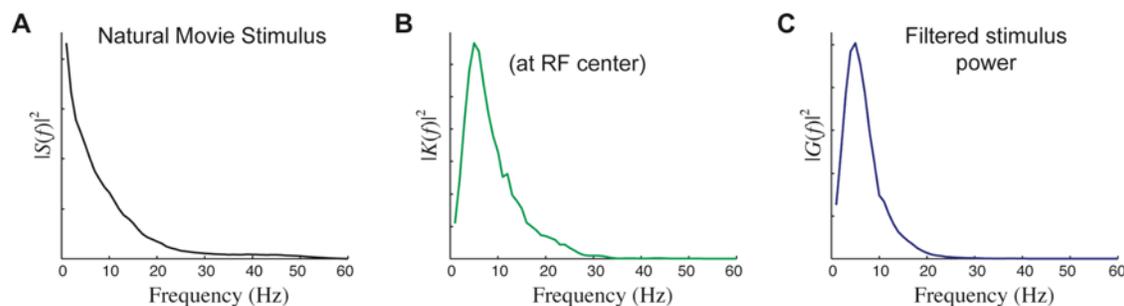


Fig. S4-2: The effect of filtering by the spatiotemporal receptive field. **A.** The power in the natural movie averaged across pixels. **B.** Square-magnitude of the transfer function at the center pixel of the STRF from Fig. S1-1. **C.** Resulting power in the filtered stimulus.

The effect of spike-timing variability (jitter) on the representation of the stimulus

In the main text, we demonstrate the effects of jitter on the representation of the stimulus that can be reconstructed from the neuron's spike train (Figs. 2 and 3). Jitter considered in this context is essentially noise, and as a result it is no surprise that it impacts the ability to reconstruct the stimulus from the response. However, as we demonstrate in the main text (Fig. 2e), degradation of the reconstructed stimulus only occurs for sufficiently large jitter. As we demonstrate below, the level of precision at which this occurs can be estimated given the effect of jitter on the coherence between the stimulus $s(t)$ and the reconstructed stimulus $s^*(t)$.

The coherence $\gamma^2(\omega)$ is given by:

$$\gamma^2(\omega) = \frac{\langle |S(\omega) \overline{S^*(\omega)}|^2 \rangle}{\langle |S(\omega)|^2 \rangle \langle |S^*(\omega)|^2 \rangle}$$

where $S(\omega)$ and $S^*(\omega)$ are the Fourier transforms of $s(t)$ and $s^*(t)$, respectively, and the bar denotes complex conjugate. [Note that we are using $\omega = 2\pi f$ instead of f to be consistent with common conventions.] Consider a temporally precise neuronal response $r_0(t)$, such that the PSTH is a series of δ -functions. For simplicity, we consider the case where the jitter added to each spike follows a Gaussian distribution with a "jitter" σ_j , given by $\mathcal{N}(\tau; \sigma_j) = A \exp(-\tau^2/2\sigma_j^2)$.

The reconstructed stimulus is given by the convolution of the optimal reconstruction filter $h(\tau)$ (see Methods below) with the response $r(t)$:

$$s^*(t) = \int d\tau h(\tau) r(t - \tau)$$

$$S^*(\omega) = H(\omega) R(\omega)$$

Thus, the coherence becomes:

$$\gamma^2(\omega) = \frac{\langle |S(\omega) \overline{R(\omega)}|^2 \rangle}{\langle |S(\omega)|^2 \rangle \langle |R(\omega)|^2 \rangle}$$

Note that even for a high-firing rate neuron for the LGN (20 Hz), the average inter-spike interval is 50 ms. The effect of the jitter on the power of the response $\langle |R(\omega)|^2 \rangle$ will be negligible, such that $\langle |R(\omega)|^2 \rangle \approx \langle |R_0(\omega)|^2 \rangle$. However, the jitter will affect the correlation between the stimulus and response, such that:

$$\langle |S(\omega) \overline{R(\omega)}|^2 \rangle = \langle |S(\omega) \overline{R_0(\omega)} e^{-\omega^2 \sigma_j^2 / 2}|^2 \rangle$$

so

$$\gamma^2(\omega) = \gamma_0^2(\omega) e^{-\omega^2 \sigma_j^2}$$

where $\gamma_0^2(f)$ is the coherence for the case of the precise spike train. Thus, in this simple example, the ratio of coherences (main text, Fig. 3d) falls off as a Gaussian with standard deviation given by the *attenuation frequency* $f_A = 1/(\sigma_j 2\pi\sqrt{2})$.

The attenuation frequency of the coherence varies continuously with jitter, but its effects on the quality of the stimulus reconstruction depend on the amount of stimulus power as a function of frequency. As jitter is increased, the attenuation frequency decreases from infinity, and the reconstruction is ultimately affected when f_A overlaps with frequencies represented in the stimulus.

Thus, the role of precision can be understood through a comparison between the frequencies represented by the filtered stimulus (the first section of this supplemental) and those attenuated by the jitter (this section). The interplay selects for a level of relative precision $\sim 3\times$, and is reflected in LGN responses recorded in the context of both SUN and natural movie stimuli (main text, Fig. 3).

ADDITIONAL METHODS

Linear Reconstruction. The optimal linear reconstruction filter corresponds to the best choice of $h(\tau)$ that minimizes the mean-squared difference between the stimulus $s(t)$ and reconstructed stimulus $s^*(t)$. It satisfies the following equation (Bialek et al., 1991; Dan et al., 1998; Stanley et al., 1999):

$$\int d\tau' h(\tau') \int dt r(t-\tau) r(t-\tau') = \int dt s(t) r(t-\tau)$$

where $r(t)$ is a sum of delta functions representative of the neuronal spike train. The reconstruction filters were estimated numerically for a given stimulus and set of experimentally observed spike times.

The reconstruction information I_{REC} is calculated from the coherence spectra $\gamma^2(f)$ between stimulus $s(t)$ and reconstructed stimulus $s^*(t)$: $I_{REC} = \int df \log_2[1-\gamma^2(f)]$. In the case of the natural movie reconstruction (main text, Fig. 3d), only the reconstruction information of the center pixel of the receptive field is reported.

REFERENCES

- Bialek W, Rieke F, de Ruyter van Steveninck RR, Warland D (1991) Reading a neural code. *Science* 252:1854-1857.
- Dan Y, Alonso JM, Usrey WM, Reid RC (1998) Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nat Neurosci* 1:501-507.
- Rucci M, Edelman GM, Wray J (2000) Modeling LGN responses during free-viewing: a possible role of microscopic eye movements in the refinement of cortical orientation selectivity. *J Neurosci* 20:4708-4720.
- Stanley GB, Li FF, Dan Y (1999) Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus. *J Neurosci* 19:8036-8042.

Supplemental #5: Phase-locking and the temporal scales of the neuronal response

[Accompanying Butts *et al.*, 2007]

One possible explanation for the temporal precision of LGN neuron responses is phase locking of the response to the CRT refresh rate (Wollman and Palmer, 1995; Williams *et al.*, 2004). While most LGN neurons in our study displayed some degree of phase-locking to the monitor refresh, we demonstrate in here that (1) it plays only a minor role in affecting the time scales of the LGN response to spatially uniform noise (SUN) stimuli and plays no discernable role in the time scale of the response to natural scenes, and (2) that the effects that we describe in the main text do not result from such phase locking.

To examine the degree of phase-locking between the neuronal response and the monitor refresh, we use the methods described in Williams *et al.* (2004). For each neuron in each stimulus condition, we measured a *power ratio* (PR) from the power spectrum of the neuron's spike response, which is given by the ratio of the power at the monitor refresh rate (120 Hz) to the average power at surrounding frequencies (± 10 Hz). Fig. S5-1 shows the power spectrum of an example LGN neuron with significant phase locking (*left*) and that of a different neuron recorded simultaneously that had no phase locking (*right*).

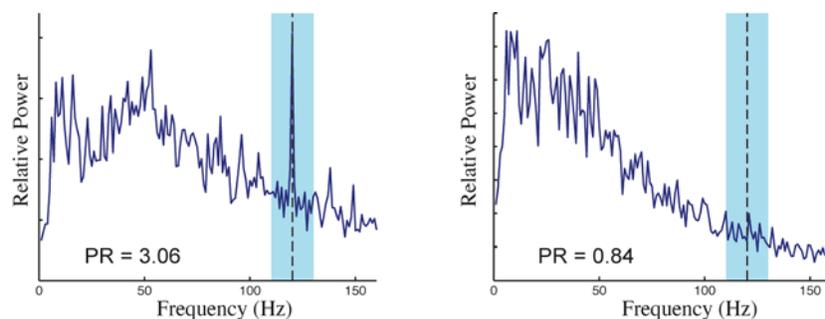


Fig. S5-1: The power spectra of two different X-cell responses to spatially uniform noise (SUN) stimuli. *Left:* significant phase-locking. *Right:* no phase-locking. The Power Ratio (PR) is the ratio of the power at the monitor refresh (*dashed line*) to the average power within ± 10 Hz (*blue shading*).

The distribution of power ratios for all cells considered in this study is shown in Fig. S5-2 for the SUN stimulus (A) and natural movie (B) – both had a monitor refresh rate of 120 Hz.

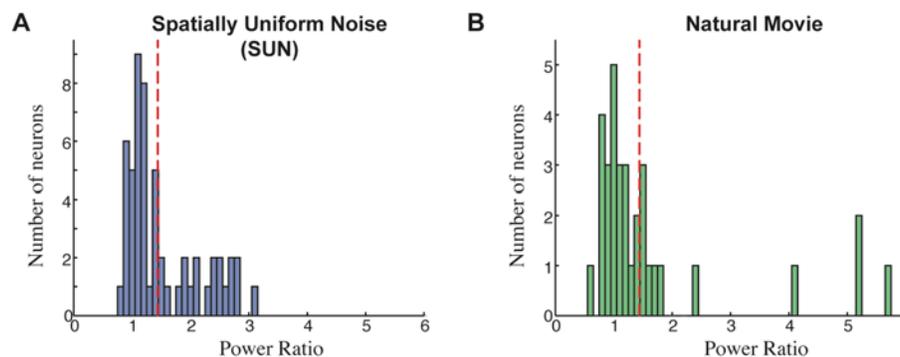


Fig. S5-2: Power ratios of neurons in SUN (A) and natural movie (B) stimuli, with the threshold for significant phase locking (Williams *et al.*, 2004) shown as a *dashed red line*.

By the criteria reported in Williams *et al.*, 40% of our cells are “significantly phase-locked” (*dashed red line*) with SUN stimuli, and 33% phase-locked with natural movies. Since phase locking can affect the temporal structure of spike trains, we took advantage of the range of PR magnitudes to estimate the effect of phase-locking on the time scale of the response PSTH.

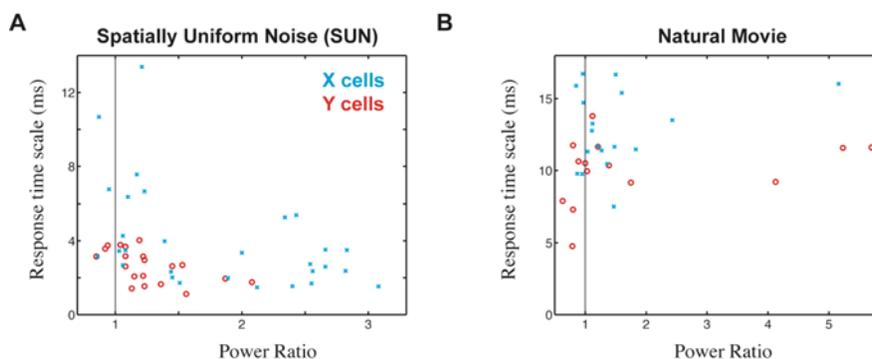


Fig. S5-3: Correlation between power ratio and response time scales for SUN (A) and natural movie (B) stimuli. Though they are correlated in SUN conditions, the correlations only account for a 2 ms difference between the most and least locked neuron. There was no discernible correlation in the natural movie condition.

In fact, the phase-locking ratio is correlated with the degree of temporal precision in the SUN condition (Fig. S5-3A), though could account for at most a 2 ms difference in response time scale between the significantly-phase-locked and not-significantly-phase-locked neurons. In the natural movie conditions, there was no correlation between the power ratio and response time scale.

The temporal precision of the response is thus only slightly affected by phase-locking, and the results presented in this paper – which hinge on much larger time-scale differences between the stimulus and response structure – are independent of phase locking to the monitor refresh.

REFERENCES

- Williams PE, Mechler F, Gordon J, Shapley R, Hawken MJ (2004) Entrainment to video displays in primary visual cortex of macaque and humans. *J Neurosci* 24:8278-8288.
- Wollman DE, Palmer LA (1995) Phase locking of neuronal responses to the vertical refresh of computer display monitors in cat lateral geniculate nucleus and striate cortex. *J Neurosci Methods* 60:107-113.