Proportional spike-timing precision and firing reliability underlie efficient temporal processing of periodicity and envelope shape cues

Y. Zheng1,4 and M. A. Escabí1,2,3

1Department of Biomedical Engineering, University of Connecticut, Storrs, Connecticut; 2Department of Electrical and Computer Engineering, University of Connecticut, Storrs, Connecticut; 3Department of Psychology, University of Connecticut, Storrs, Connecticut; and 4Laboratoire Psychologie de la Perception, CNRS, Université Paris-Descartes & Ecole Normale Supérieure, Paris, France

Submitted 13 December 2010; accepted in final form 29 April 2013

Zheng Y, Escabí MA. Proportional spike-timing precision and firing reliability underlie efficient temporal processing of periodicity and envelope shape cues. J Neurophysiol 110: 587–606, 2013. First published May 1, 2013; doi:10.1152/jn.01080.2010.—Temporal sound cues are essential for sound recognition, pitch, rhythm, and timbre perception, yet how auditory neurons encode such cues is subject of ongoing debate. Rate coding theories propose that temporal sound features are represented by rate tuned modulation filters. However, overwhelming evidence also suggests that precise spike timing is an essential attribute of the neural code. Here we demonstrate that single neurons in the auditory midbrain employ a proportional code in which spike-timing precision and firing reliability covary with the sound envelope cues to provide an efficient representation of the stimulus. Spike-timing precision varied systematically with the timescale and shape of the sound envelope and yet was largely independent of the sound modulation frequency, a prominent cue for pitch. In contrast, spike-count reliability was strongly affected by the modulation frequency. Spike-timing precision extends from sub-millisecond for brief transient sounds up to tens of milliseconds for sounds with slow-varying envelope. Information theoretic analysis further confirms that spike-timing precision and firing reliability can efficiently signal shape and periodicity temporal cues.

inferior colliculus; periodicity; repetition; envelope shape; modulation; jitter; reliability; noise; mutual information

Unlike other sensory systems, the auditory system must encode fast temporal modulation cues that vary in their frequency over several orders of magnitude. Perceptually relevant temporal cues in natural sounds can vary from as little as a few Hertz for slow word transitions in speech and vocalizations, ~1–20 Hz for rhythmic context, to substantially faster modulation frequencies, such as those observed within the pitch range (up to ~800 Hz) (Burns and Viemeister 1976; Greenberg 1999; Moore 1997; Pollack 1969; Plomp 1983). Although periodicity is an important acoustic attribute that influences the perception of pitch and rhythm, the envelope shape is a complementary temporal cue that influences the perceived loudness and timbre of a sound (Akeroyd and Patterson 1995; Irino and Patterson 1996). The envelope shape is defined by the temporal characteristics of the sound envelope within a single period, including structural features, such as the temporal attack and decay of the waveform and its duration. In music perception, for instance, the shape of the temporal envelope is a distinguishing factor that contributes to the quality of musical instruments (Paquette and Peretz 1997; Risset and Wessel 1982). Instruments and sounds with a fast attack time produce a percussive percept, while instruments with slow attack times tend to have a tonal quality (Akeroyd and Patterson 1995; Irino and Patterson 1996).

Numerous studies have demonstrated the importance of precise spike timing on the neural representation of temporal sound cues. Starting with the auditory nerve (AN), phase locking to temporal amplitude modulations is evident throughout the auditory system. In mammals, AN fibers can phase lock to temporal modulations up to and exceeding 1 kHz (Cariani and Delgutte 1996; Joris and Yin 1992). Yet this ability to phase lock to high-modulation frequencies is severely degraded at higher levels of the auditory system. In the central nucleus of the inferior colliculus (CNIC), temporal phase locking is limited to ~300 Hz (Krishna and Semple 2000; Langner and Schreiner 1988; Nelson and Carney 2007; Rees and Palmer 1989; Zheng and Escabí 2008), and this upper limit is further reduced to ~25 Hz for the majority of neurons in the auditory cortex (Joris et al. 2004). Despite this systematic degradation in temporal-following ability, neurons throughout the auditory system can encode the shape of transient sounds with remarkably high temporal fidelity (Heil 2001; Zheng and Escabí 2008). Such high fidelity is preserved even in the auditory cortex where neurons can respond to a sound onset with millisecond spike-timing precision (Heil 1997a, 1997b).

The CNIC is critically positioned for the analysis of temporal sound cues as it receives convergent input from all ascending brain stem nuclei and is a major convergent site of synaptic integration (Adams 1979; Oliver 2000). Unlike AN fibers, CNIC neurons exhibit rate tuning to amplitude modulations, and this finding has supported the view that periodic information is represented by a bank of rate-tuned modulation filters (Langner and Schreiner 1988; Schreiner and Langner 1988). Two recent studies, however, have demonstrated that rate-tuning properties are not invariant to changes in other temporal acoustic parameters, including the envelope shape (Krebs et al. 2008; Zheng and Escabí 2008), thus emphasizing the limitation of the rate-tuning hypothesis. In particular, the envelope shape and period of a sound can independently affect the temporal response pattern, and such changes can potentially be used to discriminate shape and periodicity cues (Zheng

Address for reprint requests and other correspondence: M. A. Escabí, Electrical and Computer Engineering, Univ. of Connecticut, 371 Fairfield Way, U4157, Storrs, CT 06269 (e-mail: escabi@engr.uconn.edu).
and Escabi 2008). Although these studies have argued that CNIC neurons can convey envelope sound information on the basis of the temporal firing pattern, how the shape and periodic structure of a sound are simultaneously represented by neurons is still poorly understood.

This study addresses how the periodic structure of an envelope is concurrently encoded with the shape of the envelope. Specifically, we examine the impact of precise spike timing and spiking reliability on the neural code for periodicity and shape cues in CNIC neurons. We first tested the hypothesis that sound modulation frequency and shape are reflected in the spiking-timing precision and firing reliability of CNIC responses. To do this, we develop a novel shuffled correlogram method to quantify the relationship between spiking-timing precision and reliability as a function of sound modulation frequency and envelope shape. Spike-timing precision covaried with the envelope shape of the sound and was largely independent of the sound modulation frequency. In contrast, spiking reliability systematically decreased with increasing modulation frequency. We next tested the hypothesis that an information theoretic optimal temporal encoding strategy mirrors and reflects the observed spike-timing precision and reliability trends. Specifically, the measured information efficiency and total information were strongly dependent on the spiking reliability and spike-timing precision in a manner that reflected the sound periodic structure and shape. We propose that spike-timing precision and reliability are integral attributes of the neural code that enable signaling of periodicity and shape temporal cues in the CNIC.

**METHODS**

**Animal Preparation and Recording**

Data from these experiments and the experimental procedures have been described in detail previously (Zheng and Escabi 2008). Briefly, surgery was performed on adult cats (N = 5) under pentobarbital sodium (25–30 mg/kg) and acepromazine (0.25 mg/kg). After insertion of an endotracheal tube, the pia mater were retracted, and the animal was placed in a stereotaxic assembly with hollow ear bars (David Kopf Instruments, Tujunga, CA). The anterior fossae was exposed, and the overlying occipital cortex was aspirated. The bony tentorium was placed in a stereotaxic assembly with hollow ear bars (David Kopf Instruments, Tujunga, CA). The anterior fossae was exposed, and the overlying occipital cortex was aspirated. The bony tentorium was removed to fully expose the inferior colliculus. Following the exposure, the animal was maintained in a reflexive state via continuous infusion of ketamine HCl (6–10 mg·kg⁻¹·h⁻¹) and diazepam (0.4–0.7 mg·kg⁻¹·h⁻¹) in lactated Ringer solution. The infusion rate was adjusted according to physiological criteria (heart rate, breathing and reflexes). Every 12 h the animal was given dexamethasone (1.2 mg/kg) to prevent brain edema and atropine (0.04 mg/kg) to reduce salivation. All animals were housed and handled according to procedures approved by the University of Connecticut Animal Care and Use Committee, and in accordance with guidelines set by the National Institutes of Health, US Department of Agriculture, and the American Veterinary Medical Association.

Neuronal recordings from 65 single units were obtained from the CNIC. Data from these units have been described previously, although the present results consist of a separate analysis of the data (Zheng and Escabi 2008). Epoxy-coated (4–8 ΜΩ at 1 kHz, AM Systems, Carlsborg, WA) or glass-coated tungsten electrodes (5–15 ΜΩ, 1–2-μm tip diameter) were advanced with a Burleigh microdrive (ULN 6000, Burleigh Instruments, Fishers, NY) at an angle of ~30° relative to the sagittal plane and roughly orthogonal to the CNIC frequency band lamina (Oliver and Morest 1984). Neuronal traces were amplified and passed through a window discriminator. Spike event times were then acquired digitally at a rate of 12.2 kHz (RA16PA, RX6; Tucker Davis Tech., Alachua, FL). Initially, noise bursts (50-ms duration; 300-ms intervals) were delivered to find single neurons. Neurons within the central nucleus were identified physiologically by presenting tone-pips (50-ms duration, 10-ms rise-fall time; 300-ms interval) to measure the frequency tuning of each unit and to identify the low- to high-frequency tonotopic gradient (Merzenich and Reid 1974). For the neurons studied, best frequencies (BF) extended from 1.5 to 40.8 kHz (mean BF, 13.9 ± 9.4 kHz; median BF, 11.1 ± 16.9 kHz).

**Acoustic Stimuli and Delivery**

The stimuli used for these experiments have been described previously (Zheng and Escabi 2008). Briefly, 5-s segments of periodic noise bursts (PNB) and sinusoidal amplitude modulated noise (SAMN) were presented at a fixed sound level [80-dB peak sound pressure level (SPL)]. Sounds were delivered diotically over calibrated speakers (EC1 Electrostatic Diaphragms, Tucker Davis Technologies, Alachua, FL; frequency response 1–47 kHz, ±3 dB SPL). The stimulus paradigm consists of orthogonal acoustic conditions that can be used to dissociate neural responses to shape and periodicity cues (Zheng and Escabi 2008). Specifically, although we can vary the modulation frequency of PNB, the envelope waveform shape of the sound is fixed for each modulation cycle (250 µs duration; 50 µs rise and decay time, etc.). In contrast, for SAMN, the envelope waveform shape varies with the modulation frequency such that the envelope rise times are inversely related to the modulation frequency. Thus varying the modulation frequency changes the shape and period of the SAMN sound (variable shape and period), while for PNB it strictly changes the sound period (constant shape, variable period). For both PNB and SAMN, the noise carriers consisted of independent broadband (0–48 kHz) uniformly distributed noise (unfrozen). In two experiments, PNB and SAMN repetition rates spanned 5–500 Hz (n = 12 recording sites) in logarithmic increments (0.44 octave steps, 15 modulation frequencies), while in the latter three experiments (n = 53 recording sites) the upper modulation frequency was extended to 1,341 Hz (18 modulation frequencies). Both SAMN and PNB were presented in random interleaved order (10 repeats for each stimulus condition) with a 500-ms pause between consecutive conditions. Data in this study are only shown for modulation frequencies up to 100 Hz because significant synchronized responses were not observed beyond this modulation frequency. The first 500 ms from each stimulus trial were discarded prior to analysis because neuronal responses exhibited a fast adaptation during the initial ~200 ms (Zheng and Escabi 2008).

**Shuffled Correlogram**

We measured shuffled correlograms (Joris et al. 2006; Zheng and Escabi 2008) to characterize the spiking precision and reliability of CNIC neurons. We have previously demonstrated that neurons in the CNIC exhibit distinct temporal response patterns to PNB and SAMN. Here we extend this procedure to characterize the spike-timing precision and reliability of neurons and their dependence on the stimulus period and the envelope shape. The shuffled autocorrelogram was computed as the mean autocorrelogram between all possible 4-cycle response segments in the sound:

\[
\Phi_{\text{s shuffled}}(\tau) = \frac{1}{N(N-1)} \sum_{k=1}^{N} \sum_{\tau k} \Phi_{k}(\tau)
\]

where \(\Phi_{k}(\tau)\) is the autocorrelogram between the k-th and l-th 4-cycle response segments and \(N(N-1)\) is the number of spike train segments being correlated. Because the PNB and SAMN are periodic stimuli, the data were first broken up into disjoint 4-cycle segments, and the correlograms were computed with a circular cross-correlation between each segment, as described previously (Zheng and Escabi 2008).
The circular cross-correlation ensures that there are no edge artifacts at the boundaries between consecutive segments of the data. The shuffled correlogram was measured for both PNB and SAMN at all of the tested modulation frequencies.

Prior to computing the shuffled correlograms, the PNB responses were resampled at a rate of 2 kHz while the SAMN responses were sampled at 50 samples per stimulus period, as described previously. This was done because neurons exhibited millisecond to sub-millisecond spike-timing precision, regardless of modulation frequency for PNB. For SAMN the spiking precision was dependent on, and varied considerably with, the sound modulation frequency. The procedure was also tested using 5-kHz and 10-kHz sampling rate for both SAMN and PNB. These produced similar results, although the correlograms tended to be noisier at these higher sampling resolutions. In particular, at low-modulation frequencies (e.g., 5 Hz), the SAMN correlograms were excessively noisy because spikes were not tightly temporally aligned as for PNB. Thus by sampling the data at 50 samples/period (SAMN) and 2 kHz (PNB), we significantly reduced measurement noise. Details for the shuffled correlation analysis and sampling procedure are provided in Zheng et al. (2008).

### Quantifying Spike-timing Jitter and Spike-count Reliability

We consider a neuron model that produces a quasi-periodic spiking response and contains three forms of spiking “noise” or variability observed in vivo and that are evident in the PNB and SAMN response dot-rasters. First, spike-count reliability errors were included in the model to account for the neuron’s trial-to-trial variability. Second, stochastic spike-timing errors (jitter) were included in the model to account for variability in the spike arrival times. Finally, random spiking noise was included to account for the fact that some neurons produced a baseline DC firing rate that was not modulated by the sound. The expected shuffled autocorrelogram for the model spike train is expressed as (see Appendix A):

$$\Phi_{\text{shuffled}}(\tau) = \bar{x}^2 \cdot f_m \cdot \sum_n \Phi_{f_{\text{in}}} (\tau - n \cdot T) + 2 \cdot \lambda_{\text{period}} \cdot \lambda_{\text{noise}} + \lambda_{\text{noise}}^2$$

(2)

where $\bar{x}$ is the average number of reliable phase-locked spike per period of the stimulus; $f_m$ is the modulation frequency of the sound; $T = 1/f_m$ is the periodic modulation period; $\lambda_{\text{period}} = \bar{x} \cdot f_m$ is the spike rate for the periodic component of the neural response; $\lambda_{\text{noise}}$ is the spike rate associated with the undriven firing of the neuron (i.e., nonphase locked); $f(t)$ is the probability distribution function for the spike-timing jitter errors; and $\Phi_f(\tau)$ is the autocorrelation of the spike-timing jitter errors $[f(t)$ correlated with itself]. If the spike time jitter is assumed to be normal with zero mean and variance $\sigma^2$, the shuffled correlogram is expressed as:

$$\Phi_{\text{shuffled}}(\tau) = \bar{x}^2 \cdot f_m \cdot \sum_n \frac{1}{\sqrt{4 \pi \sigma^2}} \cdot e^{-\frac{(\tau - n \cdot T)^2}{4 \sigma^2}} + 2 \cdot \lambda_{\text{period}} \cdot \lambda_{\text{noise}} + \lambda_{\text{noise}}^2$$

(3)

Given that the total spike rate is assumed to be the sum of the driven periodic response and undriven spiking noise, Eqs. 2 and 3 are both constrained by:

$$\lambda_{\text{total}} = \lambda_{\text{period}} + \lambda_{\text{noise}} = \bar{x} \cdot f_m + \lambda_{\text{noise}}$$

(4)

To quantify the spike-timing precision and reliability of CNIC units as a function of the stimulus condition, the optimal response parameters were derived by fitting Eqs. 3 and 4 to the measured shuffled autocorrelogram of each unit at each sound condition using generalized least squares optimization. The response spike-timing jitter ($\sigma$) and the reliable spike count per cycle ($\bar{x}$) were derived for each modulation frequency for both PNB and SAMN. These response metrics were then plotted as a function of modulation frequency (e.g., see Fig. 3). We point out that the reliability parameter ($\bar{x}$) is actually the average number of reliable spikes evoked by one cycle of stimulus. For SAMN the number of reliable spikes per cycle was often greater than 1, while for PNB the number of spikes per cycle rarely exceeded 1, so that the parameter $\bar{x}$ provides an appropriate metric of the spiking reliability (i.e., the probability of a hit or miss of a single action potential), as described in Appendix A (last paragraph).

Finally, to quantify the fractional number of spikes that were reliably phase-locked to the stimulus, we computed the temporal coding fraction ($F$) as the ratio of phase-locked power to total power in the spike train:

$$F = \frac{1}{T} \int_0^T \Phi_{\text{period}}(\tau) d\tau \over \frac{1}{T} \int_0^T \Phi_{\text{shuffled}}(\tau) d\tau = \frac{\lambda_{\text{period}}^2}{\lambda_{\text{total}}^2} = \frac{\lambda_{\text{period}}^2}{(\lambda_{\text{period}} + \lambda_{\text{noise}})^2}$$

(5)

where $\Phi_{\text{period}}(\tau)$ represents the shuffled autocorrelogram of the periodic response component. As can be seen, this metric lies between 0 and 1, where 1 indicates that 100% of the action potentials are reliably phase-locked to the periodic sound.

Significance testing of the parameters was performed by considering a Poisson model of identical spike rate as a null hypothesis. The Poisson neuron assumption was chosen as a reference because the action potentials for such a neuron would be randomly distributed in the cycle histogram, and the autocorrelogram will not exhibit a synchronized peak (i.e., $\bar{x}$ and $\lambda$ are both 0). The Poisson model was simulated and the data were bootstrapped to derive error bounds for the real data and the Poisson simulation. A $t$-test was then performed for the real data to test against the Poisson model at significance level of $P < 0.05$ for all parameters. Significant temporal synchrony was not observed beyond 500 Hz, and, therefore, data conditions beyond this range were not included in the present findings.

### Model Validation

The quality of the model was assessed with a cross-validation procedure in which half of the neural data were used for model optimization, and the remaining half for model validation. To do this, the dot-rasters at each modulation condition were first subdivided into non-overlapping 4-cycle segments. Odd and even segments were used to compute independent shuffled correlograms that were used for model validation [$\Phi_f(t)$] and optimization [$\Phi_f(t)$], respectively. The validation procedure was performed for every neuron at all sound conditions (PNB and SAMN, multiple modulation frequencies) that yielded statistically significant synchrony ($P < 0.05$, as described in the previous paragraph). Because, as shown in the RESULTS section, the response properties of neurons (i.e., baseline firing rate, spike-timing jitter, firing reliability) can still change substantially with the specific stimulus condition (e.g., PNB vs. SAMN, 5 Hz vs. 100 Hz), the validation procedure is not designed to assess how well the model generalizes across stimulus conditions. Instead, our intent is to use the model to derive meaningful parameters of the neural response (spike-timing jitter and reliability), which can then be used to characterize the relationship between a particular sound and the response pattern.

Two error metrics were defined to assess the quality of the model fits. First, we considered the fractional model error:

$$E = \frac{\text{var}[\Phi_f(t) - \Phi_{\text{model}}(t)]}{\text{var}[\Phi_f(t)]} \times 100\%$$

(6)

This error metric expresses the power of the difference signal as a percentage of the power of the empirical shuffled correlogram.

Because the neural data contain both deterministic model errors as well as measurement noise, we devised a second error metric that compensates for the measurement noise and strictly captures deterministic errors resulting from deviations in the model fits. The noise-corrected error metric was obtained as (Appendix B):
Like the error metric of Eq. 6, this metric expresses the model error as a percentage of the total signal power; however, bias due to measurement noise is removed from the estimate (see APPENDIX B). As expected, both metrics yield similar results, although the corrected error metric yielded slightly smaller estimates of the errors.

Information Theoretic Analysis

We consider how much information is contained by the temporal response pattern for each stimulus. Specifically, we are interested in identifying how spike-timing precision and firing reliability contribute to the representation of periodic stimuli with different shapes. To do this, we defined and computed the specific temporal information (STI) for each stimulus condition (i.e., the specific stimulus: modulation frequency, PNB, or SAMN) and determined how spike-timing reliability and jitter contribute to the information encoding efficiency using an information theoretic approach. Formally, it is shown that the total mutual information can be decomposed as (APPENDIX C)

$$I(W, S) = \sum_k p(W_k, S_j) \cdot I_{STI}(W_k, S_j) + I_{EI}$$

(8)

where $I_{STI}(W_k, S_j)$ is the STI that is contained in the temporal response pattern for a specific stimulus-response ensemble ($S_j, W_k$). The second term $I_{EI}$ is the ensemble information (EI) that can be gained by comparing responses across the stimulus-responses ensembles. The decomposition of Eq. 8 is based on the assumption that both the stimuli and responses are conditionally independent (see APPENDIX C). Here, the primary goal is to decipher how the stimulus information was carried by the temporal response pattern for each specific stimulus, and thus we will focus on characterizing the STI.

The STI was calculated using a model-based approach in which we simulated model spike trains fitted to the original data. We employed this approach because we can synthesize extended spike trains with a larger number of response cycles, which is necessary to minimize bias in the error metric of Strong et al. 1998. As shown in Fig. 2, the model spike trains closely resembled the original spike trains, both in terms of the deterministic response and the statistical structure. Hence the resulting STI trends were the same for the original data and model, although numerically smaller for the model spike trains (due to the reduced positive information bias; see Information bias correction).

Measuring STI. Neural responses were first fitted to a spiking model with spike-timing jitter, reliability, and additive spontaneous firing using Eqs. 3 and 4. Simulated spike trains with 8,000 response cycles were then generated according to Eq. A2 for each stimulus condition, as illustrated in Fig. 2.

The STI was estimated from these model spike trains by adapting the procedure of Strong et al. (1998):

$$I_{STI}(k) = H_{Total}(k) - H_{Noise}(k)$$

(9)

where $H_{Total}(k)$ and $H_{Noise}(k)$ are the total and noise entropy, respectively, after conditioning on a specific stimulus-response ensembles ($S_j, W_k$). We will refer to these as the specific entropy and specific noise entropy, respectively. Note that for simplicity, we use the abbreviated notation $I_{STI}(k) = I_{STI}(S_j, W_k)$.

The specific total entropy $[H_{Total}(k)]$ was estimated by measuring the probability distribution of observing arbitrary B-bit words in the neural response (Strong et al. 1998):
larger for low-modulation frequencies, which further exacerbates the bias at low-modulation frequencies (e.g., $B = 5$ at 200 Hz vs. $B = 200$ at 5 Hz). Together, these two factors lead to a substantially larger bias at low-modulation frequencies, even though the sound duration was identically for all conditions.

To determine a suitable number of response cycles that would minimize the bias of the simulated model spike trains, we first generated simulated model spike trains according to Eq. A2, with a variable number of response cycles (500, 1,000, 2,000, 4,000, 8,000). The spike trains were optimized to match the temporal response pattern for each condition, as outlined in Eqs. 3 and 4 and shown in Fig. 2. The measured information dropped off with increasing number of cycles due to decreasing bias. For 8,000 cycles, the residual bias was small (<10%), and thus the STI was calculated using the 8,000 cycles condition.

To further correct for unaccounted residual bias, we subtracted the measured STI for a random Poisson neuron of equal spike rate: $I_{STI}(k) = I_{STI}(k) - I_{ResBias}(k)$. We refer to this quantity as the residual bias $I_{ResBias}(k)$. Note that, ideally, STI is zero for the response of a Poisson neuron that fires independently from trial to trial in response to the periodic sounds. However, due to limitations in the information calculation and data size, the measured information for the Poisson spike train was found to be greater than zero. When tested across data sizes (500, 1,000, 2,000, 4,000, 8,000 cycles), the trend for the reduction in this residual bias with increasing data size was nearly identical to that observed for reductions in the STI, indicating that the residual bias for a Poisson condition accounts for the bias in the simulated data at multiple conditions. For the 8,000 cycles simulations, the mean residual bias was small at 2.8% and 3.0% for PNB and SAM, respectively.

Temporal Noise Entropy, Reliability Noise Entropy, and Coding Efficiency

To characterize how spike-count reliability and spike-timing jitter affect the encoding and temporal information transfer of CNIC neurons, the specific noise entropy was decomposed by considering the fractional entropy that resulted from reliability and spike-timing noise. Consider a framework in which a noiseless neuron produces a perfectly reproducible response for each modulation cycle. The neural spike train and the corresponding B-bit words for the noiseless neuron are perfectly reproducible across cycles so that the noise entropy would be precisely 0 bits because there is no variability from cycle to cycle in the response words.

For a real neuron, the cycle-to-cycle spike train variability could result from either spike-count variability or spike-timing errors. For a specific stimulus-response ensemble, the specific noise entropy reflects these two forms of variability and thus can be decomposed as $H_{Noise}(k) = H_{temporal}(k) + H_{reliability}(k)$ (APPENDIX D), where the reliability noise entropy $H_{reliability}(k)$ strictly arises from spike-count reliability errors, and the temporal noise entropy $H_{temporal}(k)$ arises from spike-timing errors. For simplicity, we assume that the specific stimulus is implied throughout and will use the abbreviated notation $H_{Noise} = H_{temporal} + H_{reliability}$.

The reliability noise entropy is defined as the entropy that strictly arises from deviations in the spike count from cycle to cycle. That is, it reflects only variability due to spike-count reliability errors across cycles independently of spike-timing errors. It is computed as the entropy of the spike count distribution, which reflects cycle-to-cycle variability not due to spike-timing errors (APPENDIX D, Eq. D6). In contrast, the temporal noise entropy strictly arises from spike-timing errors and is formally derived from the conditional response word distribution when the spike count is fixed (APPENDIX D, Eqs. D4 and D5):

$$H_{temporal}(l) = \left( \sum_{w \in L} p_{Noise}(w|l) \log_2 \left[ p_{Noise}(w|l) \right] \right)$$

where $L$ is the subset of words with precisely $l$ spikes, $p_{Noise}(w|l)$ is the word distribution conditioned for $l$ spikes, $H_{temporal}(l)$ corresponds to the specific temporal entropy when the spike count is fixed, and $H_{temporal}$ is the average specific temporal entropy (averaged across all spike counts).

To quantify the contribution of spike-timing jitter and spiking reliability in the neural response and to characterize how these forms of noise depend on the sound period and envelope shape, we consider the fraction of the noise entropy that results from reliability or spike-timing errors for a specific stimulus. The reliability noise fraction and temporal noise fraction are defined as the proportion of the total spike train entropy that could be accounted for by reliability or spike-timing errors:

$$F_{reliability} = \frac{H_{reliability}}{H_{Total}} \quad (14)$$

and

$$F_{temporal} = \frac{H_{temporal}}{H_{Total}} = \frac{H_{Noise} - H_{reliability}}{H_{Total}} \quad (15)$$

Note that the reliability and temporal noise fractions both depend on the temporal analysis resolution. Furthermore, the proportion of specific noise entropy in the response (i.e., the noise fraction) is simply the sum of the reliability and temporal noise fractions:

$$F_{Noise} = F_{reliability} + F_{temporal} \quad (16)$$

Based on the definition of coding efficiency (Rieke et al. 1995), it is easy to show that these quantities are intimately related to the specific temporal coding efficiency of the neuron according to the following relationship:

$$E = \frac{I_{STI}}{H_{Total}} = 1 - \frac{H_{Noise}}{H_{Total}} = 1 - F_{Noise} \quad (17)$$

Thus the temporal encoding efficiency decreases as either the reliability or temporal noise fraction increase.

Efficiency and Spike-timing Resolution

We also determined how encoding efficiency depend on the resolution of the neural spike train. To do this, we varied the effective spike train resolution by synthetically adding spike-timing jitter (see Fig. 6) to the neural spike trains prior to computing the STI, efficiency, and noise fractions. This procedure is analogous to that employed previously to determine the role of precise spike timing (Kayser et al. 2010; Lu and Wang 2004). Information metrics were calculated for each sound condition using different amounts of added spike-timing jitter (uniformly distributed, 1- to 200-ms range, in 0.38 octave steps). This procedure is used to characterize how temporal resolution affects encoding efficiency for both PNB and SAM.

RESULTS

We examined how envelope periodicity and shape cues influence spike-timing precision and reliability and how these factors influence information transmission in the cat CNIC. Specifically, we hypothesized that the envelope shape and period will systematically affect the spike-timing precision and reliability so that these temporal response attributes can potentially convey substantial stimulus information. Data consist of 65 single units that were driven with PNB and SAM (Fig. 1, top; shown for 5, 7, and 10 Hz). Both sounds span identical
modulation frequencies; however, their envelope shapes are distinctly different. Specifically, although the period of the PNB can be changed by adjusting the modulation frequency parameter, the shape of the envelope (i.e., the attack and decay times) is constant for all modulation conditions (Fig. 1, bottom left; consisting of a brief 250-μs pulse). Thus differences in neural firing pattern for different PNB modulation frequencies will strictly reflect changes in the period of the sound. In contrast, the envelope shape for SAMN covaries with the sound modulation frequency so that both acoustic attributes can theoretically affect the neural response (Fig. 1, bottom right). For low-modulation frequencies, the envelopes exhibit slow attack and decay times, and these become faster with increasing modulation frequency. These stimulus conditions thus allow us to parse out how pure repetition (PNB) or combined repetition and shape cues (SAMN) are concurrently encoded (Zheng and Escabi 2008).

The temporal structure of the neural responses to periodic stimuli was strongly dependent on the envelope shape. Example responses are illustrated for three units as 400-ms dot-raster segments for 10-Hz modulation frequency (Fig. 2, A–C, middle panel). As previously demonstrated, the neural responses exhibit distinct discharge patterns for PNB and SAMN that are complementary to the shape of the envelope waveform (Zheng and Escabi 2008). First, units often displayed precisely phase-locked response to the brief noise burst of the PNB, as evidenced by the tightly alignment episodes in the dot-rasters (Fig. 2A, middle panel) and the sharp events in the shuffled correlogram of a single neuron (Fig. 2D). Second, the SAMN spike trains are often more variable, although the spiking pattern accurately reflects the envelope structure of the sound (Fig. 2, B and C, middle panel). This behavior is seen in the response shuffled correlogram, which displays a temporally broad periodic pattern that closely mimics the SAMN stimulus envelope as seen from the shuffled autocorrelograms (Fig. 2E, F).

The Contribution of Envelope Periodicity and Shape to Spiking Reliability and Precision

The example response patterns of Fig. 2 demonstrate that neurons can encode the sound envelope waveform by their precise discharge pattern. Neural precision clearly varies for the example units; however, the factors that govern precision and reliability of neural spiking in the CNIC are unknown. One possibility is that spiking precision and reliability strictly reflect intrinsic properties of the neuron, such as its membrane time constant, refractory period, or the synaptic variability. For this scenario, it is expected that precision and reliability are largely stimulus independent, because precision and reliability would be dominated by the neuron’s biophysical properties. Alternately, it is also possible that spiking precision and reliability are strongly influenced by the stimulus characteristics. Specifically, although it is generally assumed that periodicity is the primary acoustic factor affecting temporal response properties, we hypothesize that the envelope waveform shape strongly influences the unit spiking properties. As illustrated in Fig. 2, A–C, the discharge spike-timing precision (jitter) for the select neurons appear to depend on the stimulus condition. Responses to PNB are typically characterized by precisely phase-locked action potentials, while the discharge patterns to SAMN at identical modulation frequency lack precisely phase-locked spikes, and the firing pattern appears to reflect the envelope shape. These differences suggest that the stimulus envelope shape and period may independently influence the precision and reliability of firing.

We first characterized how shape and periodicity influence spike-timing precision and firing reliability. To do this, we developed a stochastic neural model that incorporates three
Fig. 2. Temporal reliability and spike-timing jitter analysis. A–C, middle panel: example neural response to PNB (A) and SAMN (B and C) at 10-Hz modulation frequency are shown as dot-rasters for three single neurons [A: best frequency (BF) = 17.5 kHz, B: BF = 4.8 kHz, C: BF = 20.7 kHz]. For reference, the top panel depicts the 400-ms segment of the corresponding PNB and SAMN sounds. Note that the vertical alignment of the action potentials (dots) is indicative of the precision of the spike arrival times (jitter). For PNB, the neural response is characterized by phasic periodic pattern, while for SAMN the action potentials are broadly distributed across the sound modulation cycles. A–C, bottom panel: dot-rasters for the proposed neural model that incorporates spike-timing and spiking reliability errors (see METHODS; APPENDIX A) accurately replicates the observed neural response patterns. The best-fit model was obtained for each condition by finding the spiking reliability, spike-timing jitter, and baseline firing rate of the model that most closely matched the neural response. D–F: shuffled correlograms for the responses shown in A–C (black), respectively. The shuffled correlogram of the model is superimposed (red). G–J: summary data for the raw (G and I) and noise-corrected model errors (H and J; see METHODS; APPENDIX C) are shown as box-whisker plots as function of sound condition [PNB (I and J) and SAMN (G and H); multiple modulation frequencies]. In all panels, errors are normalized as percentage of the total response power. The notches in the box designate the median errors, and the limits of the box determine the upper and lower quartile of the errors. Whiskers lengths are set to 1.5 times the interquartile range. Neurons that exceeded this range are shown as +. The model fitting was performed only for conditions with statistically significant firing reliability relative to a Poisson neuron of identical spike rate (P < 0.05; see METHODS).
forms of spiking noise seen in vivo: spike-timing jitter, spiking reliability, and random asynchronous spikes (see METHODS and APPENDIX). Dot-rasters from the model fits to the neural data are shown in Fig. 2, A–C (bottom, red dot-rasters). Visually, the model replicates the temporal response pattern quite well. For each unit and stimulus condition, we first computed the shuffled correlogram and developed a closed-form solution for the shuffled correlogram of the spiking model under the assumption of normally distributed spike-timing jitter (METHODS, Eq. 3). The model parameters (jitter, number of reliable spikes, and baseline firing rate) were iteratively optimized to find the model configuration that most closely resembled the neural data (see METHODS). Shuffled correlogram model fits are illustrated for the examples of Fig. 2, D–F (superimposed red curve). Consistent with the model dot-rasters, the predicted shuffled correlograms provide a good approximation to the curve. Consistent with the model dot-rasters, the predicted shuffled correlograms under the assumption of normally distributed spike-timing jitter (METHODS, Eq. 3). The model parameters (jitter, number of reliable spikes, and baseline firing rate) were iteratively optimized to find the model configuration that most closely resembled the neural data (see METHODS). Shuffled correlogram model fits are illustrated for the examples of Fig. 2, D–F (superimposed red curve). Consistent with the model dot-rasters, the predicted shuffled correlograms provide a good approximation to the curve.

Multiple modulation conditions. As can be seen, the noise-corrected errors are small for the majority of neurons and conditions, implying that the model accurately fits the neural data (SAMN median < 2%, PNB < 10% for most conditions).

Spiking reliability and precision were strongly dependent on the stimulus condition, as shown for an example CNIC neuron (Fig. 3). Shuffled correlograms are shown for a single neuron at different repetition rates for PNB (Fig. 3A, black) and SAMN (Fig. 3C, black) along with the model fits (red). The shuffled correlograms for PNB are characterized by a narrow peak at zero lag and at intervals corresponding to the stimulus period (Fig. 3A). As can be seen, the unit exhibited precise time-locking to the stimulus (~1-ms resolution) over a wide range of modulation frequencies. For SAMN, the shuffled correlograms were generally much broader and lacked the high temporal precision observed for PNB (Fig. 3C). Furthermore, while the temporal spiking precision was relatively fixed for PNB, it was much more variable for SAMN. Note that the width of the correlogram peaks for SAMN become progressively narrower with increasing modulation frequency.

Fig. 3. Spike-timing precision and reliability analysis for a single neuron [BF = 19.0 kHz, best modulation frequency (BMF) (PNB) = 26 Hz, BMF (SAMN) = 100 Hz]. A and C: shuffled correlograms for the single unit (black) are shown for PNB and SAMN, respectively, over a range of modulation frequencies (5, 10, 50, 186 Hz). The shuffled correlogram of the optimal (best-fit) neuron model is shown for each condition as a superimposed red curve. B and D, top to bottom: the rate modulation transfer function (black) for this neuron exhibited a predominant high-pass pattern for PNB and a notch response pattern for SAMN. The synchrony modulation transfer function (superimposed green curve) displayed low-pass pattern for PNB and band-pass pattern for SAMN. Synchrony data points with nonsignificant value are not plotted (Raleigh statistic, P < 0.001). By comparison, the spiking reliability transfer function (i.e., the number of temporally reliable spikes as a function of modulation frequency) demonstrates that spiking reliability decreases systematically with increasing modulation frequency. The spike-timing jitter transfer function for the select neuron shows how spike-timing jitter varies with the sound modulation frequency. For PNB (B), the unit exhibited low jitter (<2 ms), as observed for frequencies between 5 and 69 Hz, where the response was statistically significant (t-test, P < 0.05). For SAMN (D), the temporal precision decreased systematically with the sound modulation frequency. For reference, gray lines represent the expected estimated jitter for a Poisson neuron of identical spike rate (Avissar et al. 2007). B and D, bottom: temporal coding fraction transfer functions for the selected neuron are shown for PNB and SAMN, respectively. The temporal coding fraction represents the proportion of the spike train power that is attributed to reliable temporally phase-locked action potentials. For PNB, the temporal coding fraction exhibits a low-pass pattern indicating that reliable, temporally locked activity was present for low frequencies. By comparison, the temporal coding fraction of SAMN is tuned, indicating that reliable temporally locked activity was dominant within the vicinity of 100-Hz modulation frequency.

Dots on the curves of B and D indicate significant results relative to a Poisson neuron with identical rate modulation transfer function (t-test, P < 0.05).
The relationship between firing reliability and spike-timing jitter for the example neuron exhibited distinct trends for PNB and SAMN (Fig. 3, B and D). First, spike-timing jitter was exceedingly small for PNB (<2 ms; Fig. 3B, third panel) for all modulation frequencies that produced significant results (5–69 Hz; tested against a Poisson neuron of identical firing rate, t-test, P < 0.05). In contrast, spike-timing jitter was substantially larger (up to ~13 ms) and decreased systematically with increasing modulation frequency for SAMN (Fig. 3D, third panel). Although the neuron exhibited a high-pass rate modulation transfer function pattern for PNB and a notch rate modulation transfer function pattern for SAMN (Fig. 3, B and D, top), the number of reliable spikes exhibited consistent patterns with modulation frequency (Fig. 3, B and D; second panel). The number of reliable spikes is roughly flat (~0.5 spikes/cycle) up to around 20 Hz for PNB and decreases systematically with increasing modulation frequency. Similarly, for SAMN, the number of reliable spikes decreased systematically with increasing modulation frequency. Finally, the temporal coding fraction quantifies the fractional power in the neural response that is phase-locked to the stimulus envelope and which potentially conveys temporal information (see METHODS). For this unit, the temporal coding fraction for PNB was close to 1 below 97-Hz modulation frequency (Fig. 3B, bottom), indicating that close to 100% of the action potentials were phase-locked to the sound and contributed to the temporal information within this range of modulation frequencies. In contrast, the temporal coding fraction for SAMN exhibited a band-pass behavior for this unit with a maximum at ~100 Hz (Fig. 3D, bottom).

Spike-timing jitter and reliability for SAMN and PNB exhibited distinct trends across our population of units. First, there was not a strong dependence between jitter and reliability for PNB as evident by the lack of correlation between these parameters (Fig. 4A; r = −0.1 ± 0.05, nonsignificant). The scatter plot is clustered at fine temporal precision with most jitter values below 1 ms. For most modulation frequencies and units, the number of reliable spikes per stimulus rarely exceeded 1 spike/stimulus (average = 0.36 spikes/stimulus between 5- and 259-Hz modulation frequency), suggesting that the PNB envelope is explicitly represented by precisely phase-locked unitary action potentials. Furthermore, average firing rates are relatively independent of modulation frequency (Fig. 5A, red) and were weakly related to the firing reliability (r = 0.22 ± 0.48, mean ± SD). In contrast, the response synchrony exhibited a low-pass pattern (Fig. 5B, red). Both the average spiking reliability (Fig. 5C, red) and temporal coding fraction (Fig. 5D, red) exhibited a low-pass trend for PNB with modulation frequency. This is consistent with the idea that a low-pass...

Fig. 5. Population response trends. A: the average normalized firing rate across neurons is shown as a function modulation frequency for SAMN (black) and PNB (red). For both SAMN and PNB, firing rates are relatively independent of modulation frequency. B: the response synchrony exhibits a low-pass trend for PNB. For SAMN, the overall synchrony is lower and exhibits a band-pass pattern across the neural population. C: the average reliability decreases monotonically with increasing modulation frequency for both SAMN and PNB. D: a similar low-pass trend is observed for the mean temporal coding fraction (i.e., the proportion of the spike train power that is phase locked to the sound), indicating that phase-locked activity contributes most for low frequencies. E: for SAMN, the mean spike-timing jitter decreased systematically with increasing modulation frequency. In contrast, for PNB, the mean spike-timing jitter was small (~300 μs) and relatively constant for modulation frequencies below 200 Hz (inset shows for PNB). Gray line indicates the theoretical spike-timing jitter that would be measured for a Poisson neuron (Avissar et al. 2007). Error bars indicate standard error of the mean.

J Neurophysiol • doi:10.1152/jn.01080.2010 • www.jn.org
filtering mechanism limits the amount of temporal phase-locking to the stimulus envelope. Furthermore, unlike spiking reliability, the average spike-timing jitter was small and relatively constant (∼300 μs) across modulation frequencies for PNB (5–259 Hz) (Fig. 5E, red, bootstrap t-test, $P < 0.05$). Thus the temporal coding precision for PNB does not depend strongly on the sound modulation frequency and is likely more influenced by the constant envelope shape.

In contrast to the highly stereotyped PNB response pattern, the SAMN response pattern was more varied across modulation frequencies. For each condition, the spike-timing jitter and the number of reliable spikes for different units were strongly clustered for different modulation frequencies (Fig. 4B). The clustering of these response parameters was strongly dependent on the modulation condition, leading to a strong correlation between jitter and reliability ($r = 0.65 \pm 0.05$, mean ± SD; $P < 0.01$). Spike rates were relatively constant across modulation frequencies (Fig. 5A, black) and were weakly related to the firing reliability ($r = 0.35 \pm 0.41$; mean ± SD). The average vector strengths (VS) exhibited a band-pass trend (Fig. 5B, black) for SAMN and was also weakly correlated with the spike-timing jitter ($r = 0.22 \pm 0.35$; mean ± SD). Similar to PNB, the average reliability (Fig. 5C, black) and the temporal coding fraction (Fig. 5D, black) decreased with increasing modulation frequency. Although the firing reliability and temporal coding fraction trends exhibit a similar low-pass filter behavior for SAMN and PNB, the average number of reliable spikes was on average higher and temporal coding fraction was lower for SAMN. This suggests that SAMN tended to produce on average more phase-locked action potentials (higher number of reliable spikes) to the stimulus envelope compared with PNB, although these were accompanied by a higher level of undriven random firing (thus the lower temporal coding fraction). Furthermore, unlike for PNB where the spike-timing jitter was roughly constant with modulation frequency (Fig. 5E, red), spike-timing jitter decreased systematically with increasing modulation frequency for SAMN (Fig. 5E, black).

Overall, the systematic and highly consistent relationship between envelope shape and jitter and modulation frequency and firing reliability support the hypothesis that spike-timing jitter is strongly influenced by the envelope shape of the sound, whereas spiking reliability is much more affected by the sound period (see Discussion).

Specific Temporal Information

The highly consistent trends observed in the correlogram analysis demonstrate that spike-timing precision and reliability can potentially encode the shape and repetition of acoustic envelopes, respectively. This raises several questions regarding the structure of the neural code in the CNIC. First, it is of interest to determine how much information is carried by the temporal signature of the neural response for a specific stimulus. Second, as demonstrated, response variability in the form of spike-timing jitter and firing reliability vary systematically with the stimulus parameters (shape and modulation frequency), and these in turn affect the information transfer. Thus it would be of interest to determine how the sound periodicity and envelope shape contribute to the temporal information transfer of the neuron.

The procedure for measuring the STI from neural responses is illustrated in Fig. 6 (see Methods). The STI was designed to quantify the available information from the periodic temporal response pattern to a “specific” sound. It is estimated with a model-based approach in which the neural response were fitted to the model (as in Fig. 2) and synthetic spike trains were generated with large number of response cycles to minimize measurement bias (8,000 simulated cycles; see Methods). To examine the effects of spike-timing resolution on STI, we varied the effective spike train resolution by synthetically adding spike-timing jitter to the model spike trains (1–200 ms in 0.38 octave steps) (Kayser et al. 2010; Lu and Wang 2004). Cycle dot-rasters from a single neuron are shown for PNB (Fig. 6A) and SAMN (Fig. 6B) at 10-Hz modulation frequency, along with model spike trains with varying amounts of added spike-timing jitter (Fig. 6, A and B: 0, 5, 50, 100 ms). For both sounds (PNB and SAMN at 10 Hz), the total spike train entropy (specific entropy) was relatively constant for multiple analysis spike train resolutions (Fig. 6, C and D, blue curve). However, the specific noise entropy increases with decreasing spike train resolution (Fig. 6, C and D, red curve) because the added spike-timing jitter adds spike-timing variability to the neural spike train. The STI is obtained by subtracting the specific noise entropy from the specific total entropy and correcting for bias in the measurement (Fig. 6, C and D, black curve; see Methods). Note that, when the analysis spike train resolution roughly corresponds to or exceeds one period of the stimulus (i.e., 100 ms for this example), the specific entropy and specific noise entropy are equal and the STI is precisely zero. This occurs because, when the added spike-timing jitter is equal to or greater than the cycle period, the spike-timing patterns are completely randomized, and there is no consistent firing pattern across stimulus trials.

Several distinct trends were observed in the neural activity that reflected the SAMN and PNB stimulus structure. First, the measured STI on a per-cycle basis (bits/cycle) decreases systematically with increasing modulation frequency for both PNB and SAMN. This behavior is seen for individual examples (Fig. 7, A and B, left panels), as well as the neural population (Fig. 7C, left panel). Furthermore, for the examples and population, the magnitude of the STI is larger for PNB (bootstrap t-test with Bonferroni correction, $P < 0.05$ for 5–134 Hz), consistent with the fact that temporal responses are more precise for PNB (Fig. 5E) and thus can convey more stimulus information through precise spike timing. The decreasing trend with modulation frequency is largely attributed to the fact that temporal reliability decreases with increasing modulation frequency for both PNB and SAMN (Fig. 5C). By comparison, when we measured the available STI per unit time (i.e., the STI rate; bits/s), a distinctly different picture is obtained. For both PNB and SAMN, the temporal information that is conveyed on average could be tuned for individual neurons (Fig. 7, A and B, right panel) and tended to be larger in magnitude for PNB compared with SAMN. Likewise, for the neural population, the STI rate was tuned as a function of modulation frequency (Fig. 7C, right panel) and larger for PNB (bootstrap t-test with Bonferroni correction, $P < 0.05$ from 5–134 Hz). The STI rate peaked at 97-Hz modulation frequency for
STI rate pattern for PNB and SAMN reflect differences in cycle, and, in turn, temporal information can be conveyed by the transient PNB envelopes leads to higher information per cycle, particularly true for PNB, where the precise spike timing to the STI conveyed by one cycle is relatively small. This is factors separately influence the coding efficiency. To do this, we consider the fraction of the total stimulus information that could be attributed to reliability and spike-timing jitter and reliability errors contribute to the temporal noise entropy. The difference between the specific noise entropy (red) and the reliability noise entropy (dotted black) are shown as a function of the effective spike train resolution (i.e., added jitter; see METHODS). Note that the total specific entropy is relatively independent of the amount of added jitter, while the specific noise entropy increases with the amount of added jitter. The coding efficiency of a neuron is defined as the fraction of the total spike train entropy that contains stimulus-related information per cycle decreases with modulation frequency, as a consequence of unreliable firing, substantial stimulus information can be accumulated over time. This occurs because, although the STI per cycles is high at low-modulation frequencies, there are relatively few cycles for which to accumulate temporal information (e.g., 5 cycles over 1 s at 5 Hz). In contrast, substantial temporal information can be gained for high-modulation frequencies by observing responses over time, even though the STI conveyed by one cycle is relatively small. This is particularly true for PNB, where the precise spike timing to the transient PNB envelopes leads to higher information per cycle, and, in turn, temporal information can be conveyed up to ~200 Hz. As proposed below, such differences in the STI rate pattern for PNB and SAMN reflect differences in the neural response that are associated with the sound envelope shape.

The Contribution of Spike-timing Jitter and Reliability on Temporal Coding Efficiency

As demonstrated, the period and envelope shape of SAMN and PNB have a pronounced effect on the temporal precision and reliability of neural responses. These systematic differences ultimately influence the temporal information transfer of CNIC neurons. Here we examine how these acoustic (shape and periodicity) and neuronal response (jitter and reliability) factors separately influence the coding efficiency.

The coding efficiency of a neuron is defined as the fraction of the total spike train entropy that contains stimulus-related information (Rieke et al. 1995). Noting that spiking noise limits the neural response that are associated with the sound envelope shape.
errors (see methods). The temporal ($F_{\text{temporal}}$) and reliability ($F_{\text{reliability}}$) noise fractions were defined as the fraction of the total entropy that was attributed to spike-timing (jitter) or spiking reliability errors. The temporal coding efficiency is then $E = 1 - F_{\text{reliability}} - F_{\text{temporal}}$ (see methods and appendix d). Figure 8, A and B (left panels), shows the temporal coding efficiency trends from two single neurons. For both PNB and SAMN, the temporal coding efficiency decreases systematically with the sound modulation frequency, although it tends to be somewhat higher for the PNB sound.

For each unit, we decomposed the temporal coding efficiency into the reliability and temporal noise fractions and plotted these as a function modulation frequency (Fig. 8, A and B, middle and right panels). The reliability and temporal noise fractions represent the proportion of the total entropy that can be accounted by reliability and spike-timing errors, respectively (see methods and appendix d). Two distinct trends are observed. First, the reliability noise fraction for the select neurons exhibited similar trends for SAMN and PNB, where it tends to increase systematically with increasing modulation frequency (Fig. 8, A and B, middle panels). Thus the response entropy associated with reliability noise is lowest at low-modulation frequencies and tends to dominate at high-modulation frequencies. This is consistent with the measured reliability trends in which firing reliability decreases with modulation frequency (Fig. 5C). By comparison, the temporal noise fraction tends to decrease with modulation frequency and is lower in magnitude for PNB. The lower noise fraction observed for PNB reflects the fact that spike-timing noise is lower for PNB as a result of the higher spike-timing precision (Fig. 5E) and the proportion of temporally phase-locked spikes is higher for PNB (Fig. 5D).

Similar behavior is seen in the population trends (Fig. 8C). As for the example single neurons, the temporal coding efficiency decreases with increasing modulation frequency, and the overall magnitude was higher for PNB compared with SAMN (bootstrap t-test with Bonferroni correction, $P < 0.05$ from 5–259 Hz), indicating a higher signal-to-noise ratio for PNB. Interestingly, the reliability noise fraction trend was identical for PNB and SAMN (Fig. 8C, middle panel, bootstrap
Bonferroni correction, nonsignificant), indicating that an equal proportion of the total response entropy could be accounted for reliability errors for both SAM and PNB. The reliability noise fraction increases systematically with increasing modulation frequency, reflecting the fact that firing reliability decreases with increasing modulation frequency (Fig. 5C). The similarity in the reliability trend for SAMN and PNB is consistent with the fact that spiking reliability depends primarily on the sound repetition, which is identical for both sounds. By comparison, the temporal noise fraction decreases with modulation frequency and is lower for PNB (Fig. 8, right panel) (bootstrap t-test with Bonferroni correction, \( P > 0.05 \) from 5–360 Hz). This is consistent with the observation that spike-timing jitter and hence the amount of noise attributed to spike-timing errors is smaller for PNB compared with SAMN (Fig. 5E). Together, these findings indicate that the reduced temporal coding efficiency for SAMN (relative to PNB) is largely attributed to larger proportion of spike-timing noise for SAMN.

Since the primary acoustic difference between PNB and SAMN is the envelope shape, we hypothesize that observed disparity in temporal coding efficiency can be attributed to spike-timing differences arising from differences in the envelope shape. As demonstrated, the fractional response entropy attributed to spike-timing errors (temporal noise fraction, right panels), by comparison, decreases with modulation frequency and is larger in magnitude for SAMN. C: similar behavior is observed for the average population results (error bars indicate SEM).

Fig. 8. Relationship between temporal coding efficiency, temporal noise fraction, and reliability noise fractions. A and B, left panel: the temporal coding efficiency is shown for SAMN and PNB for two neurons (same neurons as in Fig. 7). For both examples, temporal coding efficiency decreases systematically with increasing modulation frequency for both SAMN and PNB, although the efficiency is higher for PNB. A and B: in contrast, the proportion of noise due to reliability errors (reliability noise fraction, middle panels) increases systematically with modulation frequency and is roughly equal in magnitude for SAMN and PNB. A and B: the proportion of noise attributed to spike-timing errors (temporal noise fraction, right panels), by comparison, decreases with modulation frequency and is larger in magnitude for SAMN. C: similar behavior is observed for the average population results (error bars indicate SEM).
effective temporal resolution of the PNB and SAMN responses by adding synthetic spike-timing jitter (as in Fig. 6, A and B) [Kayser et al. 2010 (no. 7041); Lu and Wang 2004 (no. 2186)] prior to computing the STI and temporal coding efficiency. As seen from selected examples (Fig. 6) and all other neurons tested, systematically adding spike-timing jitter to the neural responses produces a systematic reduction in the STI. This behavior is evident in the population average results since the STI decreases systematically with increasing spike-timing jitter for all modulation frequencies (Fig. 9, A and B). When the added spike-timing jitter is equal to or exceeds the modulation period, the temporal pattern of the response is completely randomized, and, consequently, no temporal information is available (the STI zero).

Given that PNB responses exhibit higher efficiency than SAMN and that reliability noise fraction is identical for both sounds (Fig. 8C), we asked whether the primary differences between PNB and SAMN responses could be accounted by differences in spike-timing resolution. To test for this possibility, we measured the equivalent PNB spike-timing resolution that is required to achieve comparable efficiency as SAMN (at each modulation frequency). That is, we added synthetic spike-timing jitter to the PNB responses until the efficiency of PNB at each modulation frequency was matched to the SAMN efficiency with zero-added spike-timing jitter (dashed black contour, Fig. 9C). The population average SAMN and PNB efficiencies are shown in Fig. 9, C and D, respectively, along with the matched efficiency contours for SAMN and PNB (Fig. 9, C and D; dashed black contours). As can be seen, adding spike-timing jitter to both PNB and SAMN responses reduces the temporal coding efficiency. Interestingly, the added PNB spike-timing jitter required to match the SAMN temporal coding efficiency with zero added jitter (dashed black contour) closely follows the modulation period (gray curve). On average, the added PNB spike-timing jitter required to match the SAMN efficiency corresponded to 31% of the modulation period (Fig. 9E). Thus PNB response can be synthetically modified to resemble SAMN responses by adding spike-timing jitter that is roughly a constant proportion of the modulation period. This result mirrors and is consistent with the fact that spike-timing jitter was small and independent of the modulation frequency for PNB while it covaried proportional to the time scale of envelope shape for SAMN (Fig. 5E).

**DISCUSSION**

Current theories of amplitude modulation coding aim to account for aspects of pitch and rhythm perception associated with the periodic structure of a sound. With few exceptions (Lu et al. 2001; Neuert et al. 2001; Plessnitzer et al. 2000; Zheng and Escabí 2008), most studies do not account for relevant loudness and timbre cues associated with the envelope shape of a sound (Irino and Patterson 1996). As demonstrated, the firing
reliability and spike-timing precision were strongly affected by the sound period and shape in a manner that reflected the sound structure. Spike-timing precision covaried with the sound envelope shape, while the firing reliability was strongly dependent on, and decreased systematically with, increasing modulation frequency. These trends were closely tied and influenced the temporal information transfer and temporal coding efficiency. These results thus outline an encoding strategy in which the temporal firing pattern of CNIC neurons could simultaneously and efficiently represent periodicity and shape temporal cues.

The Role of Spike-timing Precision and Firing Reliability

The proposed framework employs the concept of spike-timing precision and firing reliability for studying the neural representation of shape and periodicity. Because firing reliability and spike-timing precision are essentially separate forms of variability in the temporal response pattern, our results demonstrate that neural “noise” or variability are inherently dependent on the sound structure. This brings up the possibility that spiking noise itself is an integral component of the neural code (Stein et al. 2005). Specifically, the observed dependence between temporal acoustic structure, reliability, and spike-timing noise could provide a means of signaling shape and periodicity-based cues.

Although firing rates were not systematically related to the sound shape or period, both jitter and reliability exhibited highly consistent response patterns from neuron to neuron, thus providing a robust signal to encode the acoustic envelope structure. For both SAMN and PNB, firing reliability was primarily affected by the sound period, while firing precision was strongly affected by the envelope shape. Firing reliability and the fraction of temporally locked action potentials decreased systematically with increasing modulation frequency. The decreasing nature of these trends may arise because of cell membrane and synaptic low-pass filtering which would limit temporal synchronization to all of the stimulus cycles with increasing modulation frequency (Geis and Borst 2009). Functionally, this trend could be related to the fact that pitch saliency degrades with increasing modulation frequency (Burns and Viermeister 1976). By comparison, firing precision was roughly constant for PNB at all modulation frequencies, although it varied systematically with the changing shape of the sound for SAMN.

The fact that central auditory neurons can respond with remarkable high temporal precision seems to be at odds with the fact that they cannot follow fast temporal modulations. Even within the auditory cortex, neurons can exhibit millisecond temporal firing precision to transient sound events, yet their ability to follow temporal modulations breaks down at ~20 Hz (Elhilali et al. 2004; Heil and Irvine 1997; Joris et al. 2004). A similar dichotomy is evident in our data since neurons did not effectively follow temporal modulations beyond 300 Hz, although they could phase-lock with microsecond precision for PNB. We propose that this apparent dichotomy is related to separate encoding mechanisms. As demonstrated, spike-timing precision is strongly affected by the envelope shape of the sound. In contrast, the limiting ability to follow fast temporal modulations is related to the observed reduction in firing reliability with increasing modulation frequency. Thus our results may provide an explanation for this apparent paradoxical behavior on the basis of distinct response attributes and acoustic cues.

The proposed framework offers several advantages over other conventional approaches used to study temporal coding. Traditionally, the VS is used to quantify the precision of firing (Goldberg and Brown 1969). Despite its utility, the VS does not distinguish reliability from timing errors in the response pattern and strictly quantifies relative spike-timing precision (i.e., the proportional width of the cycle histogram). By comparison, the proposed jitter metric measures the absolute precision of firing (in ms). The correlation index proposed by Joris and colleagues (2006) has some appeal over VS; however, this metric depends both on reliability and precision of the responses. Given that reliability and precision are uniquely different forms of neural variability and, as shown here, can vary independently for different sound cues, the proposed reliability and spike-timing metrics should be of high appeal, because they can decompose and independently quantify these forms of neural variability. Furthermore, although not demonstrated here, the proposed framework and shuffled correlogram approach for measuring reliability and precision can be extended to the more general case of aperiodic signals (Chen and Escabi 2010).

The role of firing reliability and precision has been shown to be relevant in auditory structures outside the midbrain. In the AN, firing reliability and spike-timing precision decrease with increasing characteristic frequency for low-frequency tones (Avisser et al. 2007), analogous to our results for SAMN. Similar trends for reliable firing have also been described in the auditory cortex of awake primates (Malone et al. 2007). Collectively, these studies support the view that firing reliability and spike-timing precision could provide an alternative and more general framework to study the neural representation of temporal sound information.

Proportional vs. Absolute Spike-timing Precision

As demonstrated, CNIC neurons can generate spikes with extraordinary temporal precision (on the order of hundreds of microseconds), assuming that the sound structure is itself sufficiently precise. This result may seem somewhat trivial, given that the first-spike timing precision of neurons in the auditory system has been shown to be contingent on the shape of the rising envelope of a sound (Heil 1997b, 2001; Heil and Neubauer 2003). However, these previous studies strictly examined the first-spike to brief transient envelopes, which are functionally different from the subsequent spikes (Zheng and Escabi 2008). By comparison, we have examined the temporal structure of the neural response when both periodicity and shape cues are present and consider all of the action potentials in the neural response (not simply the first spike).

Based on our analysis, firing reliability and precision appear to be independently affected by the sound period and envelope shape, respectively. The findings suggest that the basic relationships between envelope shape and timing precision previously observed for ramped waveforms (Heil and Irvine 1997) generalizes for the case of periodic sounds. Furthermore, spike-timing precision is not strongly dependent on the sound modulation frequencies. This can be inferred from the fact that, for PNB, fine temporal precision resulted from the sub-millisecond
structure of the stimulus events (250-μs events of PNB) and was not dependent on the sound modulation frequency. By comparison, spike-timing precision for SAMN varied with the envelope shape from several milliseconds to tens of milliseconds (Fig. 5E). Surprisingly, although the firing reliability of the first-spike improves for sharper rising envelopes (Heil 1997a, 1997b), this was not the case for our periodic sounds, since reliability of firing was dominated by the modulation frequency for PNB and SAMN. Since PNB and SAMN differ only on the basis of the envelope shape (they cover identical modulation frequencies), it can be concluded that the spike-timing precision and reliability were largely contingent on the shape and period, respectively.

The role of proportional spike-timing precision likely extends beyond auditory modality, as visual neurons can exploit similar rules for encoding natural visual scenes (Butts et al. 2007). The temporal firing precision of visual cortex neurons adapts to the temporal resolution of natural visual inputs in a manner that optimizes the encoding strategy. Visual and auditory receptive field studies have demonstrated that changes in neural sensitivity are strongly dependent on the statistics of relevant natural stimuli (Escabi et al. 2003; Hosoya et al. 2005; Lesica et al. 2007; Nagel and Doupe 2006; Theunissen et al. 2000). Although our result is based on responses to artificial stimuli, it provides compelling evidence that temporal firing statistics systematically change and are contingent on the temporal resolution of the sound.

**Temporal Information Transmission, Coding Precision, and Efficiency**

High levels of information transfer and efficiency have been reported previously for peripheral auditory neurons, with efficiency approaching 50% (Rieke et al. 1995). In an analogous manner, neurons in the CNIC could achieve high efficiency and information transfer. However, the total STI and efficiency largely depended on the stimulus condition and the response spike-timing precision. For PNB, the coding efficiency could exceed 40% as a result of the high spike-timing precision, while SAMN responses generally had larger spike-timing errors and thus lower efficiency (~20% maximum).

The information theoretic framework allowed us to identify how shape and periodicity cues contribute to the information that can be conveyed by the firing reliability and spike-timing precision. Ultimately, reliability and precision are critical factors that limit the transmitted temporal information and which affect the coding efficiency. As demonstrated, the entropy associated with spike timing and reliability noise systematically covaried with the modulation frequency and shape of the sound, respectively. In particular, the proportion of noise entropy attributed to reliability errors (reliability noise fraction) was essentially identical for SAMN and PNB, thus reflecting the common periodicity for both sounds. By comparison, the proportion of entropy attributed to spike-timing errors (temporal noise fraction) was larger for SAMN, consistent with the larger spike-timing errors observed for SAMN responses (compared with PNB). By adding synthetic spike-timing jitter to PNB responses, the jitter required to match the SAMN efficiency was nearly a constant proportion of the modulation period. This behavior is consistent with the observation that spike-timing jitter covaries proportional to the sound envelope shape for SAMN (Fig. 5E).

The available temporal information in neural responses was dependent on both the envelope shape and modulation period. Unlike the STI, which decreased systematically with increasing modulation frequency, the STI rate (i.e., total information per unit time) is tuned with modulation frequency for both SAMN and PNB. For SAMN, the STI rate peaks at 26 Hz, while, surprisingly, the maximal STI rate shifted toward modulation frequencies between 186 and 259 Hz for PNB (Fig. 7). This demonstrates that a substantial amount of temporal information can be accumulated over time at these relatively high-modulation frequencies, even though the amount of information per cycle is relatively low. Furthermore, it demonstrates that responses for sounds with transient envelopes (such as PNB) can convey substantially more information for high-modulation frequencies as a consequence of precise spike timing to the sharp envelope shape.

Our goal of measuring the stimulus information in the periodic temporal response pattern is motivated by the observation that envelope shape and periodicity are perceptually relevant acoustic cues (Akeroyd and Patterson 1995; Burns and Viemeister 1981; Irino and Patterson 1996; Viemeister 1979) that are reflected in the temporal response of inferior colliculus neurons (Krebs et al. 2008; Zheng and Escabi 2008). A limiting factor of our information theoretic approach is that it requires knowledge of the sound period a priori. This brings up the question of how the brain would extract such periodic information from the neural responses. Although our approach does not directly address this question, several schemes have been proposed that could extract information from the neural response to periodic signals, such as a central bank of tuned filters or cross-correlation networks (Dau et al. 1997; Shamma and Klein 2000). Future studies will need to consider how and to what extent central networks of neurons utilize this reliable and precise spike-timing information to produce a meaningful interpretation of the sound.

**Summary**

Our findings provide evidence for a proportional spike-timing and reliability code in the CNIC, where the reliability of firing and spike-timing precision covaried with the sound period and envelope shape, respectively. As demonstrated, the envelope shape and period of a sound greatly impact the spike-timing precision, firing reliability, and, consequently, the available temporal information that can be theoretically transmitted by CNIC neurons. These response attributes alter the information transfer of the stimulus and provide a means of signaling shape and repetition based temporal cues.

**APPENDIX A: SHUFFLED CORRELOGRAM MEASURE OF SPIKE-TIMING JITTER AND RELIABILITY**

Consider a phenomenological spiking model in which a hypothetical noiseless neuron produces a phase-locked response to a periodic stimulus of frequency $f_m$. The hypothetical neuron produces a perfectly phase-locked response with precisely $M$ spike per stimulus period $T = 1/f_m$.
\[ s_{\text{noiseless}}(t) = \sum_{i} \sum_{m=1}^{M} \delta(t - i \cdot T) \quad (A1) \]

where the Dirac delta function, \( \delta(t) \), represents a single spike. The index \( i \) sums over the stimulus periods, and \( m \) sums over the \( M \) spikes per stimulus period. The reference hypothetical neuron has 100% firing reliability and infinite precision and thus a VS of 1. Clearly, the synchronized responses of real neurons will never conform to this highly idealized response. Thus, to create a realistic stochastic model that accurately replicates the forms of variability observed for spiking neurons in vivo, we introduce three forms of noise into the spiking model: spike-timing jitter, spiking count reliability errors, and additive spiking noise (baseline firing rate). The stochastic model produces a quasi-periodic response pattern with a variable number of synchronized and unsynchronized spikes per period of the form:

\[ s_i(t) = \sum_{m=1}^{M} \delta(t - i \cdot T - \Delta t_{m,i,k}) \cdot s_{m,i,k} + \sum_{j} \delta(t - t_{j,k}) \quad (A2) \]

Here the index \( k \) designates the response to the \( k \)-th stimulus trial. The index \( i \) sums over each of the stimulus periods and is necessary to produce a quasi-periodic spiking pattern. Two forms of noise are introduced into this quasi-periodic term. First a Bernoulli random variable \( s_{m,i,k} \) models the firing reliability of the response. This variable assumes a value of 1 with the hypothetical neuron which produces \( M \) spikes per period, the model neuron produces an average of \( x = M \cdot p \) reliable phase-locked spikes per period. Second, spike-timing jitter is introduced by incorporating spike-timing errors \( \Delta t_{m,i,k} \) to the original noiseless action potentials (\( m \)-th action potential, from the \( i \)-th period and \( k \)-th trial). We assume that the spike-timing errors follow an arbitrary distribution, \( f(\Delta t) \). In the second sum of Eq. A2, the index \( j \) sums over individual spikes for the additive baseline noise component of the response \( s_{\text{noise}}(t) \). We assume that the noise spike times, \( t_j \) (\( j \)-th spike noise from the \( k \)-th stimulus trial), are Poisson distributed with spike rate \( \lambda_{\text{noise}} \). Finally, we assume that the all of the noise terms (jitter, reliability, and baseline noise) are independent but identically distributed across trials, so that the responses from separate trials are independent and share identical statistics.

We now derive the expected shuffled correlogram of the proposed spiking model, and we will subsequently use the closed form solution to characterize neuronal responses to periodic stimuli in the inferior colliculus. The expected shuffled correlogram between the \( k \)-th and \( l \)-th response trials is obtained by performing an average across time (\( \langle \cdot \rangle \) and taking an expectation \( (E[\cdot]) \) with respect to each of the random variables:

\[ \Phi_{\text{shuffled}}(\tau) = \left\langle \mathcal{E} [s_i(t) \cdot s_l(t + \tau)] \right\rangle = \sum_{i} \sum_{l} \mathcal{E} \left[ \mathcal{E} [s_{\text{period}}(i,t) \cdot s_{\text{period}}(l,t + \tau)] \right] + \sum_{i} \mathcal{E} \left[ \mathcal{E} [s_{\text{period}}(i,t) \cdot s_{\text{noise}}(l,t + \tau)] \right] + \mathcal{E} \left[ \mathcal{E} [s_{\text{noise}}(i,t) \cdot s_{\text{noise}}(l,t + \tau)] \right] \quad (A3) \]

Since the noise components for the \( k \)-th and \( l \)-th trials are independent and likewise the noise is independent of the response we have that:

\[ \Phi_{\text{shuffled}}(\tau) = \sum_{i} \sum_{l} \mathcal{E} \left[ s_{\text{period}}(i,t) \cdot s_{\text{period}}(l,t + \tau) \right] + 2 \cdot \lambda_{\text{period}} \cdot \lambda_{\text{noise}} + \lambda_{\text{noise}}^2 \quad (A4) \]

To evaluate the first term of Eq. A4, we combine the first term of Eq. A2, assume independence of the firing reliability and spike-timing jitter, and use the independence to reorder the expectation and time-average operators as follows:

where \( \lambda_{\text{period}} \) is the firing reliability of the response. This variable assumes the form of the expected shuffled correlogram of the proposed spiking model:

\[ \mathcal{E} \left[ s_{\text{period}}(i,t) \cdot s_{\text{period}}(l,t + \tau) \right] = \sum_{i} \sum_{l} \mathcal{E} \left[ s_{\text{period}}(i,t) \cdot s_{\text{period}}(l,t + \tau) \right] \]

Next we note that:

\[ \mathcal{E} [s_{\text{period}}(i,t) \cdot s_{\text{period}}(l,t + \tau)] = p \quad (A6.1) \]

\[ \mathcal{E} [\delta(t - i \cdot T - \gamma)f(\gamma)d\gamma = f(t - i \cdot T) \quad (A6.2) \]

\[ \mathcal{E} [\delta(t + \tau - i \cdot T - \gamma)f(\gamma)d\gamma = f(t + \tau - i \cdot T) \quad (A6.3) \]

so that Eq. A5 simplifies:

\[ \langle M \cdot p \rangle = \left\langle \sum_{i} f(t - i \cdot T) \cdot \sum_{l} f(t + \tau - i \cdot T) \right\rangle \quad (A7) \]

Expanding the time-average operator and assuming infinite spike train duration in Eq. A7 yields

\[ \langle M \cdot p \rangle = \sum_{i} \sum_{l} \lim_{D \to \infty} \frac{1}{D} \int f(t - i \cdot T) \cdot f(t + \tau - i \cdot T)dt \]

\[ = \sum_{i} \sum_{l} \lim_{D \to \infty} \frac{1}{D} \cdot \Phi_{f}(\tau + (i - \tau)) \]

where \( T_D \) is the spike train duration, and \( \Phi_{f}(\tau) \) is the autocorrelation function of the spike-timing jitter distribution, \( f(\Delta t) \). Making a change of variables \( n = i - \tau \) allows us to collapse one of the sums:

\[ = \sum_{i} \sum_{n} \frac{N}{n} \cdot \Phi_{f}(\tau - n \cdot T) \]

\[ = \sum_{i} \sum_{n} f_m \cdot \Phi_{f}(\tau - n \cdot T) \]

where \( N \) is the number of stimulus periods over \( T_D \) seconds. To achieve convergence, we have assumed that averaging is performed over an infinite amount of time so that \( N/T_D \) converges to the modulation frequency, \( f_m \). Finally, combining Eq. A4 with Eq. A9 leads to the desired solution:

\[ \Phi_{\text{shuffled}}(\tau) = \sum_{i} \sum_{n} \Phi_{f}(\tau - n \cdot T) + 2 \cdot \lambda_{\text{period}} \cdot \lambda_{\text{noise}} + \lambda_{\text{noise}}^2 \quad (A10) \]

We point out that, in Eq. A10, the number of reliable spikes per cycle, \( \bar{x} \), is precisely equal to the firing reliability, \( p \), for the case where the hypothetical neuron produces 1 spike per cycle.

**APPENDIX B: VALIDATION USING A CORRECTED ERROR METRIC**

Here we devise two error metrics that are used to quantify the quality of the shuffled correlogram model (derived APPENDIX A) relative to the neural data. Our goal is to identify strictly deterministic errors in the model without the influence of measurement noise resulting from data limitations (i.e., insufficient averaging, noisy responses such as weak phase locking, etc.).

We start by assuming that the measured correlogram for each half of the data consists of an ideal deterministic correlogram \( \Phi(t) \) (expected for infinite data averaging) that has been corrupted by measurement noise \( n_i(t) \) for the validation half and \( n_j(t) \) for the
time and response variability. To compensate for this noise dependent on the model error, we estimated the noise power in the data as

\[
\text{var}(\varepsilon(t)) = \text{var}(\Phi(t) - \Phi_{\text{model}}(t)) + \text{var}(n(t)) = \sigma_n^2 + \sigma_{\varepsilon}^2
\]

In an analogous manner, the noise power in the empirically derived correlogram can be estimated by noting that the measurement noise is independent for each data segment so that

\[
\text{var}(\varepsilon(t)) = \text{var}(\Phi(t) - \Phi_{\text{model}}(t)) = \text{var}(n(t)) = \sigma_n^2 + \sigma_{\varepsilon}^2
\]

Therefore, we estimated the noise power in the data as

\[
\sigma_n^2 = \frac{1}{2} \text{var}(\Phi_{t}(t) - \Phi_{\text{model}}(t))
\]

Finally, we note that the power of the validation data contains both a signal and a noise component.

\[
\text{var}(\Phi(t)) = \text{var}(\Phi(t) + n(t)) = \sigma_n^2 + \sigma_{\varepsilon}^2
\]

By considering the power of the measured shuffled correlogram (Eq. B7), difference (Eq. B4), and noise signals (Eq. B6), two error metrics were devised. First, the model error was expressed as the power of the model difference signal

\[
\text{var}(\Phi_{t}(t) - \Phi_{\text{model}}(t)) = \sigma_n^2 + \sigma_{\varepsilon}^2
\]

Note that this metric essentially removes the noise power from both the numerator and denominator of Eq. B8 and, as desired, captures the fractional power of the error signal.

### APPENDIX C: RELATIONSHIP BETWEEN STI, ENSEMBLE INFORMATION, AND MUTUAL INFORMATION

In this appendix, we show the mutual information for an entire stimulus-response ensemble is a weighted sum of the STI and the EI. The STI quantifies the contribution to the MI that is provided by the temporal response pattern to a “specific stimulus”, while the EI accounts for the information that can be gained by comparing responses across the stimulus-response ensembles.

The proposed framework is motivated by previous work, which attempted to define the information associated with one stimulus (Butts 2003; DeWeese and Meister 1999). Here we have a similar goal; however, our stimuli and responses are conditionally independent, and we will include this constraint on our formulation.

That is, for the entire stimulus and response ensemble, S and W can be partitioned into independent stimulus-responses subspaces so that

\[
S = S_1 \cup S_2 \ldots S_{l-1} \cup S_l \quad \text{and} \quad W = W_1 \cup W_2 \ldots W_{l-1} \cup W_l.
\]

Here \(S_k, (k = 1 \ldots L)\) represents the ensemble of periodic features that comprise the \(k\)-th stimulus (i.e., the specific stimulus), and \(W_k\) corresponds to the ensemble of responses for \(S_k\). The neural responses subspaces, \(W_k\), consist of the ensemble of response words measured for the \(k\)-th stimulus. This partitioning of the stimulus and response space is chosen so that the acoustic features \(s\) and response words \(w\) are exactly one period duration and span all possible phases of the sound. Conditional independence of the stimulus-response ensembles requires that \(S_k \cap S_l = \emptyset\) and \(W_k \cap W_l = \emptyset\), for \(k \neq l\), and that the responses from \(S_k\) are exclusively contained within \(W_k\); therefore, \(p(S_k) = p(W_k) = p(S_k, W_k)\).

We seek to quantify the information associated with the temporal response pattern for each stimulus. We start by noting that the MI between the stimulus and response ensemble is expressed as:

\[
I(W, S) = \sum_{w \in W} \sum_{s \in S} p(w, s) \log_2 \frac{p(w, s)}{p(w)p(s)}
\]

where \(s\) and \(w\) correspond to the stimulus and responses, respectively. Given conditional independence of the stimuli and responses, we rearrange Eq. C1 so that we sum across each of the specific stimulus \(s \in S_k\) and response \(w \in W_k\) ensembles:

\[
I(W, S) = \sum_{k} \sum_{w \in W_k} \sum_{s \in S_k} p(w, s) \log_2 \frac{p(w, s)}{p(w)p(s)}
\]

Next we apply the law of total probability by conditioning on the specific stimulus and response ensembles and note that whenever \(w \in W_k\) and \(s \in S_k\), as in the inner two sums of Eq. C2, we have that

\[
p(w, s) = \sum_{l} p(w, s|W_l)p(W_l) = p(w|W_k)p(W_k)
\]

\[
p(s) = \sum_{l} p(s|S_l)p(S_l) = p(s|S_k)p(S_k)
\]

The mutual information can thus be expressed as

\[
J Neurophysiol · doi:10.1152/jn.01080.2010 · www.jn.org
Simplifies to

\[ I(W, S) = \sum_{w} \sum_{S_k} p(w, S_k) \cdot \sum_{r \in S_k} p(w, r | S_k) \cdot \log_2 \frac{p(w, S_k)}{p(w)} \]

Next we note that, in the second sum of Eq. C4

\[ \sum_{S_k} \sum_{w} p(w, S_k) = 1 \]  
(C5)

Rearranging terms and incorporating Eq. C5 the mutual information simplifies to

\[ I(W, S) = \sum_{w} \sum_{S_k} p(w, S_k) \cdot \log_2 \frac{p(w, S_k)}{p(w)} + \sum_{w} \sum_{S_k} p(w, S_k) \cdot \log_2 \frac{p(w)}{p(S_k)} \]

We next define the STI

\[ I_{STI}(W_k, S_k) = \sum_{w} \sum_{r \in S_k} p(w, r | S_k) \cdot \log_2 \frac{p(w, r | S_k)}{p(w)} \]  
(C7)

and the EI

\[ I_{EI} = \sum_{w} \sum_{S_k} p(w, S_k) \cdot \log_2 \frac{p(w)}{p(S_k)} \]  
(C8)

so that the total mutual information is expressed as:

\[ I(W, S) = \sum_{w} \sum_{S_k} p(w, S_k) \cdot I_{STI}(W_k, S_k) + I_{EI} \]  
(C9)

Thus, under the imposed conditional independence stimulus-response constraint, the mutual information is a weighted sum of the information associated with the temporal response pattern to a specific stimulus (STI) and the information gained by comparing the responses across the conditionally independent stimulus-response ensembles (EI).

**APPENDIX D: TEMPORAL AND RELIABILITY NOISE ENTROPY**

Here we demonstrate that the response noise entropy (Eq. 12) can be decomposed into separate components that strictly arise from spike-timing and firing reliability errors: \( H_{\text{Noise}} = H_{\text{Reliability}} + H_{\text{Temporal}} \) for each stimulus-response ensemble. To facilitate the derivation, we assume that the specific stimulus-response ensemble is implied (\( S_k, W_k \)), and we thus employ a simplified notation \( p_{\text{Noise}}(w) \) in place of the noise distribution for each specific stimulus [i.e., \( p_l(s | w \theta) \) in Eq. 12].

We start by partitioning the noise entropy into disjoint response subsets, each of which contains precisely \( l \) spikes per word

\[ H_{\text{Noise}} = -\langle \sum_{w} p_{\text{Noise}}(w) \log \frac{p_{\text{Noise}}(w)}{p_{\text{Noise}}(w)} \rangle \]

Here, \( w \in L \) is the response subset containing all of the words with precisely \( l \) spikes and the average operator, \( \langle \rangle \), is taken with respect to the cycle phase as in Eq. 12. When \( w \in L \), \( p_{\text{Noise}}(w) = \sum_{n} p_{\text{Noise}}(w | n) p(n) = p_{\text{Noise}}(w | l) p(l) \), where \( n (n = 1, 2, 3, \ldots) \) is the possible number of spikes in a word, and \( p(l) \) is the probability of observing a word with precisely \( l \) spikes (i.e., the spike count distribution). Combining with Eq. CI and rearranging terms yields

\[ H_{\text{Noise}} = -\langle \sum_{w} \sum_{l} p_{\text{Noise}}(w | l) p(l) \log \frac{p_{\text{Noise}}(w | l) p(l)}{p_{\text{Noise}}(w | l)} \rangle \]

\[ = -\langle \sum_{l} \sum_{w} p_{\text{Noise}}(w | l) p(l) \log \frac{p_{\text{Noise}}(w | l) p(l)}{p_{\text{Noise}}(w)} \rangle \]

\[ - \langle \sum_{w} p_{\text{Noise}}(w) \log \frac{p_{\text{Noise}}(w)}{p_{\text{Noise}}(w)} \rangle \]

(D2)

Next, we note that \( \sum_{w} p_{\text{Noise}}(w) = 1 \) in the second sum since the sum takes place over the entire subspace containing all words with precisely \( l \) spikes. Combining and rearranging terms for the first sum yields

\[ H_{\text{Noise}} = -\langle \sum_{l} \sum_{w} p_{\text{Noise}}(w | l) p(l) \log \frac{p_{\text{Noise}}(w | l) p(l)}{p_{\text{Noise}}(w)} \rangle \]

\[ - \langle \sum_{w} p_{\text{Noise}}(w) \log \frac{p_{\text{Noise}}(w)}{p_{\text{Noise}}(w)} \rangle \]

(D3)

From Eq. D3, we define

\[ H_{\text{temporal}}(l) = -\langle \sum_{w} p_{\text{Noise}}(w | l) p(l) \log \frac{p_{\text{Noise}}(w | l) p(l)}{p_{\text{Noise}}(w)} \rangle \]  
(D4)

as the noise entropy when there is precisely \( l \) spikes per word. Because there is an exact number of spikes (\( l \)) for all of the words used to compute the entropy in Eq. D4, this entropy strictly arises from temporal variability in the response. When there are no spikes present (\( l = 0 \), \( p_{\text{Noise}}(w | l = 0) = 1 \), so that \( H_{\text{temporal}}(0) = 0 \) (i.e., because the zero word occurs with 100% certainty). Therefore, we can define the total temporal noise entropy by averaging across all possible spike counts

\[ H_{\text{temporal}} = \langle \sum_{l} p(l) \cdot H_{\text{temporal}}(l) \rangle \]  
(D5)

where \( p(l) \) is the spike count probability distribution. Equation D5 represents the expected entropy associated with spike-timing errors after averaging across all possible spike counts. Thus the temporal noise entropy strictly reflects temporal variability so that firing reliability errors do not contribute.

From Eq. D3, we next define the reliability noise entropy as

\[ H_{\text{Reliability}} = -\langle \sum_{l} p(l) \log \frac{p_{\text{Noise}}(w | l) p(l)}{p_{\text{Noise}}(w)} \rangle \]  
(D6)

which corresponds to the entropy in the spike count distribution. This component of the noise entropy reflects the variable number of spikes per stimulus cycle and thus can be seen as arising from spike-count reliability errors across response cycles (misfiring).

Finally, after combining the above two equations, we have that the noise entropy can be decomposed into a reliability and temporal component:

\[ H_{\text{Noise}} = H_{\text{Reliability}} + \langle \sum_{l} p_{\text{Noise}}(w | l) \cdot H_{\text{temporal}}(l) \rangle = H_{\text{Reliability}} + H_{\text{temporal}} \]  
(D7)

**ACKNOWLEDGMENTS**

We thank S. Kuwada and D. Kim for numerous discussions and H. L. Read for help with surgical procedures.

**GRANTS**

This work was supported by a grant from the National Institute of Deafness and Other Communications Disorders (DC-006397).

**DISCLOSURES**

No conflicts of interest, financial or otherwise, are declared by the author(s).
Ascending projections to the inferior colliculus.

REFERENCES


J Neurophysiol • doi:10.1152/jn.01080.2010 • www.jn.org