Cross-Correlation and Joint Spectro-Temporal Receptive Field Properties in Auditory Cortex

Masahiko Tomita and Jos J. Eggermont
Department of Physiology and Biophysics, Department of Psychology, University of Calgary, Calgary, Alberta, Canada

Submitted 25 June 2004; accepted in final form 25 August 2004

Tomita, Masahiko and Jos J. Eggermont. Cross-correlation and joint spectro-temporal receptive field properties in auditory cortex. J Neurophysiol 93: 378–392, 2005. First published September 1, 2004; doi:10.1152/jn.00643.2004. Recordings were made from the right primary auditory cortex in 17 adult cats using two eight-electrode arrays. We recorded the neural activity under spontaneous firing conditions and during random, multi-frequency stimulation, at 65 dB SPL, from the same units. Multiple single-unit (MSU) recordings (281) were stationary through 900 s of silence and during 900 s of stimulation. The cross-correlograms of 545 MSU pairs with peak lag times within 10 ms from zero lag time were analyzed. Stimulation reduced the correlation in background activity, and as a result, the signal-to-noise ratio of correlated activity in response to the stimulus was enhanced. Reconstructed spectro-temporal receptive fields (STRFs) for coincident spikes showed larger STRF overlaps, suggesting that coincident neural activity serves to sharpen the resolution in the spectro-temporal domain. The cross-correlation for spikes contributing to the STRF depended much stronger on the STRF overlap than the cross-correlation during either silence or for spikes that did not contribute to the STRF (OUT-STRF). Compared with that for firings during silence, the cross-correlation for the OUT-STRF spikes was much reduced despite the unchanged firing rate. This suggests that stimulation breaks up the large neural assembly that exists during long periods of silence into a stimulus related one and maybe several others. As a result, the OUT-STRF spikes of the unit pairs, now likely distributed across several assemblies, are less correlated than during silence. The cross-correlation for spikes contributing to the STRF depended much stronger on the STRF overlap than the cross-correlation during either silence or for spikes that did not contribute to the STRF (OUT-STRF). Thus the cross-correlation for spikes during silence was enhanced. Reconstructed spectro-temporal receptive fields (STRFs) for coincident spikes showed larger STRF overlaps, suggesting that coincident neural activity serves to sharpen the resolution in the spectro-temporal domain. The cross-correlation for spikes contributing to the STRF depended much stronger on the STRF overlap than the cross-correlation during either silence or for spikes that did not contribute to the STRF (OUT-STRF). Compared with that for firings during silence, the cross-correlation for the OUT-STRF spikes was much reduced despite the unchanged firing rate. This suggests that stimulation breaks up the large neural assembly that exists during long periods of silence into a stimulus related one and maybe several others. As a result, the OUT-STRF spikes of the unit pairs, now likely distributed across several assemblies, are less correlated than during silence. Thus the ongoing network activity is significantly different from that during stimulation and changes after the arousal during stimulation.

INTRODUCTION

Neurons in auditory cortex do not fire independently of each other; even those neurons that are separated by several millimeters show significant correlations in their firing times (Eggermont 1992). We have previously shown that cortical network activity, such as reflected in electroencephalographic (EEG) rhythms, contributes considerably to the overall correlation (Eggermont and Smith 1995, 1996) and can mask the effect of direct anatomical connections, e.g., from the thalamus, on the correlation of firing times. Generally, a cortical neuron receives only a small fraction of its input from specific thalamic afferents and a much larger fraction from other cortical neurons. Yet, analyzing stimulus-response correlations of cortical cells, such as the frequency-tuning curve (FTC) or spectro-temporal receptive field (STF), tacitly implies that this cell is a simple single input-single output system, i.e., does not allow inferences about the role the multiple inputs play. One reason that the single-input single-output concept works fairly well is that the firings of the specific thalamic afferents to the cortex are synchronized by the stimulus, whereas the majority of intra- and inter-cortical neural inputs arrive asynchronously (Abeles 1991). Modeling results (Mikula and Niebur 2003) suggest that the synchrony effect is optimal for cross-correlation coefficients between the afferent neurons of around 0.01, a value typically within the range observed in cortical neurons under spontaneous conditions (Eggermont 2000). Because the timing of the sensory stimulus that induces the activity in the specific afferents is known, the firings related to the specific afferent inputs can be extracted by stimulus response correlation. Thus extracting the stimulus-evoked spikes from the remaining, asynchronous background activity in cortical neurons could provide a way to assess the role of these two groups of firings on neural synchrony. Here, neural synchrony is defined as the correlation under stimulus conditions, neural correlation is defined as the correlation under spontaneous firing conditions or after correction for stimulus-related correlations (Eggermont 1994).

Role of background activity in shaping the efficacy of neural connections

Due to the subthreshold nature of excitatory postsynaptic potentials (EPSPs), input from a single thalamic cell will generally not suffice to make the cortical neuron fire unless the neuron is bursting (Salinas and Sejnowski 2000). However, when other coincident arrivals of spikes occur, the neuron can be activated. These coincident spikes may result from other synchronized thalamic inputs (Alonso et al. 1996; Roy andAlloway 2001) or from diffuse asynchronously arriving background activity. This background level could function as a control parameter that determines the operating point of the target neuron and thereby the efficacy of the otherwise subthreshold connection (Aertsen et al. 1994). For low synaptic efficacies, as found for cortico-cortical connections, the functional connection strength showed a monotonic increase with background firing rate (Eggermont 1992). As we will show, one can separate the spikes recorded during continuous stimulation in those that contribute to the STRF (IN-STRF spikes) and those that do not (OUT-STRF spikes). For IN-STRF spikes, the peak correlation strength between two neurons will likely depend strongly on the overlap of the STRFs of these neurons, whereas for the “spontaneous spikes” would not be or less dependent thereof. Such a functional separation of

Address for reprint requests and other correspondence: J. J. Eggermont, Dept. of Psychology, University of Calgary, Calgary, Alberta T2N 1N4, Canada (E-mail: eggermon@ucalgary.ca).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
the spikes allows the correlation of the OUT-STRF spikes for both cells; this is a measure that can directly be compared with the correlation under spontaneous firing conditions (i.e., during silence). This would allow inferences of the effect of stimulation on background cortical network activity and of the effect of this background activity on the stimulus induced correlations.

Neural correlation and joint STRF properties

Receptive fields of neurons are determined by the contributions of their afferent inputs and the balance between excitation and inhibition therein. The EPSPs and the inhibitory postsynaptic potentials (IPSPs) recorded from neurons in primary auditory cortex have the same frequency-tuning curves (Wehr and Zador 2003). These postsynaptic potential-based tuning curves are broader than those based on the spike firings of the same neuron that result from this input (Tan et al. 2004). Thus at the edges of the postsynaptic potential-based tuning curves the inhibitory inputs prevent the excitatory ones from firing the neuron.

The first exploration of coincident firing and STRFs was done in the visual cortex of the cat (Ghose et al. 1994) using reverse correlation to calculate the joint receptive field for correlated spikes of the two neurons recorded on a single electrode. They suggested that coincident firings showed greater spatial and temporal resolution compared with that displayed by the single units. We will explore the same but using simultaneously recorded activity from \(\sim 16\) electrodes, thereby exploring a larger range of STRF overlaps.

Neurons of the ventral part of the medial geniculate body project to primary auditory cortex with a wide divergence (of the order of several millimeters) of their axon collaterals, whereby each cortical cell receives overlapping input from \(\sim 30\) thalamic neurons with different CFs (Miller et al. 2001). The number of thalamocortical axon collaterals that are shared by two cortical cells (Lee et al. 2004), and/or the synchronous activation of two cortical neurons by highly synchronized activity of different thalamocortical cells, will determine the degree of neural synchrony in the firings of the cortical cells. Additional aspects that determine the strength of the neural correlation include the synaptic strengths and the pattern of firing (e.g., bursting) of the thalamocortical cells. Similar aspects apply to the contributions of the horizontal fiber activity to neural synchrony.

The receptive field corresponding to the synchronous spikes shared by the two cortical neurons will be close to those of the thalamic neurons that provide axon collaterals to both cortical neurons. Cortical neurons that are separated by more than several millimeters will not share axon collaterals from the same thalamic cells. In case a multi-frequency stimulus is used and frequencies that activate both neurons are simultaneously present, a stimulus induced neural synchrony could result that would show a functional receptive field that is similar to the union of the two individual cell’s receptive fields.

Synchrony in the firings of cells in auditory cortex of the anesthetized cat during long periods of silence may be related to the sleep spindles induced by ketamine anesthesia that are similar to those in slow wave sleep (Steriade 2001). However, sound stimulation may disrupt part of this correlation, potentially by “arousing” the cat and affecting the spindles. In the course of that process, sound stimulation may be breaking up the large-assembly sleep state into a number of smaller, mutually independent, assemblies that represent a dynamic state resembling that observed in the alert animal (Miller and Schreiner 2000). The decrease in neural synchrony, both in the EEG and in the peak cross-correlation coefficient that is often observed during stimulation, may thus reflect the reduced neural assembly size in a stimulus-processing brain.

The goals of this study are to show that neural synchrony during long periods of silence is not the appropriate control for neural synchrony during equally long periods of stimulation and that coincident neural activity serves to sharpen the STRF in both the spectral and temporal domain.

Here, we show that stimulation reduces the correlation in background activity and as a result the signal-to-noise ratio of correlated activity in response to the stimulus is enhanced. The cross-correlation for IN-STRF spikes depended much stronger on the STRF overlap than the cross-correlation during either silence or for OUT-STRF spikes. Reconstructed STRFs for coincident spikes are smaller than their single-electrode counterparts and showed larger weighted STRF overlaps largely due to the reduction of the unit’s STRF size. This suggests that coincident neural activity serves to sharpen the resolution in the spectro-temporal domain.

METHODS

The care and the use of animals reported in this study was approved (No. BI 2001–021) by the Life and Environmental Sciences Animal Care Committee of the University of Calgary. All animals were maintained and handled according to the guidelines set by the Canadian Council of Animal Care.

Animal preparation

All animals were deeply anesthetized with the administration of \(25\) mg/kg of ketamine hydrochloride and \(20\) mg/kg of pentobarbital sodium, injected intramuscularly. A mixture of \(0.2\) ml of acepromazine (0.25 mg/ml) and \(0.8\) ml of atropine methyl nitrate (25 mg/ml) was administered subcutaneously at \(-0.25\) ml/kg body wt. The tissue overlying the right side of the skull overlying the temporal lobe was removed and two \(8\)-mm holes (centered \(-9\) mm posterior and \(21\) mm ventral from bregma and \(-5\) mm posterior and \(17\) mm ventral from bregma) were trephined in the skull. If necessary, the holes were enlarged with small bone rongeurs to ensure that the primary auditory cortex (AI) was fully exposed. The dura was then cut back, and a photo was taken of the exposed cortical surface such that the brain surface vascular pattern could be used as spatial reference for locating electrode placements. We used the location of the electrodes as well as the characteristic frequency of the neurons to ensure that the electrodes were located in AI. First, the anterior and posterior ectosylvian sulci (roughly delineating AI) were visible; the electrode arrays were usually located between their tips or slightly more dorsal. Second, the tonotopic gradient reverses between AI and anterior auditory field (AAF), passing from an antero-posterior gradient in AI to a dorsoventral gradient in AAF. The exposed cortical surface was covered with light mineral oil to prevent the tissue from drying. Throughout the experiment, light anesthesia (sufficient to ensure that pinna reflexes were absent) was maintained with ketamine hydrochloride (5–10 mg · kg\(^{-1}\) · h\(^{-1}\)). The acepromazine atropine methyl nitrate mixture was administered approximately every 2 h to control fluid secretion in the airways. The body temperature was monitored and maintained around 37\(^\circ\)C with a thermostatically controlled heating blanket. After the experiment, the animals were administered a lethal dose of pentobarbital sodium.
Acoustic stimulus presentation

Stimuli were generated in MATLAB and transferred to the DSP boards of a TDT-2 (Tucker Davis Technologies) sound-delivery system. Acoustic stimuli were presented in an anechoic room from a speaker system [Fostex RM765 in combination with a Realistic Super-Tweeter that produced a flat spectrum (±5 dB) ≤40 kHz measured at the cat’s head] placed 30° from the midline into the contralateral field, ~50 cm from the cat’s left ear. Calibration and monitoring of the sound field was accomplished with a condenser microphone (Bruel and Kjær 4134) placed above the animal’s head, facing the speaker, and a measuring amplifier (Bruel and Kjaer 2636). Prior to acute recordings peripheral hearing sensitivity was determined using auditory brain stem response (ABR) thresholds (details in Noreta et al. 2003).

Cortical tuning properties

Frequency-tuning curves were measured by randomly presenting 27 gamma-tone pips with frequencies covering five octaves (e.g., 1.25–40 kHz) in equal logarithmic steps and presented at eight different stimulus levels in 10-dB steps (e.g., 5–75 dB SPL) at a rate of 4/s such that each intensity-frequency combination was repeated five times. The envelope of the gamma tones is given by

\[ \gamma(t) = (4t^2) \exp(-4t) \]  

with \( t \) in ms. The duration of the gamma tones at half-peak amplitude was 15 ms, and the envelope was truncated at 50 ms. Because of the exponential decay, the amplitude value at the 50-ms point is only 3.7 \( \times 10^{-6} \), i.e., 109 dB down from the value at the peak and thus causing no audible transients.

STRFs were obtained by presenting multi-frequency stimuli consisting of randomly presented gamma-tone pips, as defined in the preceding text. Here, tone pips for each of 81 frequencies in five octaves were presented according to a Poisson process (Blake and Merzenich 2002), with similar average rate but different randomization for each frequency. Each frequency was presented at a rate of 0.25 Hz so that the aggregate tone-pip rate was 20/s. On average the tone pips did not overlap, i.e., the tone-pip duration was equal to the average presentation interval but because of the random nature of the presentations, a great deal of temporal overlap ensued. Thus the probability of two overlapping tone pips, as calculated from the Poisson distribution, is equal to 0.184, the probability of three overlapping tone pips is 0.061, and for four overlapping tone pips it is only 0.015.

Recording and spike separation procedure

Two arrays of eight electrodes (Frederic Haer) each with impedances between 1 and 2 MΩ were used. The electrodes were arranged in a 4×2 configuration with inter-electrode distance within rows and columns equal to 0.5 mm. Each electrode array was oriented such that all electrodes were touching the cortical surface and then were manually and independently advanced using a Narishige M101 hydraulic microdrive (1 drive for each array). Some dimpling of cortex manually and independently advanced using a Narishige M101 hydraulic microdrive (1 drive for each array). Some dimpling of cortex was done again to form a multi-unit spike train, thereby eliminating potential contributions from thalamo-cortical afferents or fast spikes from interneurons.

Data analysis

FREQUENCY TUNING. To assess frequency-tuning properties, the peak number of action potentials in the poststimulus time histogram (PSTH, 5-ms bins) calculated over the first 100 ms after gamma-tone presentation was estimated. This long window was used to capture long-latency activity especially at near-threshold intensity levels. The results were calculated per stimulus intensity and were combined into an intensity-frequency-rate profile from which tuning curves, rate-intensity functions, and iso-intensity rate-frequency contours could be derived (Eggermont 1996) using routines implemented in MATLAB. The frequency-tuning curve was defined for a firing rate at 25% of the maximum peak-firing rate (FRmax). This was −10–20% above the background firing-rate, but as the latter was dependent on the level of stimulus-induced suppression, the tuning curve criterion based on a percentage of peak firing rate was preferred over that based on increase over background activity. The values of firing rate given in this paper are in spikes/s. The number of spikes in a 5-ms bin is divided by the number of repetitions (5 repetitions over 3 frequencies = 15) times the bin size in seconds (0.005 s), i.e., by 0.075. This gives very high values of FRmax when the MSU spikes are well synchronized to the onset of stimulus. The threshold was determined as the lowest intensity that produced visible time locked responses to the tone pip.

The STRF was determined by constructing PSTHs for each of the preceding gamma tones in a 200-ms window. For that purpose, each spike elicited was plotted several times in the appropriate frequency bins and in the 200-ms time window after each of the preceding gamma tones (because spike latency is a priori unknown). If the gamma tones had no effect on a spontaneously firing neuron, the entire matrix of 81 frequency bins by 100 (2 ms wide) time bins would be filled uniformly. If certain frequencies consistently produce excitation especially at near-threshold intensity levels. The results were calculated per stimulus intensity and were combined into an intensity-frequency-rate profile from which tuning curves, rate-intensity functions, and iso-intensity rate-frequency contours could be derived (Eggermont 1996) using routines implemented in MATLAB. The frequency-tuning curve was defined for a firing rate at 25% of the maximum peak-firing rate (FRmax). This was −10–20% above the background firing-rate, but as the latter was dependent on the level of stimulus-induced suppression, the tuning curve criterion based on a percentage of peak firing rate was preferred over that based on increase over background activity. The values of firing rate given in this paper are in spikes/s. The number of spikes in a 5-ms bin is divided by the number of repetitions (5 repetitions over 3 frequencies = 15) times the bin size in seconds (0.005 s), i.e., by 0.075. This gives very high values of FRmax when the MSU spikes are well synchronized to the onset of stimulus. The threshold was determined as the lowest intensity that produced visible time locked responses to the tone pip.

For STSFs, contour lines or color-coded images are plotted at a certain intensity level. STSF overlap was measured by taking the contour line representing 50% above the mean value of the difference between the maximum response and the mean value of the STSF for each of the neurons, calculating the number of pixels for each neuron (each pixel defined as: 1 frequency bin by 1 2-ms time bin) and the number of overlapping pixels. This 50% value was determined on the basis of giving the strongest correlation between the cross-correlation coefficient and the weighted overlap, but values between 20 and 40% gave similar results. The weighted STSF overlap was defined as [overlap size/sqrt(STSF size 1 × STSF size 2)], which scaled between 0 (no overlap) and 1 (complete overlap). The minimum latency of the 30% excitation contour line, the temporal duration of the 30% excitation contour line, the frequency of the center of gravity and the spectral bandwidth of the 30% contour lines were determined for each stimulus.

CROSS-CORRELATION. Cross-correlograms were calculated using custom-made programs in Matlab. Quantification of neural correlation was done on the basis of the cross-correlation coefficient

\[ R(\tau) = \frac{N_{\text{cross}}(\tau) - \bar{E}}{\sqrt{\sum_{\tau} (N_{\text{cross}}(\tau) - \bar{E})^2}} \]

where \( N_{\text{cross}}(\tau) \) is the number of coincidences in the bin corresponding to lag time \( \tau \), \( E \) is the expected value for coincidences under the assumption of independent spike trains

\[ E = \frac{N_A N_B}{N} \text{ with } N = T/\Delta \]

where \( N_A \) and \( N_B \) are the number of spikes in the recording, \( \Delta \) is the bin size, and \( T \) the duration of the recording |\( R(\tau) | \leq 1. \]
The peak cross-correlation coefficient was considered significant if its peak was within 10 ms from the zero-lag point, and its value exceeded the baseline by 4 SD. For firing rates <10 spikes/s and small bin sizes, SD = $(N)^{-1.5}$, with $N$ being the number of bins in the recording (Eggermont 1992). Cross-correlogram peak widths were expressed in milliseconds and were measured at half-amplitude between peak and baseline.

Spike contributing to the part of the STRF within the 30% contour line were considered as IN-STRF spikes; those not contributing to the STRF were considered OUT-STRF spikes. To calculate mean firing rates and cross-correlation coefficients for the IN-STRF and OUT-STRF conditions, we need to estimate the effective number of bins for each condition. For the IN-STRF condition, we assumed this to be proportional to the summed STRF areas minus the overlap, relative to 5 octaves by 50 ms (the average inter-tone-pip interval). Thus the number of bins for the IN-STRF spikes will equal $N \times (\text{summed STRF area} - \text{overlap})/(5 \text{ octave} \times 50 \text{ ms})$ and for the OUT-STRF spikes will equal $N \times [1 - (\text{summed STRF area} - \text{overlap})]/(5 \text{ octave} \times 50 \text{ ms})$.

All statistical analyses were performed using Statview 5 (SAS Institute). We will use $r$ to report statistical correlation (Pearson) as $R$ is reserved for the peak cross-correlation coefficient.

RESULTS

Recordings were made from the right AI in 17 adult cats. We recorded the neural activity under spontaneous firing conditions and during stimulation, at 65 dB SPL, from the same units. Multi-unit (MU) recordings (281) showed no significant firing rate changes in 50-s sections through 900 s of silence and during 900 s of stimulation, i.e., were stationary. A total of 732 cross-correlograms were calculated of which 685 were significant at a level of 4 SD. The cross-correlograms of 545 MSU pairs with peak lag times within 10 ms from zero lag time were analyzed. No correction for stimulus coordination is applied.

Illustrative example

Figure 1 shows the procedure for selecting IN- and OUT-STRF spikes and the calculation of the STRF overlap. A–C show STRFs obtained with 15-min multi-frequency stimulation in the cat primary auditory cortex. These three panels correspond to STRFs for simultaneous recordings made from three different electrodes in same array. The location of the recording electrode in the array is given as channel numbers 2, 5, and 8 on the top of each panel. D shows the STRF overlap of recordings for the three STRFs shown in A–C. Spikes contributing to the part of the STRF within this boundary were considered as IN-STRF spikes, those not contributing to the STRF were considered OUT-STRF spikes. The amount of overlap between STRFs for electrodes 2 (red) and 8 (blue) is smaller than that between STRFs for electrodes 5 (green) and 8 (blue), and the weighted overlap was 0.40 and 0.77, respectively.

Figure 2 shows cross-correlation histograms for two pair combinations of the three recordings shown in Fig. 1. The top
FIG. 2. Examples of cross-correlograms for 2 pairs of units with STRFs shown in Fig. 1. Comparisons are shown among IN-STRF, OUT-STRF, and spontaneous spikes collected during 15 min of silence.

row shows the correlograms for electrodes 2 and 8 and the bottom row the correlograms for electrodes 5 and 8. Note that the vertical axes are scaled differently. The dashed horizontal lines indicate the 4 SD boundary above the expected value (horizontal full lines). The left and middle columns show the cross-correlation histogram for the IN- and OUT-STRF spikes obtained during stimulation. The right column shows the cross-correlation histogram for a 15-min silence condition for the same electrodes and location as during stimulation. The overlap in STRF between electrodes 5 and 8 was nearly twice that between electrodes 2 and 8 (cf. Fig. 1), and the peak cross-correlation coefficient (R) was also larger for these electrode pairs.

Group data

The mean R values obtained for spontaneous and stimulus driven activity are shown in Table 1. We distinguish spontaneous activity (recorded during 15 min of silence), stimulated activity (all spikes recorded during 15 min of stimulation), and IN- and OUT-STRF spikes. The only significant difference in R is found for the OUT-STRF condition for which R was lower than the three other values (P < 0.0001). The IN-STRF peak cross-correlation was strongly correlated (r = 0.59) with the OUT-STRF peak cross-correlation. The width of the correlogram peak at half height (W) was significantly smaller (P < 0.0001) for the IN-STRF compared with the other conditions, which were not significantly different from each other. The areas (RW) under the correlogram were all significantly different (P < 0.0001) except between the IN- and OUT-STRF conditions (P = 0.47).

The geometric mean firing rates for the IN-STRF condition were significantly higher (P < 0.01) than for the other three conditions, which were not significantly different from each other. There was a significant correlation between the geometric mean firing rates for IN-STRF spikes and OUT-STRF spikes (r² = 0.13, P < 0.0001). R for IN-STRF spikes was significantly correlated with the geometric mean firing rate for IN-STRF spikes (r² = 0.19, P < 0.0001), but there was no correlation between R for IN-STRF spikes and the geometric mean firing rate for OUT-STRF spikes (r² = 0.007, P = 0.3).

STRF overlap

The weighted STRF overlap and bandwidth (BW) overlap reflect the divergence and convergence of the thalamo-cortical afferents onto cortical cells. The weighted STRF-overlap depended significantly on the spatial distance in the posterior-anterior direction [F(3,497) = 7.6; P < 0.0001], with the mean overlap (0.52 ± 0.20; mean ± SD) at zero distance in the posterior-anterior direction (but still for activity recorded on separate electrodes) being significantly larger (P < 0.05) that those at 0.5 mm (0.46 ± 0.21), 1.0 mm (0.41 ± 0.18), and 1.5 mm (0.37 ± 0.16). No other differences were significant. Surprisingly, the BW overlap (in octaves) and the STRF-duration overlap (in ms) were independent of the electrode distance in the posterior-anterior direction, suggesting that the BW and response duration are not independent factors in the determination of the STRF and its overlap. By this we mean that the STRF is not equal to the product of its marginal distributions: the BW and the STRF duration. Both weighted STRF overlap (only values >0; 0.45 ± 0.20) and the log of the BW overlap (mean = −0.21, i.e., mean BW overlap = 0.615 ± 0.28 octaves) were normally distributed (Kolmogorov-Smirnov test) across all distances and remained so for restricted distances.

The stimulus level dependence (from 10 to 60 dB SPL) of the weighted STRF overlap was studied in three sets of recordings in three different cats. A total of 58 pair comparisons were made between electrodes within an array; in 52 pairs there was an STRF overlap >0. The number of combinations that gave zero STRF overlap comprised the same set independent of stimulus level. For the 52 pairs with overlapping STRF, there was no dependence of the weighted STRF overlap on stimulus level (Fig. 3; r² = 0.011, P = 0.14). This is not surprising because STRFs show very little dependence on stimulus level (Valentine and Eggermont 2004). Six of the 52 pairs showed a significant dependence on intensity level (3 were negatively correlated and 3 positively). This suggests that the STRF overlap obtained at 65 dB SPL is representative for those at other stimulus levels.

Dependence of R on STRF parameters

A stepwise regression for R obtained during spontaneous firing was done for nine independent variables: weighted STRF overlap,
overlap in BW (octave), overlap in STRF duration (ms), electrode distance (mm), geometric mean of CFs (kHz), geometric mean firing rate, CF distance (octave), geometric mean peak latency (ms), and peak latency difference (ms), resulted in six significant dependencies. Because there was no significant dependence between R and CF distance (octave), between R and overlap in latency range (ms), and between R and geometric mean peak latency, these variables were not included in the model for multiple regression. Thus a multiple regression was run for the six remaining variables and the four conditions shown in Table 1.

Except for the spontaneous condition where all regression coefficients were significantly different from zero \( (P < 0.05) \), for the three other conditions there was no significant dependence on BW overlap. We continue with describing the five remaining significant dependencies.

First, there was a weak but significant \( (r^2 = 0.04; \text{slope of regression line}, P < 0.0001) \) dependence of R on the geometric mean of the CFs of the units such that the R values increased with increasing CFs. Because all conditions shown in Table 1 were present at any given CF combination, this CF dependence will not further be analyzed.

Second, Fig. 4 shows the dependence of R on the posterior-anterior distance \( (D) \), i.e., along the frequency-gradient, between electrodes within an array for spontaneous firing during 15 min of silence, and for the IN- and OUT-STRF spikes during stimulation. The R values are normally distributed if plotted on a logarithmic scale. One observes that the mean R decreases with increasing D both for the spontaneous condition and for OUT-STRF spikes during stimulation. An ANOVA showed a significant dependence of R on D in all three conditions \( (P < 0.005) \). Post hoc Scheffe tests showed that for the spontaneous condition, the \( D = 0.0 \text{ mm} \) data were significantly \( (P < 0.05) \) different from the \( D = 1 \) and \( D = 1.5 \text{ mm} \) data, and the \( D = 0.5 \text{ mm} \) data were significantly different from the \( D = 1.5 \text{ mm} \) data. For the IN-STRF group, the only significant difference was between the \( D = 0.0 \) and \( D = 1.5 \text{ mm} \) data. For the OUT-STRF group, the \( D = 0.0 \text{ mm} \) and \( D = 0.5 \text{ mm} \) data were both significantly different from \( D = 1.0 \) and \( D = 1.5 \text{ mm} \) data. Extending this analysis to correlations between units recorded with different electrode arrays showed no further dependence on distance. In fact, no significant differences between R values at 1 mm distance and any larger distance \((\leq 4 \text{ mm})\) could be demonstrated for any of the four conditions in Table 1.

Third, to parcel out the dependence of R on electrode distance, we analyzed separately the 247 MSU pairs with \( D = 0.5 \text{ mm} \) in the posterior-anterior direction. We will also show the values obtained when all distances were included. Figure 5 shows the geometric mean of the firing rate \( [M(FR)] \) of the units and the peak cross-correlation coefficients \( (R) \) during silence as a function of weighted STRF overlap \((0\text{-overlap data are considered separately at the end of this section})\). Locally weighted regression lines with tension 0.66 (Cleveland 1981) are drawn in. M(FR) is not related to the weighted STRF overlap \( (r^2 = 0.011, P = 0.054) \), but R is positively correlated with the weighted STRF overlap \( (r^2 = 0.12, P < 0.0001) \).

When all distances were included, the conclusions were the same, as the correlation with STRF overlap decreased to 0.009 for M(FR) and increased to 0.18 for R. The range of R values at any given STRF overlap is about a factor 10, suggesting that STRF overlap is only one of the aspects that determine R.

For \( D = 0.5 \text{ mm} \), Fig. 6 shows the scattergrams of M(FR) and R as a function of the weighted STRF overlap for IN- and OUT-STRF spikes during stimulation. The M(FR) for the OUT-STRF condition was not correlated \( (r^2 = 0.01, P = 0.1) \) with the weighted STRF overlap, whereas the M(FR) for the IN-STRF condition was positively correlated \( (r^2 = 0.13, P < 0.0001) \) with STRF overlap (Fig. 6A), albeit that the increase only occurred for STRF overlaps \( > 0.5 \). These findings did not change qualitatively when all inter-electrode distances were included (Table 1).

R was correlated with the weighted STRF overlap for both IN and OUT conditions (Fig. 6B). The dependence of R was strongest for the IN-STRF spikes \( (r^2 = 0.54, P < 0.0001) \). For OUT-STRF spikes, the correlation between R and the weighted STRF was weaker but still highly significant \( (r^2 = 0.13, P < 0.0001) \). These findings did not change qualitatively when all
distances were included: for In-STRF spikes \( r^2 = 0.36 \) (\( P < 0.0001 \)) and for OUT-STRF spikes \( r^2 = 0.21 \) (\( P < 0.0001 \)). For spontaneous firings, the dependence on weighted STRF was similar to that for OUT-STRF correlation (\( r^2 = 0.14, P < 0.0001 \)).

The peak width of the cross-correlogram was negatively correlated with weighted STRF overlap for OUT-STRF (\( r^2 = 0.04, P < 0.005 \)) and silence conditions (\( r^2 = 0.05, P < 0.001 \)), but not correlated (\( r^2 = 0.004, P = 0.3 \)) for the IN-STRF condition. When all distances were included, the values for OUT-STRF and silence were, respectively: \( r^2 = 0.1 \) (\( P < 0.0001 \)) and \( r^2 = 0.16 \) (\( P < 0.0001 \)). As a consequence, the area under the cross-correlation peak increased with weighted STRF overlap for the IN-STRF condition (\( r^2 = 0.2, P < 0.0001 \)), but was independent of weighted STRF overlap for the two other conditions (\( r^2 < 0.005, P > 0.3 \)). This conclusion also applied when all distances were included.

Fourth, the \( R \) values were correlated with weighted STRF overlap for OUT-STRF (\( r^2 = 0.04, P < 0.005 \)) and silence conditions (\( r^2 = 0.05, P < 0.001 \)), but not correlated (\( r^2 = 0.004, P = 0.3 \)) for the IN-STRF condition. When all distances were included, the values for OUT-STRF and silence were, respectively: \( r^2 = 0.1 \) (\( P < 0.0001 \)) and \( r^2 = 0.16 \) (\( P < 0.0001 \)). As a consequence, the area under the cross-correlation peak increased with weighted STRF overlap for the IN-STRF condition (\( r^2 = 0.2, P < 0.0001 \)), but was independent of weighted STRF overlap for the two other conditions (\( r^2 < 0.005, P > 0.3 \)). This conclusion also applied when all distances were included.

For 31 neuron pairs recorded from the same electrode array (all distances included) for which there was no overlap between the STRFs, the mean \( R \) during stimulation (all spikes) was significantly lower than during silence (0.016 vs. 0.024, \( P < 0.01 \)). Not surprisingly, the IN-STRF spikes for these pairs peak STRF latency (\( P < 0.0001 \)). For latency differences >10 ms, the \( R \) values were all <0.02. As expected, weighted STRF overlap was negatively correlated (\( P < 0.001 \)) with the absolute difference in peak latencies. Absolute peak latency difference increased significantly (factor 1.7 per mm; \( P < 0.05 \)) with posterior-anterior distance between the electrodes.

The data are summarized in Table 2.

**Table 2.** Overview of significant (\( P < 0.05 \)) positive (+) and negative (−) correlations.

<table>
<thead>
<tr>
<th>R</th>
<th>Silence</th>
<th>IN-STRF</th>
<th>OUT-STRF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance in posterior-anterior direction</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Geometric mean CF</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>CF difference, octave</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted STRF overlap</td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Geometric mean firing rates</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Geometric mean peak latency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geometric mean STRF duration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peak STRF latency difference</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

STRF, spectro-temporal receptive field.
were only very weakly correlated (mean $R = 0.004$, i.e., not significantly different from 0), suggesting that on average, STRF overlap is the dominant factor for IN-STRF spike correlation.

Reconstruction of STRFs based on coincident spikes

On the basis of coincidences in the peak of the cross-correlogram for the IN-STRF spikes, we can extract the coincident-spike trains of the two units involved (Fig. 7). First, a time window is established for acceptable cross-intervals for coincident spikes. Then the spikes of the two trains contributing to these cross-intervals are selected to constitute a coincident-firing spike train for each unit. These two spike trains can then be used to construct a coincident-firing STRFs for each unit. The remaining spikes for each unit form their noncoincident spike trains. Figure 8 compares STRFs for coincident-spike trains and for noncoincident spike trains in comparison with the all-spike STRF. We show the results for two units that are also featured in Figs. 1 and 2. On basis of the cross-correlogram for IN-STRF spikes, we used a time window of ±10 ms as a criterion for coincident spikes, all remaining spikes constitute the noncoincident spikes. Figure 8, top, shows the coincident-spike STRFs for electrodes 2 and 8, and in the right-most column, the 30% contour lines of the two STRFs. The middle row shows the same for noncoincident spikes, and the bottom shows the all-spike STRFs with the 30% contour lines of the coincident-spike STRFs superimposed. The bottom right panel shows the all-spike contour lines (in black) and the coincident-spike contour lines (color as in top right). The first thing one notices is that the electrode 2 contour lines for the coincident spike STRF are identical to those for the all-spike STRF, albeit that some differences in the STRFs are evident. This suggests that in this case coincident spikes provide an overall scaling of the STRF. Second, the coincident-spike STRF for electrode 8 is a subset with smaller latencies of that for the all-spike STRF. Third, one observes that the coincident-spike STRFs are not identical for the two MUs, suggesting that coincident firings in response to nearly simultaneous but dissimilar frequencies do occur.

As expected, the noncoincident-spike STRFs are complementary (middle), albeit that the noncoincident-spike STRFs overlap with those of the coincident-spike ones for the same MUs. When the criterion for coincident spikes is more confined, e.g., at ±2.5 ms, the main effect is the shrinking in the temporal domain of the STRF of electrode 8: its longest latency was reduced from 58 to 48 ms. For the electrode 2, STRF the faint green-blue area at the bottom right of the coincident-spike STRF (top left) disappeared. So the weighted overlap between the two coincident-spike STRFs increased with a stricter criterion for coincident spikes, but this was largely due to a reduced duration range.

We analyzed the weighted STRF overlap for coincident spikes (±10 ms) for 188 pairs composed of those units that responded with ≥1,000 spikes in 900 s and had a well-defined STRF. The STRF sizes for those units did decrease slightly (ratio: 0.94 ± 0.38) but significantly ($P < 0.005$) for the coincident-spike STRFs compared with the all-spike STRFs. As expected, the weighted STRF overlap increased significantly ($P < 0.0001$) from 0.46 ± 0.21 (Fig. 10, top), nearly identical to the value when all units are taken into account, see above, suggesting that firing rate does not affect the weighted overlap) to 0.67 ± 0.13 for these pairs (Fig. 10, bottom). Further limiting the extent of the coincident spike window to ±2.5 ms did significantly increase the weighted STRF overlap to 0.74 ± 0.12 (Fig. 9). The significant decrease of weighted STRF overlap for all spikes with inter-electrode distance disappeared for the coincident-spike STRF overlap, both for the ±10 ms and the ±2.5-ms window (Fig. 10).

The IN-STRF $R$ values increased with increasing weighted coincident-spike STRF overlap, here only shown for the coincidences within the ±2.5-ms window (Fig. 11), suggesting that while coincident spikes tend to occur in a smaller part of the frequency-time plane, the noncoincident spikes contributing to the STRF (cf. Fig. 11) are still important in the determination of the $R$ values. The $R$ values were also conspicuously larger for geometric-mean STRF sizes in the region between 50 and 150 pixels than for larger STRF sizes. These small STRFs were those that had high peak firing rates.

![Fig. 7. Procedure to extract coincident IN-STRF spike trains and reconstruct the coincident-spike STRFs. The STRFs of the 2 neurons are shown as the red ellipses. The 2 spike trains $A$ and $B$ (in red) represent the IN-STRF spikes. After applying a coincidence window for the cross-intervals (equal to the interval between the left or right dotted line and the center dotted line in the cross-correlogram) the 2 coincident-spike trains (in green) are obtained. The coincident-spike STRFs are shown as green circles.](http://jn.physiology.org/ by 10.220.32.247 on June 10, 2017)
**Toward population coding**

How do pair-wise coincident-spike STRFs change across an electrode array? To illustrate this, the STRFs for all electrodes in an array, and representative for recordings in AI, are plotted in Fig. 12 as color maps. This set was recorded simultaneously with the data shown in Figs. 1 and 8 but with the other electrode array. Note that the latency range for the individual panels may be different as a result of centering the STRFs for maximum visibility. The BF gradually decreases with increasing electrode number (refer to Fig. 1 for the electrode array layout) as expected for an orientation of the array where low number electrodes are more anterior than those with higher numbering. The peak latencies range between 25 (electrode 1) and 35 ms (electrodes 7 and 8). The peak responses of the all-spike STRFs for each electrode, and the pair-wise peak cross-correlation coefficients are shown in Table 3. The peak cross-correlation was not significantly related to mean firing rate if the electrode 1-electrode 2 pair was excluded (an outlier). All cross-correlogram peaks were wider than 5 ms at the half-amplitude point (range: 5–30 ms).

All pair-wise coincident-spike (within a ±2.5-ms window) STRFs that involve electrode 1 (to reduce the number of pair combinations to be shown from 28 to 7) are plotted as 30% contour lines and overlaid on the all-spike STRFs. Thus the coincident spike STRFs for electrodes 1 and 2 are shown as black lines, those between electrode 1 and electrode 3 as red lines, etc. For electrode 1, there are thus 7 contour lines that are overlaid on the all-spike STRF. For the other electrodes, there is only one contour line, and its color-line-style code can be inferred from those plots. It is noted that the individual coincident-spike STRFs generally overlap with their peak-response part (red-brown) in the all-spike STRFs. Finally in the center panel of the figure all the contour lines of the individual electrodes are superimposed (for electrode 1 we used that for the coincident spikes with electrode 2). Note that these contours in general differ from those of electrode 1 (as elucidated in Fig. 8). The large overlap regions ~10 kHz take the attention away from a narrow part of the STRF situated just <20 kHz, which is represented in the all-spike STRFs for each electrode, albeit below the 30% level for electrodes 2 and 8. Also in this region of the STRF coincident-spikes are clearly present.

A population-STRF can be defined for which all pair-wise coincident-spike STRFs overlap: this STRF is centered slightly above 10 kHz and ~37 ms latency for electrode 1 (cross-hatched in Fig. 12, top left) and does not correspond to the area with the highest peak response. The overlap region of the other electrodes’ coincident-spike STRFs is found around the same frequency-latency coordinates but is distinctly smaller in size (see center panel, cross-hatched area). The latter all-electrode overlap region is in fact nearly identical to that part of the coincident-spike STRF for electrode 5 that is situated slightly
above 10 kHz, however, is not its strongest responding part. This all-inclusive STRF is the upper boundary of an STRF based on spikes that are coincident across the spike trains for all 8 electrodes. Excluding electrode 5 from this putative synchronized assembly would increase the overlap area considerably and to the size of that shown in the panel for electrode 1.

**DISCUSSION**

We showed that OUT-STRF spikes had peak cross-correlation coefficients that were ~25% lower than those for firings during 15 min of silence. This suggests that stimulation reduces the correlation of background activity and as a result the signal-to-noise ratio of correlated activity in response to the stimulus is enhanced. In contrast, the cross-correlation coefficients for spikes contributing to the STRF were not significantly different from those during silence, albeit that their mean firing rates were a factor 2.5 higher. There was also a difference in the peak width of the correlogram between the IN-STRF on one hand and OUT-STRF and silence conditions on the other, suggesting a tighter synchrony in the IN-STRF condition. Thus although there was no difference in the peak R value between silence and stimulation, the correlogram peak was much narrower. Combined the results suggest that the OUT-STRF condition is the proper control to judge the effect of stimulation on neural synchrony. We agree that calculating a shift predictor or joint-PSTH predictor is the proper way to extract the effects of stimulation on the effective neural connectivity (Espinoza and Gerstein 1988). However, as we have previously shown (Eggermont 1994), when the firings of the neurons are tightly locked to the stimulus, as they typically are in auditory cortex even for our steady-state stimulation, the correction basically removes all or nearly all correlation such that the result is often much less than the neural correlation under spontaneous conditions. Is the difference between the OUT-STRF peak neural correlation and the IN-STRF peak neural synchrony a sign of changed effective connectivity or just the extra synchrony provided by the stimulus? We cannot answer that question using a shift-predictor as the particular stimulus we used does never repeat itself. However, our experience with the other correction procedures also suggests that these are generally over-correcting and thus generally suggest a reduction in neural correlation during stimulation. Thus in the

**FIG. 9.** Distribution of weighted STRF overlap for a subset of high-firing rate neurons. Top: the result for all-spike STRFs; middle: coincident (±10 ms) spike STRFs; bottom: coincident (±2.5 ms) spike STRFs.
The present study, all IN-STRF spike correlograms are uncorrected. In this respect, we like to remark that it is unlikely that the nervous system performs a correction for stimulus-induced correlation as estimated by the various predictors. It is the actual spike coincidences that are affecting the potential for firing in a target neuron not the stimulus-corrected ones (deCharms and Merzenich 1996). Thus our raw correlations (neural synchrony) may effectively estimate those coincident firings between neurons that could play a role in neural population coding of sound.

The cross-correlation for IN-STRF spikes depended much more strongly on the STRF overlap than the cross-correlation during either silence or for OUT-STRF spikes. This difference was enhanced when the area under the cross-correlogram peak was considered; the area increased with STRF overlap for IN-STRF spikes but was independent of overlap for OUT-STRF spikes and spikes recorded during silence. The reason for this was the decrease in peak width with increasing STRF overlap for the latter two conditions. This suggests that for the OUT-STRF and silence condition the number of synchronized firings stays the same with increasing STRF overlap, but they become concentrated in a roughly two times (at the highest overlap) smaller time window than at the lowest overlap condition. Assuming that the contribution from intra-cortical fibers does not change, for cortical units with increasing STRF separation, the spontaneous inputs from their thalamic afferents show increasing asynchrony. Because all the overlap data relate to recordings within one electrode array, the maximum distance was 1.5 mm and well within the range for sharing axon collaterals from the same thalamic cells.

Our findings for the dependence of $R$ on joint-STRF properties are summarized in Table 3. For the IN-STRF spikes the peak width did not change with STRF overlap and was on average equal to that for full overlap in the spontaneous and OUT-STRF conditions. This suggests that the inputs from different thalamic cells stayed synchronized regardless their STRF separation but that the percentage of synchronized cortical spikes decreases with decreasing overlap in correspondence with the notion that the neurons share less axon collaterals with decreasing STRF overlap. Predictably, the $R$ values were dependent on the difference in the STRF peak latencies because this determines whether activation in the strongest responding areas will be within the duration of the correlogram peak. Typically, the jitter of spike timing increases proportionally with the latency and thus the $R$ value will correspondingly decrease with increasing peak latency difference. Surprisingly, there was no such dependence on the duration of the STRF overlap. This was likely due to the wide range of overlap-duration values within the 30% contour that ranged from 2 to 50 ms and because the individual STRF centers tend to be asymmetrically located within this overlap (cf. Fig. 12).

Although the overall firing rate during stimulation was the same as during silence and there was no significant difference in the overall peak $R$ values, changes in $R$ as a result of stimulation became manifest when IN- and OUT-STRF spikes were considered separately. Compared with that for firings during silence, the cross-correlation for the OUT-STRF spikes was much reduced despite the unchanged firing rate. This suggests that stimulation breaks up the large assembly that exists during long periods of silence into a stimulus related one and maybe several others. As a result, the OUT-STRF spikes of the unit pairs, now likely distributed across several assemblies, are less correlated than during long periods of silence. This also suggests that using the correlation during long periods of silence is not a proper control condition to assess stimulus related changes in correlation. In that condition, the ongoing network activity is significantly different from that during stimulation. This may be related to the use of ketamine anesthesia, but even in awake animals, drowsiness will occur during long periods of silence and that brain state may change after arousal during stimulation.

**Selection of IN-STRF spikes**

The selection of the spikes was done on the basis of the 30% contour between maximum firing rate and mean firing rate in the STRF. This 30% contour defines a “cylinder” for firing levels below the mean and a “cone” above this level. By definition of the STRF, and the way we construct it as the frequency-dependent PSTH, there will also be a percentage of spontaneous spikes in there. The spontaneous spikes are contained in the “cylinder” volume below the mean. The ratio of the stimulus induced and spontaneous spikes will be proportionally to the volume (number of spikes) of the STRF above the mean divided by the volume (number of spikes) below the mean. This signal-to-noise ratio was on average $6.3 \pm 5.7$ with 20% of the spikes below the mean value. Thus the IN-STRF $R$
values will be contaminated by this contribution of spontaneous spikes. Given that the average $R$ value for the OUT-STRF spikes is about 25% smaller than that for the IN-STRF spikes, the latter is underestimated. Because $R$ values do not add up, it is hard to put an exact number on the underestimation, but it will be of the order of 5% (20% × 25%), based on the fraction of spikes below the mean and the difference in the $R$’s for IN- and OUT-STRF spikes.

The calculation of $R$ between IN-STRF spikes of two units also requires that the number of bins for use in Eq. 2 be adjusted to reflect that these spikes are not evenly distributed throughout all bins in the entire 900 s of the recording. If we would not adjust the number of bins according to the size of the STRFs (see METHODS), the result would be an underestimation of the expected value ($E$) and an overestimation of $R$. In comparison, for IN-STRF spikes the uncorrected $R$ would be $0.053 \pm 0.058$, whereas after adjusting the number of bins $R = 0.045 \pm 0.053$, and this was significantly lower. There was hardly any effect on the calculation of $R$ for the OUT-STRF spikes: uncorrected $R = 0.035 \pm 0.031$ and corrected $R = 0.034 \pm 0.031$. Thus the conclusion that stimulation reduces the OUT-STRF correlation compared with that during long periods of silence was not affected by the correction.

Gating role for the spontaneous activity?

The $R$ values for the IN-STRF spikes were strongly correlated with those for the OUT-STRF spikes. This is what one

---

**TABLE 3.** Response parameters and cross-correlation values for Fig. 15

<table>
<thead>
<tr>
<th>Peak</th>
<th>Electrode 1</th>
<th>Electrode 2</th>
<th>Electrode 3</th>
<th>Electrode 4</th>
<th>Electrode 5</th>
<th>Electrode 6</th>
<th>Electrode 7</th>
<th>Electrode 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electrode 2</td>
<td>0.172</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrode 3</td>
<td>0.082</td>
<td>0.078</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrode 4</td>
<td>0.071</td>
<td>0.071</td>
<td>0.050</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrode 5</td>
<td>0.036</td>
<td>0.036</td>
<td>0.082</td>
<td>0.052</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrode 6</td>
<td>0.094</td>
<td>0.079</td>
<td>0.049</td>
<td>0.066</td>
<td>0.049</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrode 7</td>
<td>0.054</td>
<td>0.051</td>
<td>0.059</td>
<td>0.037</td>
<td>0.065</td>
<td>0.071</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrode 8</td>
<td>0.036</td>
<td>0.031</td>
<td>0.048</td>
<td>0.033</td>
<td>0.078</td>
<td>0.042</td>
<td>0.103</td>
<td></td>
</tr>
</tbody>
</table>
could expect if the asynchronous spiking activity of the neurons would set the baseline for threshold crossings. In that case, one would also expect that $R$ for IN-STRF spikes would be correlated with the geometric mean of the firing rates of the two units in the OUT-STRF or silence conditions; this was not the case. An alternative reason for the correlation between the $R$ values for IN- and OUT-STRF spikes is that both depend on synchronous spiking input from the thalamus and that stimulation increases this synchrony. Thus the dependence between the $R$ values in the three conditions reflects largely the efficacy of the anatomical connections. If anatomical connections were the only factor, however, one would not expect that during stimulus conditions, the OUT-STRF cross-correlation would be so much lower than the cross-correlation in silence. Thus there is another factor beyond the STRF overlap, and this may be the degree of synchronous firing of the various thalamic cells that determine the STRFs as well as cortical rhythmic network activity (Eggermont et al. 1993). As theoretical studies have shown, a modest increase in the synchrony between the thalamic inputs to the cortical cells can result in a large increase in the synchrony of the cortical cells (Stroeve and Gielen 2001).

**STRF dependence**

It was surprising that $R$ was not dependent on the CF distance in octaves, albeit that there was a dependence on spatial distance and on STRF overlap. Thus CF distance may be contained in both the spatial distance and the STRF overlap or is not the primary factor. The latter is potentially supported by observations (Brosch and Schreiner 1999; Eggermont 1994) that for spontaneously firing neurons, the dependence on CF distance was small, and even under stimulus conditions, no strong dependence on CF distance was noted. The main changes in $R$ during spontaneous activity are within the first 0.2 mm, where the units have very similar CFs (Eggermont 1992) and when units are in different cortical areas even when they have the same CFs (Eggermont 2000). However, during stimulation, the $R$ for activity of units in different auditory fields is similar to that for units within an area, showing hardly any dependence on distance (Eggermont 2000). Within the confines of one electrode array (maximum spatial distance parallel to midline = 1.5 mm), the most variance was carried by the weighted STRF overlap. This overlap appeared not to depend on stimulus level; if anything, it decreased somewhat with level. The overlap was dependent on spatial distance but the mean values at 0- and 1.5-mm separation in the posterior-anterior direction were only 15% apart. At each distance, the distribution was normally distributed with both mean $\pm SD$ decreasing toward larger distances, likely because the overlap can only decrease and there is a lower bound.

Brosch and Schreiner (1999), in the only large parametric study on cross-correlation in primary auditory cortex beside our own previous studies, found that the width of the cross-correlogram peak was significantly dependent on spectral overlap for the multi-unit clusters and so was the peak correlation strength, but only during stimulation with noise bursts. As in Brosch and Schreiner’s study, we found a negative correlation between $R$ and correlogram peak width ($P < 0.0001$ for all conditions). Their data suggest that peak strength and peak width have different and, to some extent independent, roles in coding aspects of sound in auditory cortex. Our finding that the peak width is reduced for IN-STRF spike correlograms compared with silence and OUT-STRF spike correlograms corroborates their findings. The negative correlation of $R$ with peak-latency difference is also in accordance with Brosch and Schreiner’s findings as is our finding of lack of effect of the CF difference on the value of $R$, which we also showed previously (Eggermont and Smith 1996). What the present study adds is describing receptive field overlap in both spectral and temporal dimensions, whereas previously this was only defined on BW overlap, and demonstrating that there is a strong correlation with this weighted overlap measure for both spontaneous and stimulated correlations. We did not see a dependence of $R$ on the BW overlap for silence or OUT- or IN-STRF conditions, whereas Brosch and Schreiner only found a correlation under stimulus conditions; however, they found a strong effect of BW overlap on the correlogram peak width. We found a significant negative correlation between correlogram peak width and BW overlap for the silence and OUT-STRF conditions but not for the IN-STRF condition, suggesting that the type of stimulus (50-ms duration noise bursts repeated at a rate of 2.2/s vs. a 15-min multi-tone stimulus) has an effect on the cross-correlations or that the stimulus-correction procedure done in Brosch and Schreiner’s study somewhat affects the conclusions.

**BW overlap and overlap in STRF duration**

BW overlap and overlap in STRF duration did not depend on distance in the posterior anterior direction within an electrode array. This somewhat surprising result can be explained by the large range of these parameters at each distance. For instance, the BW overlap ranged from 0.06 to 2.9 octave for each of the distances; similarly the duration overlap ranged from 2 to 50 ms. The fact that the weighted STRF overlap did show a dependence on distance suggests that it is not equal to the product of BW overlap and duration overlap.

**Multiple single unit versus single unit**

We studied the cross-correlation between multiple single-unit activity, and likewise the STRFs were also based on more than one unit. We have previously looked at the similarity of CFs for single units recorded on the same electrode for 79 sites in AI (Eggermont 1996) and found that the mean difference was $0.15 \pm 0.24$ octave, which was not significantly different from zero. We have also previously (Valentine and Eggermont 2004) explored the difference in STRFs for single-unit versus multiple single unit for 85 multiple single-unit recordings with firing rates $>10$ spikes/s. This yielded 214 single units, and their STRFs had the same properties as those for the multiple single unit ones. Thus we do not think that the results would have been qualitatively different for single-unit data. The advantage of using multiple single units to construct the STRFs is that the signal-to-noise ratio is higher.

Evidence from neural modeling suggests that interpretation of neural correlations from multi-unit recordings may be ambiguous as it is not a linear combination of correlations for the various single-unit pairs (Bedenbaugh and Gerstein 1997; Gerstein 2000). However, changes in single-unit correlation strengths will be accompanied by comparable changes in the correlation between multi-unit activity as shown in Eggermont (2000). Thus if a number of single units contribute to each multiple-single-unit recording, the changes in the MSU $R$
values will be positively, but nonlinearly, related to those for the corresponding SU ones. So if single-unit correlations go up, so will MSU correlation. The only difference is that the MSU R values are larger than the corresponding single-unit ones but smaller than the sum of all the relevant single-unit pair R values.

STRF overlap and thalamic cell divergence

The STRFs of cortical cells are the direct result of converging thalamic input and shaped by intra-cortical inhibition (Miller et al. 2001). The weighted STRF overlap of the two cortical cells will reflect the fraction of the thalamic afferents that are shared by them. The R values, in addition, reflect the degree of synchrony in the cortical cells that likely is conveyed by the degree of synchrony in the shared thalamic afferents. This, as we have seen, depends on the fraction of shared collaterals, the activity of which is by definition synchronous but also by the amount of synchrony imposed by stimulation or, to some extent, by network activity such as burst-firing and spindling. Not surprisingly, the R values do strongly depend on the STRF overlap.

STRFs for coincident spikes

Studies comparing the receptive fields of correlated neural activity between cortical neurons with those of individual cells are rare; the only other study that we are aware of is that of Ghose et al. (1994) in visual cortex. Just as in the present study, they found that the joint receptive field for coincident spikes was smaller than that for the constituting neurons, even though the single units were recorded on the same electrode. Ghose et al. (1994) used a reverse correlation technique to calculate the average STRF before a “coincident” spike. As we have shown previously, the reverse correlation approach and the forward correlation approach for the multi-tone stimuli used here are essentially the same in the spectral domain, but the STRF is smoothed in the temporal domain in the reverse correlation approach as a result of the stimulus envelope (Valentine and Eggermont 2004). Given that the latencies and width of the correlograms are less in auditory cortex compared with visual cortex, albeit that the most common peak width was 10 ms in both studies, we only used coincident-spike windows of 10 and 2.5 ms, whereas Ghose et al. (1994) used 8–64 ms. They did not see a quantitative effect between the window lengths, whereas we did see a small but consistent contraction of the STRF when the cross-interval window was reduced from 10 to 2.5 ms. So despite that we did not use a stimulus-correction procedure and our joint STRFs could be the byproduct of stimulus locked discharges, we arrived at largely the same conclusions. Namely, that the joint STRF is more confined than that of the contributing neurons by being reduced in both the spectral (spatial) and temporal dimensions. If we assume that a joint STRF is the result of reading the input of two neurons, or in our case two MSU-clusters, through a coincidence operation by a putative downstream neuron, then one can argue again that all stimulus-evoked spikes are important and not just the ones that remain after a stimulus-correction procedure.

When using a tone-pip stimulus with Poisson distributed rates for each frequency, it happens frequently, even if the presentation rates of the 81 individual frequencies are as low as 0.25 Hz, that tone pips overlap in time. For instance, the probability of two overlapping tone pips, as calculated from the Poisson distribution, is equal to 0.184, the probability of three overlapping tone pips is 0.061, and for four overlapping tone pips, it is only 0.015. If the responses to the tone pips were independent of the presence of other tone pips, then the joint coincident-spike STRF of two units would be the union of the two individual receptive fields. However, if coincident firing occurred only for particular frequencies common to both STRFs, then the joint coincident-spike STRF would be the overlap between the two individual STRFs. We have found that the pair-wise coincident-spike STRFs are somewhat in between these two extremes; however, the STRF region that is represented in all pair-wise coincident-spike STRFs for a given electrode array is by definition this overlap region. The window chosen for coincident firing strictly determines the temporal extent of the coincident-spike STRFs, but the frequency extent is closer to the union of that for the individual STRFs. Consequently, the average weighted overlap for the coincident-spike STRFs is increased compared with that for all spikes but still ranges from 0.3–0.4 upward, depending on the window size. This suggests that coincident spikes are responsible for those parts of the STRF with the highest firing rates. This holds clearly for the pair-wise calculations and generally also for the all-electrode coincident-spike STRF, albeit that for one of the electrodes in our example (Fig. 12), this region was located outside that of the largest response. It is likely that such a unit’s presence in a neural assembly will not be weighted heavily. Thus a coincident assembly of neurons would be able to focus on the STRF centers, thereby improving the signal to noise ratio. The independence of the coincident spike STRF overlap with the distance between unit pairs recorded with the same electrode array could also allow processing of complex sounds by a coincident-firing neuronal assembly that covers a relatively large part of primary auditory cortex.

ACKNOWLEDGMENTS

A. Norella assisted with the data collection and provided helpful discussions. G. Shaw provided programming assistance.

GRANTS

This investigation was supported by grants from the Alberta Heritage Foundation for Medical Research, the Natural Sciences and Engineering Research Council of Canada, and the Campbell McLaurin Chair for Hearing Deficiencies.

REFERENCES


A. Norella assisted with the data collection and provided helpful discussions. G. Shaw provided programming assistance.

This investigation was supported by grants from the Alberta Heritage Foundation for Medical Research, the Natural Sciences and Engineering Research Council of Canada, and the Campbell McLaurin Chair for Hearing Deficiencies.


