

Online Stimulus Optimization Rapidly Reveals Multidimensional Selectivity in Auditory Cortical Neurons

Anna R. Chambers,^{1,2} Kenneth E. Hancock,^{1,3} Kamal Sen,⁴ and Daniel B. Polley^{1,3}

¹Eaton-Peabody Laboratory, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts 02114, ²Program in Neuroscience, Harvard Medical School, Boston, Massachusetts 02115, ³Department of Otolaryngology, Harvard Medical School, Boston, Massachusetts 02115, and ⁴Hearing Research Center, Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215

Neurons in sensory brain regions shape our perception of the surrounding environment through two parallel operations: decomposition and integration. For example, auditory neurons decompose sounds by separately encoding their frequency, temporal modulation, intensity, and spatial location. Neurons also integrate across these various features to support a unified perceptual gestalt of an auditory object. At higher levels of a sensory pathway, neurons may select for a restricted region of feature space defined by the intersection of multiple, independent stimulus dimensions. To further characterize how auditory cortical neurons decompose and integrate multiple facets of an isolated sound, we developed an automated procedure that manipulated five fundamental acoustic properties in real time based on single-unit feedback in awake mice. Within several minutes, the online approach converged on regions of the multidimensional stimulus manifold that reliably drove neurons at significantly higher rates than predefined stimuli. Optimized stimuli were cross-validated against pure tone receptive fields and spectrotemporal receptive field estimates in the inferior colliculus and primary auditory cortex. We observed, from midbrain to cortex, increases in both level invariance and frequency selectivity, which may underlie equivalent sparseness of responses in the two areas. We found that onset and steady-state spike rates increased proportionately as the stimulus was tailored to the multidimensional receptive field. By separately evaluating the amount of leverage each sound feature exerted on the overall firing rate, these findings reveal interdependencies between stimulus features as well as hierarchical shifts in selectivity and invariance that may go unnoticed with traditional approaches.

Key words: auditory cortex; closed-loop; genetic algorithm; online neural characterization; optimization algorithm; sparse coding

Introduction

Sensory neurophysiologists face a dilemma when choosing a stimulus feature space to explore in an experiment. Because of what is commonly referred to as the “curse of dimensionality” (Bellman, 1961), a space that captures the full range of complexities of high-dimensional, naturally occurring sensory events is intractably large, and an exhaustive search of this space is not feasible. Instead, most experimenters choose to vary one or two stimulus dimensions while holding others at fixed, predetermined values. Whereas neurons at each successive stage of the visual cortical hierarchy have been shown to operate on stimulus features of increasing complexity and abstraction, auditory cortical neurons are driven by both simple and complex sound features that vary across all possible derivations of spectral, tem-

poral, intensive, and dichotic variations in the sound pressure waveform (Wollberg and Newman, 1972; Wang et al., 1995; deCharms et al., 1998; Bizley et al., 2009; Carruthers et al., 2013). Thus, stimulus dimensionality reduction is particularly problematic for studies of sensory coding at higher levels of the central auditory pathway where neurons are driven by a restricted and unpredictable set of sound features within a vast parameter space (Hromádka et al., 2008; Schneider and Woolley, 2011).

Multiparametric stimulus search paradigms that are effective and efficient, yet as unbiased and inclusive as possible, are of great interest to the field (Edin et al., 2004; Benda et al., 2007; Lewi et al., 2011; DiMattina and Zhang, 2013). To this end, a variety of approaches for neural characterization have been described that provide alternatives to probing receptive field organization with discrete, arbitrary stimuli. These procedures aim to fully characterize the stimulus–response relationship via online system identification (Wu et al., 2006; Lewi et al., 2009; DiMattina and Zhang, 2011; Lewi et al., 2011) or are guided by specific aspects of the neural input–output function, such as mutual information (Machens, 2002; Sharpee et al., 2004; Lewi et al., 2011) and spike count (Nelken et al., 1994; Bölinger and Gollisch, 2012).

Inspired by an approach to study complex shape encoding in the visual cortex (Yamane et al., 2008), we implemented a rapid, reliable, and computationally inexpensive stimulus evolution technique to tailor stimulus features to the neural receptive field

Received Jan. 20, 2014; revised May 8, 2014; accepted May 20, 2014.

Author contributions: A.R.C., K.E.H., and D.B.P. designed research; A.R.C. performed research; A.R.C., K.E.H., and K.S. contributed unpublished reagents/analytic tools; A.R.C. analyzed data; A.R.C. and D.B.P. wrote the paper.

This work was supported by National Institutes of Health Grant R01 DC009836 to D.B.P. and Grant P30 DC05209. We thank Professors Jennifer Bizley and Stephen David for providing comments on an earlier version of the manuscript and Ross Williamson for providing analysis code and stimuli for the STRF method.

The authors declare no competing financial interests.

Correspondence should be addressed to either Dr. Anna R. Chambers or Dr. Daniel B. Polley, Eaton-Peabody Laboratory, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, MA 02114, E-mail: chamber3@fas.harvard.edu or Daniel_polley@meei.harvard.edu.

DOI:10.1523/JNEUROSCI.0260-14.2014

Copyright © 2014 the authors 0270-6474/14/348963-13\$15.00/0

according to real-time variations in firing rate. As applied here, the closed-loop optimization technique aimed to characterize the relationship between firing rate and stimulus optimality in the primary auditory cortex (A1) and central nucleus of the inferior colliculus (ICc) of awake animals. There has been much debate regarding the nature of cortical representations of optimized stimuli, with both sustained, high firing rate responses and low rate, highly precise responses reported in the literature (DeWeese et al., 2003; Wang et al., 2005; Hromádka et al., 2008). We sought to provide a specific test of whether stimulus preference is represented by the disproportionate growth of spiking during the steady-state portion of the neural response. In addition, we used this approach to illuminate whether and how stimulus invariance and selectivity changed across distinct nodes of the sensory processing hierarchy.

Materials and Methods

Surgical procedures. All procedures were approved by the Animal Care and Use Committee at Massachusetts Eye and Ear Infirmary and followed the guidelines established by the National Institutes of Health for the care and use of laboratory animals. Male CBA/CaJ mice 8–10 weeks of age were brought to a surgical plane of anesthesia with ketamine/xylazine (induction with 120 mg/kg ketamine and 12 mg/kg xylazine, with 60–80 mg/kg supplements as necessary). The core body temperature of the animal was maintained at 36.5°C with a homeothermic blanket system. Using a scalpel, a small craniotomy was centered over right primary auditory cortex or right central nucleus of the inferior colliculus, leaving the dura mater intact. The brain surface was covered with sterile ointment. Chronic implants consisted of multichannel silicon probes (177 μm^2 contact area, 100 μm contact separation; NeuroNexus Technologies) arranged in a 4×4 configuration mounted onto a bidirectional microdrive (A1) or 16×1 stationary probe configuration (all ICc, some A1). The A1 implant was positioned by first mapping the cortex to delineate the low-high-low caudal-rostral best frequency gradient that uniquely identifies the orientation of A1 and the anterior auditory field (Hackett et al., 2011; Guo et al., 2012). The ICc was identified based on a low-to-high best frequency gradient along the dorsal-ventral axis of a single-shank multichannel probe. A1 and ICc implants were performed in separate groups of mice. Bone wax (3M) was packed around the margins of the craniotomy to protect the probes and brain surface, and the microdrive (for A1 implants) and headstage connector were affixed to the skull surface using acrylic bonding material (C&B MetaBond, Parkell). Ground wires were implanted outside of the auditory cortex, and the animal was given postsurgical subcutaneous injections of buprenorphine (0.05 mg/kg) and saline (0.5 ml) to reduce pain and dehydration, respectively. A clear plastic cylinder was affixed to the perimeter of the A1 craniotomy to protect the probes and brain surface.

Recording procedures. At least 48 h after the implant surgery, the animal was placed in a small, acoustically transparent chamber built to be just large enough for its body (3.5 cm \times 7 cm). The chamber was tapered at one end to encourage the mouse to face a consistent direction, and was too narrow to allow the mouse to turn its head by more than a few degrees in either direction. The mouse was continuously video monitored during recording to ensure that it was still and its eyes open, and recording was paused if the mouse reared or moved its head position from the target position. Cortical probes were moved slightly (~ 25 – $50 \mu\text{m}$ increments) to increase single-unit yield between recording sessions. Single units were characterized with the adaptive online procedure if they exhibited a response above spontaneous rate to any of 633 randomly selected sound stimuli. All neurons tested with the evolutionary procedure were included in the dataset, regardless of efficacy. Raw signals were digitized at 32 bit, 24.4 kHz (RZ5 BioAmp Processor; Tucker-Davis Technologies) and stored in binary format. Single units were isolated online and then further denoised offline, as necessary, using a PCA-based sorting method in SpikePac software (Tucker-Davis Technologies).

Acoustic stimuli. Stimuli were generated with a 24 bit digital-to-analog converter (National Instruments, model PXI-4461) and presented via

four free-field electrostatic speakers (Tucker-Davis Technologies) positioned around the mouse restraint chamber. Stimuli were calibrated before recording sessions using a wide-band ultrasonic acoustic sensor (Knowles Acoustics, model SPM0204UD5). Other than the frequency response areas (FRAs) and spectrotemporal receptive fields (STRFs), stimuli were 400 ms in duration with a 600 ms interstimulus interval and 4 ms raised cosine onset/offset ramps.

Simulated receptive fields. For each simulated neuron, five tuning functions were chosen with replacement from a randomly varied set of five tuning “types” as follows:

1. Sigmoid

$$F(x) = \frac{1}{1 + e^{-kx}}$$

Where k was drawn randomly from the interval from -1 to 1 .

2. Gaussian

$$F(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Where μ and σ were drawn randomly from the interval from 1 to 20 and 1 to 5 , respectively.

3. Difference of Gaussians

$$F(x|\mu, \sigma_1, \sigma_2) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma_1^2}} - \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma_2^2}}$$

Where σ_1 and σ_2 were drawn randomly from the interval from 1 to 5 .

4. Sum of Gaussians

$$F(x|\mu_1, \mu_2, \sigma_1, \sigma_2) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} + \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}$$

Where μ_1 and μ_2 were drawn randomly from the interval from 1 to 20 , and σ_1 and σ_2 were drawn randomly from the interval from 1 to 5 .

The final tuning type was flat tuning (no change in response). Twenty samples were taken along each feature dimension. A maximal firing rate, chosen at random from a range typical of A1 single unit firing rates (~ 20 – 100 Hz), was selected for each modeled neuron. Tuning functions evaluated at each sample represented the neuron’s fractional response of maximal firing rate (the peak of each tuning function was 1), and the fractional responses at each feature dimension were multiplied together to reveal the firing rate evaluated at that intersection of features (i.e., the predicted firing rate to each stimulus). Spontaneous responses with random noise were also added before calculating the final response magnitude of the model neuron to each stimulus.

Simulating conventional methods of stimulus search. For one- and two-dimensional stimulus optimization, three or four stimulus features were held constant at a random, arbitrary value, whereas the remaining dimensions were varied along their entire sampling range (20 samples). Simulated neuron responses were calculated at each sample, taking the values of the other (constant) stimulus dimensions into account. The maximal response from this procedure was then normalized to the response of the neuron to its globally optimal stimulus. To simulate an iterative optimization method, we first ordered the stimulus dimensions arbitrarily for each neuron. Then, starting with the first stimulus dimension, we varied one dimension at a time along its entire sampling range and chose the “optimal” value for that dimension according to the model neuron response, whereas other, nonoptimized dimensions were held at an arbitrary constant (Wang et al., 1995). This response was calculated as described above: each tuning function evaluated at a specific intersection of feature dimensions revealed a fractional response of maximum for each model neuron. The fractional responses were multiplied together to

create the final fractional response of maximum, and this number was multiplied by the model neuron's maximal firing rate to reveal the response to each stimulus. The tested dimension was held at its optimal value, and we then iteratively optimized the remaining dimensions until peaks along every dimension were defined. Although this procedure is multidimensional in nature, the stimulus dimensions are varied one at a time rather than simultaneously, and it is nonexhaustive in that most stimulus feature combinations are not tested.

Evolutionary stimulus search procedure. A stimulus set varying along five acoustic parameters: center frequency (4–64 kHz in 0.1 octave increments), level (10–60 dB in 10 dB increments), spectral bandwidth (pure tone, 1.25 octave band in 0.25 octave increments), sinusoidal amplitude modulation frequency (0.25 octave [0 Hz], 70 Hz in 10 Hz increments), and speaker location (all permutations of left, right, top, center), yielded a pool of 177,120 potential stimuli. Thus, each individual stimulus was a sinusoidally amplitude-modulated sound (unless amplitude modulation of zero happened to be selected) whose carrier was either a pure tone or a band-limited noise of a particular spectral bandwidth. Each run of the adaptive search procedure was initialized with 50 stimuli from this pool, selected at random. Firing rate responses (the average of two repetitions) were calculated during the stimulus window (0–400 ms), and all stimuli played in the experiment thus far were rank-ordered at the end of each generation. At the conclusion of one generation, all stimulus-evoked spike counts were rank-ordered and the top 10 most effective stimuli overall were used as “breeders” for the next generation. The “offspring” for each breeder stimulus was created by randomly shifting one or more acoustic parameters to its nearest-neighbor value (e.g., if the level of a breeder stimulus was 40 dB, its offspring could have a level of 30, 40, or 50 dB, as the sampling density for level was 10 dB). After the first generation, 80% of stimuli were chosen randomly from all possible offspring identified by the evolutionary algorithm, whereas 20% of stimuli were chosen randomly from the entire acoustic feature space to avoid focusing on local maxima and mitigate potential decreases in firing rate response magnitude resulting from adaptation. The most effective stimulus from the first generation (the “yardstick” stimulus) was repeated in every subsequent generation to estimate the overall response stability across generations. Most A1 single units were subjected to two separate runs of the procedure with independently seeded starting conditions to estimate the reliability of the evolutionary procedure.

Movement artifact rejection. Two criteria were used to safeguard against contamination of neural responses by infrequent movement artifacts. First, a global ceiling on firing rate was set (200 Hz for cortex, 400 Hz for midbrain). Any stimuli corresponding to neural responses exceeding this ceiling were excluded from the dataset. Second, responses to stimuli across two repetitions were compared. If the neural responses to the two repetitions differed by >25 Hz, at least one presentation was considered an artifact and the stimulus was excluded from the dataset.

FRAs. In a subset of single units ($n = 12$), FRAs were measured to cross-reference the “best” features selected by the adaptive search procedure. FRAs were measured with pseudorandomly presented tone pips (50 ms duration, 4 ms raised cosine onset/offset ramps, 0.5–1 s intertrial interval) of variable frequency (4–64 kHz in 0.1 octave increments) and level [0–60 dB sound pressure level (SPL) in 5 dB increments]. A total of 533 unique frequency–level combinations were presented twice for a given single unit. FRA analysis was performed as outlined previously (Guo et al., 2012). Briefly, a poststimulus time histogram (PSTH) was calculated with 1 ms bin size and windowed to identify sound-driven spikes. Response onset was the point at which firing rate exceeded the spontaneous rate by at least 4 SDs. The offset was set at the point where firing rate decreased to <5 SDs above the spontaneous rate. The best frequency was associated with the highest spike count summed across all sound levels.

STRFs. In a subset of single units, STRFs were measured for cross-referencing purposes. A 60-s-long dynamic random chord stimulus was used, which consisted of tone pulses (20 ms duration, 5 ms raised cosine onset/offset ramps) randomly and independently chosen from 20 ms bins in time, 1/12 octave bins covering 4–64 kHz in frequency, and presented from 20 to 75 dB in 5 dB bins. For each single unit, the stimulus was repeated 20–40 times. Automatic Smoothness Determination was used to improve estimates (Sahani and Linden, 2003; Sahani et al., 2013),

in which Bayesian techniques are used to obtain optimal spectral/temporal smoothing and scaling parameters for the neural data. To estimate a functional transformation of stimulus input that would yield the recorded neural response, a discrete-time Wiener filter is usually used with reverse correlation (Aertsen and Johannesma, 1981). However, regression combined with Bayesian smoothing and scaling parameters avoids the overfitting to noise typical of the Wiener filter. We did not use a static nonlinearity to determine spike rate from the STRF.

Radial visualization method. The RADVIZ method is a dimensionality reduction technique meant for plotting high-dimensional data onto radial axes (Ankerst et al., 1996). Each stimulus is represented as a point on a radial plane, which is connected to five equally spaced anchor points on the perimeter of the plane by five springs. Spring constants k are determined by the five acoustic parameters of the stimulus. The position of the point is determined by the equilibrium position of the connected springs. The radial plane was superimposed onto a hue/saturation color map so that the position of the point also represented a color.

Sparseness metric. Sparseness (Willmore et al., 2011) of a response distribution was defined as follows:

$$S = 1 - \frac{E[r]^2}{E[r^2]}$$

where $E[\cdot]$ is the expected value. Using this definition, sparseness equals zero for a dense code and 1 for a sparse code. An intuitive definition for lifetime sparseness would be the peakedness of a response distribution of each neuron, where lifetime sparseness would be high for a neuron that was silent for most stimuli but occasionally exhibited high response rates. For each neuron, a lifetime sparseness metric, S_L , was obtained by calculating the sparseness of its responses to all randomly presented stimuli. These lifetime sparseness estimates were then averaged across all n units in the A1 or IC population. For the simulations, sparseness was calculated according to the responses of simulated neurons to 300 randomly chosen stimuli from the 5-dimensional feature space.

$$\bar{S}_L = \frac{1}{n} \sum_{i=1}^n S_{Li} = \frac{1}{n} \sum_{i=1}^n \left[1 - \frac{E[r_{ij}]^2}{E[r_{ij}^2]} \right]$$

In contrast to the work of Willmore et al. (2011), the stimuli used to elicit responses that went into the lifetime sparseness calculation described here were randomly chosen and were not tailored to the receptive field of the neuron. Therefore, sparseness values may be skewed upward because of the lack of a prior probability of eliciting a spike and may not be suitable for direct comparison.

Results

Low-dimensional stimulus optimization results in lower probabilities of maximizing model neuron responses, even when responses are dense

To test the effectiveness of various conventional and adaptive stimulus search methods, we first simulated the responses of generic neurons with tuning across five arbitrary stimulus dimensions (Fig. 1). Receptive field properties of the simulated neurons were determined by randomly setting the selectivity and shape of the tuning function (see Materials and Methods) with replacement from the following list: sigmoidal, Gaussian, difference of Gaussians, sum of Gaussians, and flat (no tuning). Figure 1A shows the 5-dimensional tuning for two model neurons that exemplify dense or sparse responses (see Materials and Methods). The tuning functions for each parameter were then combined in a multiplicative manner to create the 5-dimensional receptive field of each model neuron.

Nonexhaustive stimulus search was simulated with a conventional strategy in which one or two stimulus dimensions were varied whereas all others were held at arbitrary constant values (Fig. 1B, left and middle). This strategy is common in neurophysiological characterizations in which the representation of a small number of stimulus features is the focus of the study. However, in

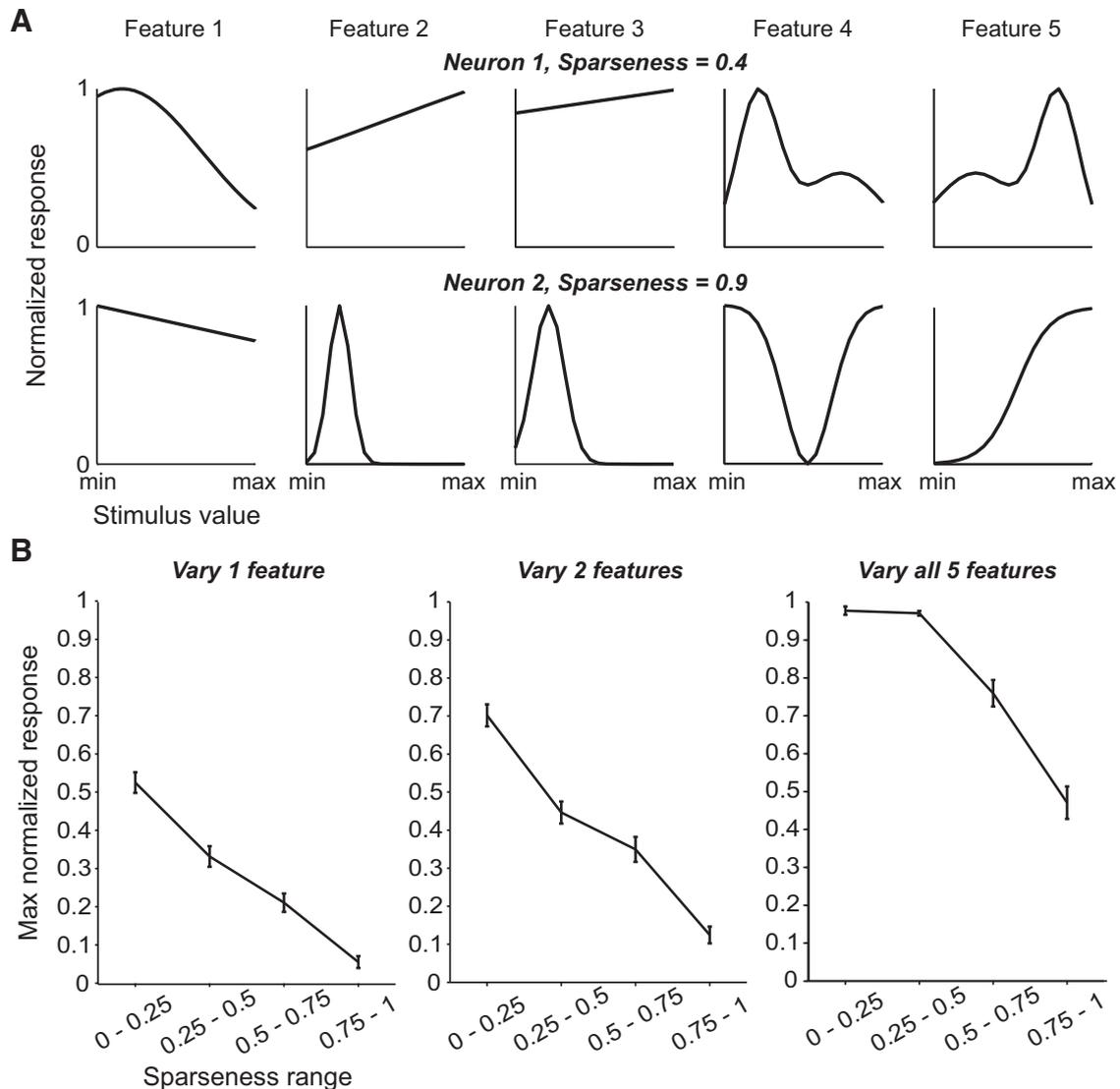


Figure 1. Varying more parameters in the search paradigm increases the likelihood of obtaining maximal responses from simulated neurons. **A**, Five tuning functions, chosen with replacement from a list of five possibilities (Gaussian, Sigmoid, Sum of Gaussians, Difference of Gaussians, No Tuning), were assigned to each model neuron to represent the change in response across variations of five arbitrary features. Shown are the five tuning functions for two example model neurons, with differing sparseness (for receptive field model parameters, see Materials and Methods). **B**, Three stimulus search strategies commonly used in sensory neurophysiology experiments were tested for their efficacy at eliciting maximal responses from model neurons. In the first strategy (left), one feature (chosen at random from five total features in the stimulus space) was varied, whereas the other four features were held at an arbitrary constant (randomly chosen for each feature). The same procedure was performed for varying two features (middle). Finally, all features were varied iteratively, in a random order (right).

model neurons with complex, multidimensional feature selectivity and sparse responses, meant to simulate receptive fields in higher-order cortical areas, these stimulus search paradigms often fail to capture the full range of responsiveness of the neuron. Even in cases of low sparseness, where neurons are likely to respond robustly to more stimuli, the one- and two- dimensional search techniques are insufficient to maximize model neuron firing rates.

Another method common to neurophysiological experiments is to iteratively optimize multiple stimulus features in a fixed order. Successive low-dimensional optimizations may work well for neurons with certain types of receptive fields, particularly if feature tuning is combined in a linear fashion to determine the neuron's output. Indeed, this method has been shown to be successful at eliciting sustained responses from auditory cortical neurons (Wang et al., 2005). However, because sound features are not varied simultaneously, nonlinear interactions between stimulus features may be missed. In our simulation, this method

was successful in eliciting maximal responses from model neurons with low to moderate sparseness values (Fig. 1B) but was not reliably effective for the most sparse responses. Depending on the particular application, neurophysiology experiments often set stimulus values that are not explicitly the focus of the experiment at constant, although not arbitrary, values. The simulations shown here represent a generic dataset outside of any particular sensory modality. Nevertheless, they demonstrate that, even in cases of low sparseness, where neurons are likely to respond robustly to more stimuli, the one- and two- dimensional search techniques are insufficient to maximize model neuron firing rates.

A nearest-neighbor stimulus search strategy successfully optimizes stimuli along five dimensions for simulated neurons

Given that conventional, low dimensional, or iterative search methods could not reveal sparse responses in model neurons, we

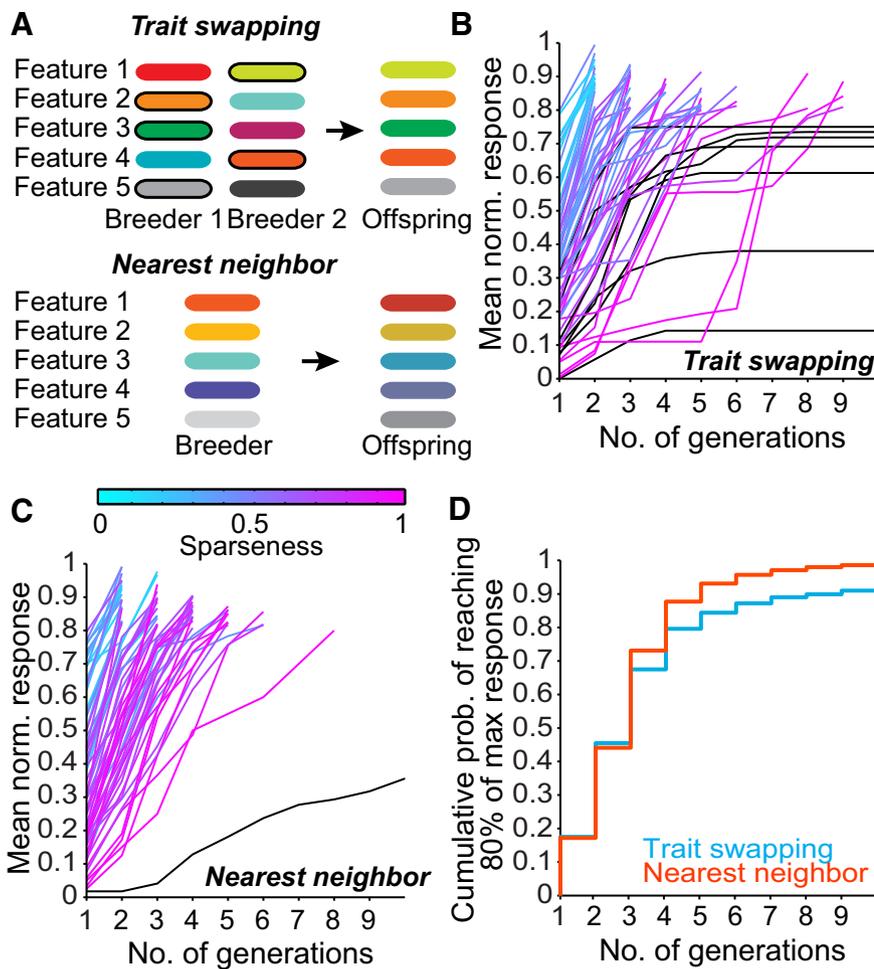


Figure 2. Two closed-loop, evolutionary search strategies identify maximum response regions with differing degrees of efficiency. A closed-loop evolutionary stimulus search was simulated (**A**) using two different evolutionary algorithms. In both cases, an initial generation of 50 stimuli was randomly chosen from the 5-dimensional stimulus space. In subsequent generations, 10 stimuli were chosen at random and 40 stimuli were chosen based on one of the two evolutionary strategies. In the first strategy (“trait swapping”), pairs of “breeder” stimuli were chosen from the top 10 stimuli of the previous generation, ordered by response magnitude. For each stimulus feature, the child stimulus inherited its trait from one of the two parents, chosen at random (black outline). In the second strategy (“nearest neighbor”), a 5-dimensional cloud of nearest neighbors was defined around each breeder stimulus, and offspring traits were chosen at random from this cloud. **B, C**, Mean response magnitude of the breeder stimuli at each successive generation, for 80 runs each of the two search strategies. The algorithm ran until a mean breeder response magnitude of ≥ 0.8 was reached. Black lines indicate runs where the criterion was not met after 10 generations. Color scale indicates sparseness. **D**, Cumulative probability distribution of the number of generations necessary to reach the criterion for both search strategies, for 1000 runs of each.

sought to implement a closed-loop, adaptive search technique that would simultaneously vary at least five stimulus dimensions while avoiding null response regions of stimulus space and effectively capturing maximal responses. Beginning with a set of 50 randomly seeded stimuli, model neurons were tested with an evolutionary search procedure meant to resemble processes of natural selection and social behavior (Bleeck et al., 2003; Kennedy and Eberhart, 1995). In each subsequent generation, 10 stimuli were chosen at random and 40 stimuli were chosen based on model neuron feedback, using one of two potential strategies. In the first strategy (“trait swapping,” Fig. 2A, top), “offspring” stimuli were created from a combination of traits from two breeder “parents.” Breeding pairs were selected at random from the top 10 most effective stimuli. The “genetic penetrance” was set to 100% for each stimulus dimension, meaning that each trait in the offspring was copied from one of the two parent stimuli at random. In the second strategy (“nearest neighbor,” Fig. 2A, bot-

tom), a 5-dimensional cloud of nearest neighbors to each breeder stimulus was defined, and each offspring stimulus dimension was set to one of the neighboring values for the corresponding stimulus.

After each generation of 50 stimuli, the stimuli were reordered to update the 10 most effective overall breeders, which were then used to seed the next generation. The simulations ran until the mean response magnitude of the breeder stimuli reached at least 80% of the model neuron’s maximal response magnitude. In most runs of the simulation, this was accomplished in ≤ 10 generations. However, the trait swapping approach produced more runs that failed to reach the criterion level of effectiveness within 10 generations (Fig. 2B, black lines), indicating that the algorithm may have perseverated in a local maximum of the simulated neuron’s receptive field. Overall, the nearest neighbor algorithm was able to identify effective regions of the 5-dimensional response manifold more reliably and rapidly as determined from 1000 simulated neurons equally distributed along the sparseness axis ($p < 0.01$, two-sample Kolmogorov–Smirnov test; Fig. 2C,D).

Evolutionary stimulus search rapidly shifts the distribution of firing rate responses of A1 neurons toward maximal observed values

Having tested our stimulus evolutionary procedure *in silico*, we then performed extracellular single-unit recordings from the primary auditory cortex of awake mice to confirm its utility *in vivo*. We populated the sensory feature space with acoustic parameters to which cortical neurons have been shown to display selectivity: carrier frequency, spectral bandwidth, intensity, amplitude modulation rate, and location (Merzenich et al., 1975; Phillips and Irvine, 1981; Schreiner and Mendelson, 1990; Middlebrooks et al., 1994; Read et al., 2001; Bizley et al., 2009; Yin et al., 2011; David and Shamma, 2013) (Fig. 3A). Upon permuting across a typical range and sampling resolution, this created a set of $\sim 175,000$ potential stimuli. When stimuli were selected at random, A1 single-unit responses (Fig. 3B, top; normalized to peak firing rate observed for the unit) were heavily skewed toward minimal or no response; high firing rates were extremely rare (Fig. 3B, bottom, C).

Beginning with a randomly seeded set of 50 stimuli, neurons were tested with the nearest-neighbor algorithm of the evolutionary search procedure (Fig. 3D). Each stimulus was 400 ms long with 600 ms between trials, such that six generations of the evolutionary search procedure would take 10 min. From our simulation experiments, six generations often elicited maximal responses from even the most sparse model neurons (Fig. 2C), and 10 min was well within the time frame of single unit experiments *in vivo*. For each subsequent generation of 50 stimuli, 39

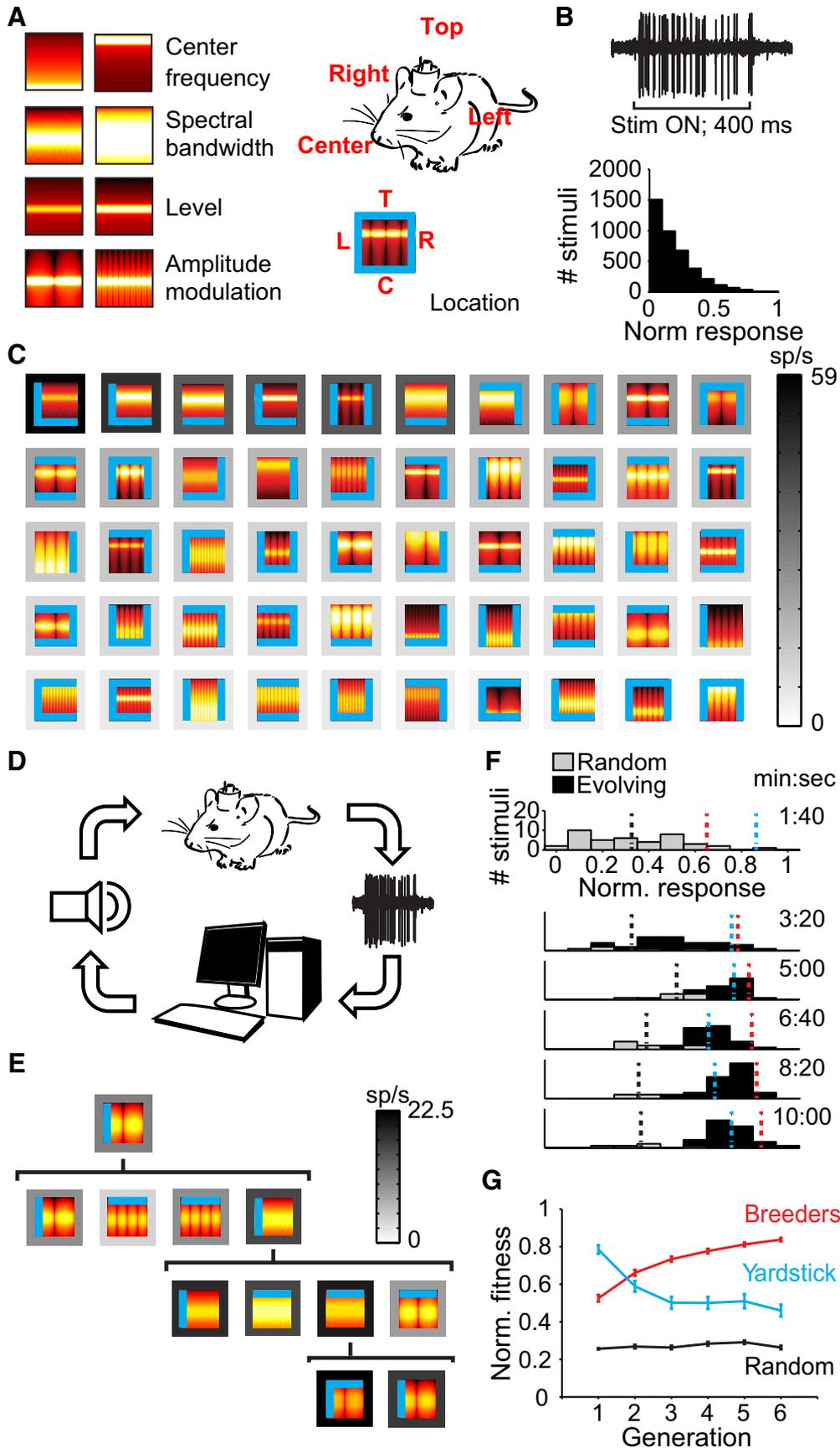


Figure 3. Adaptive stimulus search rapidly shifts the distribution of A1 single-unit responses toward high firing rates. **A**, The relevant stimulus space was constructed using five basic acoustic features, shown in the thumbnails from their minimum (left) to maximum (right) values in the experiment. A diagram shows the location of the four speakers relative to the animal's head. For location (right bottom), the presence of a blue line indicates an active speaker. **B**, Example SU recording trace from A1. Bottom, Distribution of normalized firing (Figure legend continues.)

offspring evolved from the top 10 fittest breeders. Ten stimuli were chosen at random to avoid perseverating in local peaks within the preferred stimulus manifold, and increase the likelihood of arriving at a globally optimal stimulus. A single “yardstick” stimulus was repeated in each subsequent generation to gauge the level of response adaptation. Firing rate responses were averaged across two repetitions of each stimulus. Figure 3E shows an example of an evolving stimulus subset from one breeder and the progressive growth of sound-evoked spike rate across subsequent generations of its offspring.

Firing rate adaptation has been identified as a potential hindrance to online stimulus optimization (DiMattina and Zhang, 2013). The tendency for neural responses to lose sensitivity for similar stimuli repeated over time could interfere with the adaptive procedure, especially in later generations when the algorithm might have converged on a narrow region of the stimulus manifold. In our experiments, however, the mean firing rate to each generation of breeder stimuli increased (Fig. 3F, red line) despite the concurrent growth of response adaptation (Fig. 3F, blue line), reflected in the tendency of neuronal firing rate to decrease over repeated presentations of the “yardstick” stimulus. This trend was observed in the group data, where responses to the same stimulus adapted over generations (repeated-measures ANOVA, $F_{(5,50)} = 26.77, p < 0.001$; Fig. 3G, blue line), and there was a significant increase in responses to breeder stimuli over generations ($F_{(5,50)} = 173.1, p < 0.0001$; Fig. 3G, red line), whereas responses to randomly selected stimuli did not change over generations ($F_{(5,50)} = 0.73, p = 0.6$; Fig. 3G, black line).

The evolutionary procedure consistently converged onto an effective stimulus set across separate runs initialized with distinct random stimuli

In a subset of neurons ($n = 31$), the evolutionary procedure was run twice to verify convergence onto a reliably effective stimulus subset. A radial visualization technique (see Materials and Methods) assigned each stimulus a color based on its five parameter values (Fig. 4A). Stimuli from two runs of the procedure for each of three example A1 single units are ranked according to firing rate (Fig. 4B–D). Each bar represents a single stimulus whose height and color represent its firing rate response and stimulus attributes, respectively.

←

(Figure legend continued.) rates of A1 single units ($n = 50$) to all randomly chosen stimuli from a 5-dimensional feature space. Top, **C**, Example responses of one A1 unit to 50 randomly chosen stimuli. Each thumbnail represents one stimulus: spectrogram represents frequency content (y -axis), temporal modulation (x -axis), and level (brightness). Blue lines indicate location of speakers playing the stimulus: left, top, right, and center speakers (clockwise from left side of thumbnail). Grayscale background represents the unit firing rate to the corresponding stimulus. Thumbnails are rank-ordered by firing rate response. **D**, Diagram of closed-loop stimulus search procedure, adapted from Benda et al. (2007). **E**, Example of a breeder stimulus, initially selected at random (top thumbnail, same schematic as in **C**), which evolved to elicit higher firing rates in subsequent iterations (bottom 3 rows) by adjusting features to nearest neighbor values. **F**, Distribution, per generation (from top to bottom, generations 1–6) of normalized firing rate responses of an A1 single unit to randomly chosen (gray bars) versus evolving (black bars) stimuli. Elapsed time is indicated to the right of each histogram. Top histogram, First generation, comprised entirely of randomly chosen stimuli. Dashed lines indicate mean firing rate responses to random (black), breeder (red), and yardstick (blue) stimuli, per generation. **G**, Mean normalized fitness (firing rate) \pm SEM per generation for randomly chosen stimuli, breeder stimuli (top 10 stimuli), and yardstick stimuli (top stimulus from first generation, repeated in each subsequent generation). Despite significant adaptation of firing rate to effective stimuli (blue line), fitness of the breeder stimuli increases across generations (red line) and remains significantly higher than randomly chosen stimuli (black line).

These examples convey three important points about the evolutionary procedure and its effectiveness. First, color converges strongly toward the right side of each bar plot, indicating that, within each run, maximum firing rates tended to be elicited by very similar stimuli. Second, color convergence between two consecutive runs that were initially seeded with independent, random stimuli converged onto highly similar effective stimuli, indicating that the evolutionary procedure was reliable at identifying highly effective regions of the receptive field for each neuron. Third, performing separate runs with highly similar convergence confirmed that, instead of focusing on local maxima, the procedure truly converged on regions of receptive field space that reliably drove A1 neurons at maximal firing rates.

We quantified the degree of convergence by independently correlating the averaged parameters of the top 10 stimuli from each of two runs (Fig. 4E). Significant correlations are noted for the comparison of each stimulus property ($r > 0.43, p < 0.05$ for all; as a categorical variable, stimulus location was not amenable to parametric correlation).

The stimulus search procedure shows agreement with other well-tested methods but allows for simultaneous testing of variation along many feature dimensions

Stimulus preference defined by the evolutionary procedure (Fig. 5A) was cross-validated against pure tone receptive field estimates (Fig. 5B) and the STRF obtained with a dynamic random chord stimulus (deCharms et al., 1998; Linden et al., 2003) ($n = 20$; Fig. 5C). We observed a substantial agreement in preferred frequency between the FRA, STRF, and evolutionary search (Fig. 5A–C), across the sample (FRA/ES: $r = 0.98, p < 0.01$; STRF/ES: $r = 0.96, p < 0.01$; Fig. 5D). Importantly, though, multidimensional search consistently elicited higher peak firing rates than a pure tone at the best frequency for every neuron tested in A1 and the ICc, an auditory midbrain structure (Fig. 5E).

A subset of neurons ($n = 20$) was subjected to a subsequent test in which a single stimulus parameter was varied while the other four were held at their optimal value, as determined from the evolutionary search procedure (Fig. 5F). This procedure revealed the leverage of each individual sound dimension on the firing rate of the neuron. As the example in Figure 5F illustrates, A1 neurons may display bandpass, low- or high-pass selectivity across multiple feature dimensions, demonstrating the utility of unbiased search strategies. A glyph plot (Fig. 5F, bottom right) was constructed from this unit where each spoke represents the normalized difference in firing rate between the peak and the trough of the tuning function across each parameter. Glyph plots of all A1 single units subjected to this procedure are shown in Figure 5G. The heterogeneity of spoke lengths relative to other spokes in each glyph as well as other glyphs in the population further illustrate the complex interplay of combinatorial sensitivity and tolerance to multidimensional stimuli.

A1 neurons display a sparse code across some acoustic dimensions, with a dense code across others

As shown in Figure 6A, only variations across center frequency, and to a lesser extent location, could modulate the optimized stimulus response of A1 units from maximal to minimal firing rates. Other stimulus features modulated the firing rate by $\sim 50\%$ on average (significant differences across parameters indicated with horizontal lines; paired t test, with Bonferroni correction, significance level $p < 0.005$), indicating that spike rate remained relatively invariant across large variations in multiple parameters.

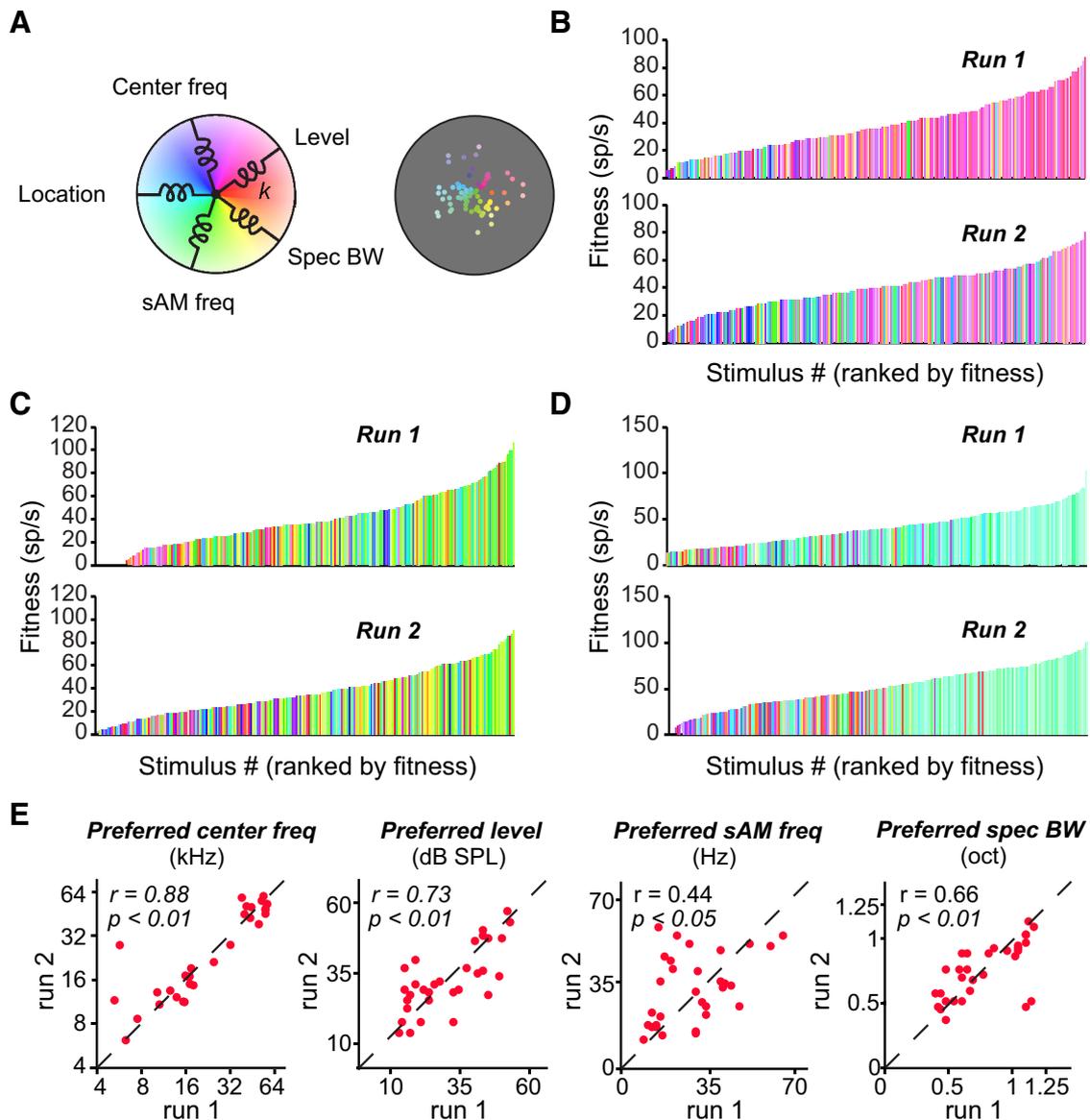


Figure 4. Adaptive stimulus search reliably converges on consistent areas of feature space across runs initialized with distinct pedigrees. **A**, Individual stimuli were plotted onto a circular hue/saturation map using the radial visualization technique (see Materials and Methods) according to the values of the five stimulus parameters, shown as equally spaced anchor points on the color map. The position of each stimulus determines its color. Right circle, The most “fit” stimulus (in terms of firing rate) from each of 78 runs of A1 single units is plotted onto the circular map. **B–D**, Example sets of stimuli from two separate runs of three individual A1 single units. Each stimulus is represented by a vertical bar, and bars are rank-ordered on the x-axis by firing rate response, which is also represented by the y-axis height (most effective stimuli appear to the right of each plot and are tallest). Color represents stimulus features; therefore, more similar colors are more closely spaced on the hue/saturation map with radial visualization. Plots show the following: (1) convergence to a set of effective stimulus features, and (2) stable convergence to the same features across two separate runs initialized with distinct randomly chosen stimuli. **E**, Correlation of best features determined from two separate runs of the adaptive search procedure from the same A1 single unit. The most effective value across all parameters (shown in separate plots) is determined by taking the mean value of the top 10 stimuli from each iteration of the adaptive search procedure.

This was somewhat surprising given that neural codes have been theorized to maximize efficiency such that fewer metabolically expensive action potentials would ideally be used to represent stimuli (a sparse code) (Barlow, 1961; Olshausen and Field, 1996). As an extension of these ideas, in a system where stimulus representation is progressively less linear as one ascends the processing hierarchy, at each stage the most efficient stimulus representation would be increasingly sparse. This hypothesis has been made in the visual (Willmore et al., 2011) and auditory (Wang, 2007; Atencio et al., 2012) systems.

Experimental evidence indicated, however, that in the visual system sparseness may instead be balanced between different stages of the sensory processing hierarchy. This preservation of

sparseness can be attributed to balanced increases of tolerance and selectivity across identity-preserving and identity-determining variations, respectively (Rust and DiCarlo, 2012). Consistent with the observation that intensity invariance, although not commonly seen in subcortical areas, is present in A1 single units (Sadagopan and Wang, 2008), we observed that the fractional change in peak firing rate ((peak-trough)/peak) across level variation was lower in A1 than for ICc neurons, indicating a higher tolerance of A1 maximal responses to variations of intensity (Fig. 6B). Further, we calculated the number of SDs between the peak of the tuning functions and the mean (z -score) as a proxy of selectivity and found that center frequency was associated with higher z -scores in A1 than ICc (paired t test, $p < 0.01$), consistent

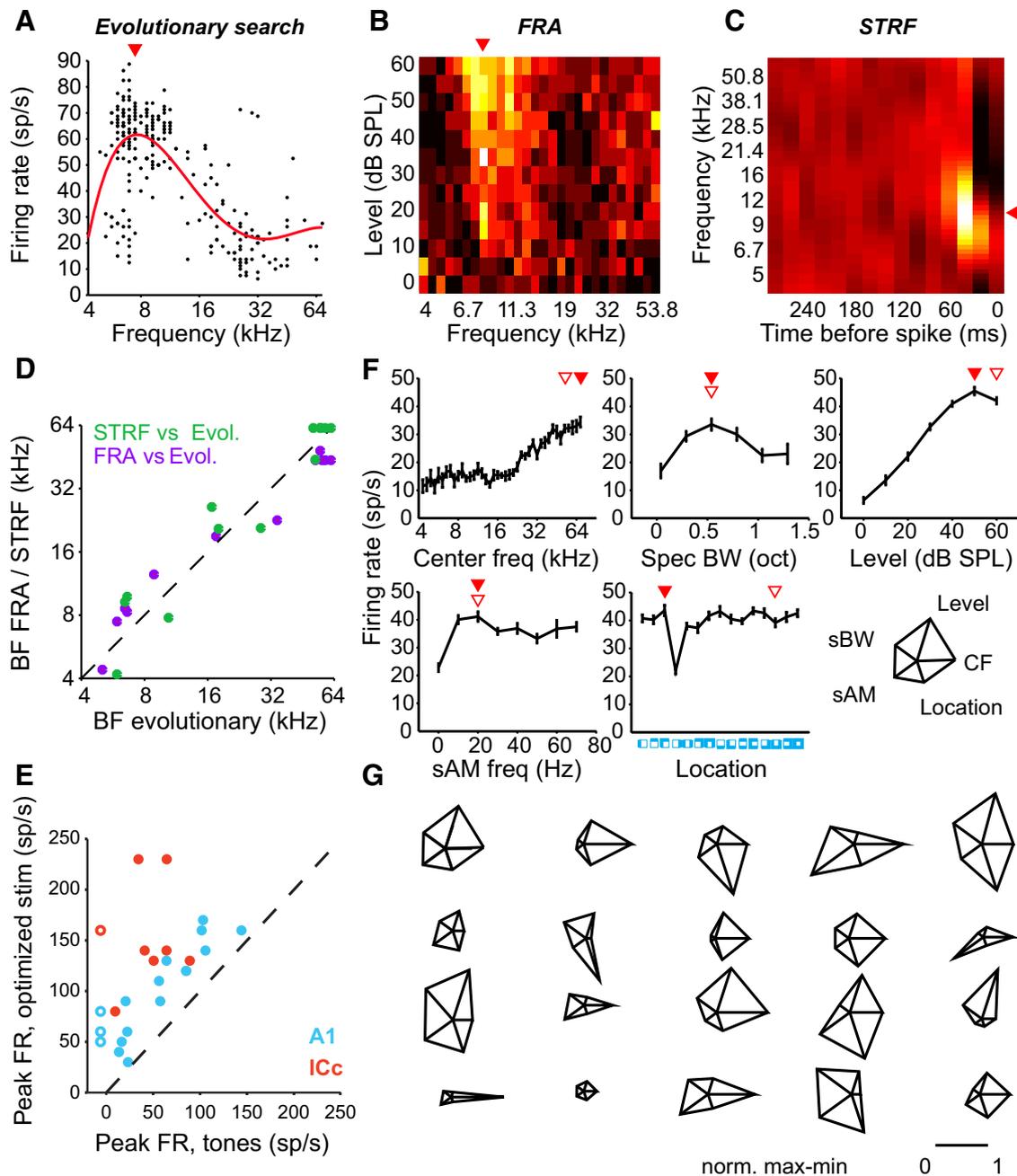


Figure 5. Adaptive stimulus search performs well against conventional methods but allows for unbiased search across more stimulus dimensions. **A**, Scatter plot of firing rate versus center frequency (kHz) of all stimuli from one run of the evolutionary procedure in an example neuron. Stimuli were varying across four other dimensions not plotted here. Red line indicates a fourth-order polynomial fit of the mean firing rate across frequency. Red arrow indicates the preferred frequency. **B**, A frequency–response area measurement from the same neuron as in **A**. Red arrow indicates the best frequency of the FRA. **C**, An STRF measurement from the same neuron as in **A** and **B**. Red arrow indicates best frequency, which closely corresponds to values obtained from the FRA measurement and evolutionary search. **D**, Summary of concordance between the best frequency obtained with the evolutionary procedure versus the FRA (purple) and the STRF (green) ($n = 12$). **E**, Peak firing rates from A1 (blue) and ICc (red) were calculated as the peak 50 ms response bin across all stimuli presented. Open circles represent units that did not display a significant response above spontaneous rate to any pure tones but were responsive to one or more stimuli from the evolutionary search procedure. **F**, Example one-dimensional tuning functions from a single A1 neuron where each dimension of the preferred stimulus (obtained from the evolutionary search procedure) was varied at a time while others were held constant. Filled and unfilled arrows indicate the actual peak and the optimized value obtained for that parameter in evolutionary search, respectively. Glyph plot represents the difference in firing rate between the maximum and minimum values of each one-dimensional tuning function, meant to estimate the leverage of each stimulus dimension on the maximal response of the neuron. **G**, Glyph plots constructed from all A1 neurons subjected to one-dimensional variations of the preferred stimulus. Spokes indicate same parameters as schematized in **F** ($n = 20$).

with sharper frequency tuning observed in primate A1 (Bartlett et al., 2011) (Fig. 6C). As in the visual system, the shifting balance of selectivity and tolerance for stimulus features in two hierarchically related brain areas was associated with no net change in lifetime sparseness (see Materials and Methods) between ICc ($n = 12$) and A1 ($n = 50$) units (Fig. 6D,E).

Although overall responses were not significantly more sparse in A1 than ICc, a prominent hypothesis for sound coding at higher levels of the central auditory pathways holds that the sustained nature of stimulus-evoked responses, rather than the overall rate, is indicative of stimulus preference (Wang et al., 2005; Wang, 2007). Namely, responses in A1 single units in awake an-

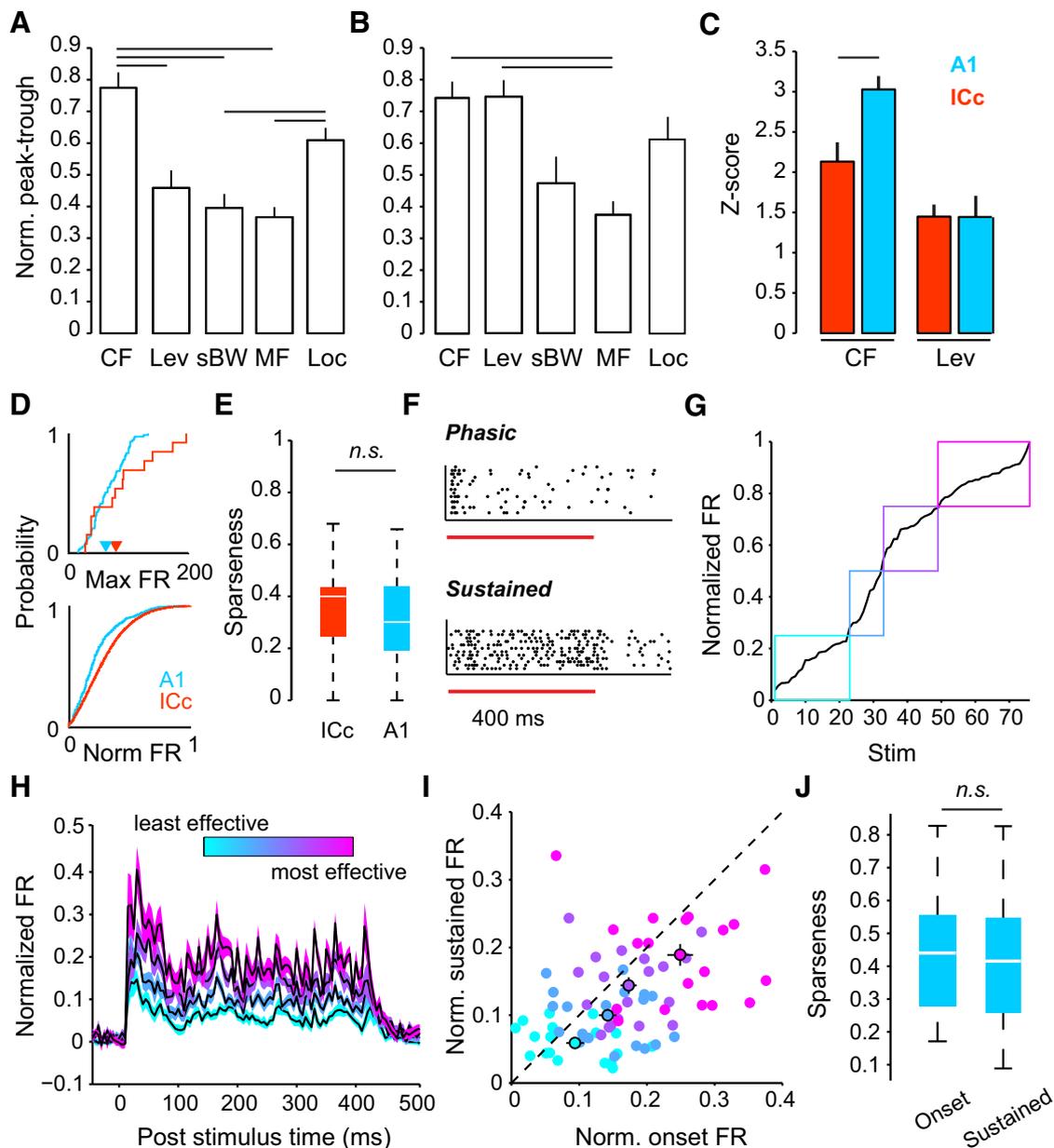


Figure 6. The evolutionary stimulus search procedure allows the observation that stimulus preference is not disproportionately indicated by a tonic response in awake A1, and A1 units are not sparser than IC units. **A**, Mean (\pm SEM) normalized peak minus trough firing rates across A1 units ($n = 20$) for one-dimensional tuning functions in which one parameter (on x -axis) was varied while other parameters remained at optimized values. Black lines indicate significant differences between indicated groups ($p < 0.005$, paired t test, with Bonferroni correction for multiple comparisons). **B**, Same as in **A**, but for IC units ($n = 8$). **C**, z -score of center frequency and intensity tuning functions of IC and A1 neurons. **D**, Cumulative probability distribution of maximum firing rates (top, triangles on x -axis represent the median values, A1 = 58 sp/s; IC = 78 sp/s) and normalized firing rates (bottom) in A1 and IC to randomly selected, five-dimensional stimuli (A1 $n = 50$, IC $n = 12$). **E**, Lifetime sparseness (median \pm IQR) of A1 and IC neurons to 100 randomly selected, 5-dimensional stimuli. n.s., Not significant. **F**, Example raster plots of phasic and tonic responses from cortical single units. **G**, Distribution of firing rate responses (black line) from one example A1 neuron (76 stimuli, obtained by starting at the preferred stimulus from evolutionary search and varying one parameter at a time along its relevant range). Colored boxes represent quartiles according to proportion of maximal response. **H**, For A1 neurons, mean normalized PSTHs to stimuli falling within each quartile of the response distribution (indicated by color corresponding with Fig. 4C, $n = 20$). PSTHs are normalized to the peak 20 ms bin for each neuron. **I**, Onset (0–50 ms) versus sustained (51–400 ms) firing rate (mean, normalized) to stimuli in each quartile of the response distribution in A1. Mean \pm SEM along each axis indicated in black outlined circles. **J**, Lifetime sparseness (median \pm IQR) of A1 units in response to 100 randomly chosen 5-dimensional stimuli using only the onset (left) or sustained (right) portions of the response.

imals are theorized to systematically shift from phasic (onset) to tonic (sustained) firing (Fig. 6F) as the stimulus is progressively fit to the most effective regions of the multidimensional receptive field. We therefore organized A1 neural responses into an onset portion (0–50 ms) and a sustained portion (51–400 ms) to gauge potentially independent effects of stimulus optimality on firing rate and sparseness during the two epochs. We hypothesized that, if the aforementioned model were correct, the steady-state por-

tion of the response would grow disproportionately as stimuli more closely approximated the preferred region of the multidimensional receptive field.

Contrary to the model's prediction, we found that onset and sustained firing rates tended to increase together across the population of recorded neurons (Fig. 6G–I), as shown in the population PSTHs in response to progressively optimized stimuli (Fig. 6G,H; colors indicate which quartile of the stimulus–response

function for each neuron contributed to the PSTH; one example of this function for an A1 neuron is shown). We found no evidence indicating disproportionate increases in the steady-state epoch of the response (paired t test, $p > 0.05$ for all quartiles). As expected, sparseness values were not significantly different when the firing rate was calculated only from the onset or sustained epoch (Fig. 6*J*). The same results were observed in the ICc (paired t test, $p > 0.05$ for all quartiles; Fig. 7*A, B*). For direct comparison with previous findings in marmoset A1, we plotted the onset response index, or the ratio of the onset to sustained response, for the preferred and nonpreferred stimuli of A1 and IC neurons (Fig. 7*C*). Our results indicate that the transition from the non-preferred to preferred receptive field was not accompanied by a shift from onset to sustained firing.

Discussion

Early stages of auditory processing feature a variety of biophysical and synaptic specializations that allow neurons to precisely synchronize action potential timing to spectral, temporal, and dichotic features within the acoustic source signal. As afferent activity is propagated through the central auditory neuroaxis, the functional organization changes to support more integrative and contextual processing. At the level of the auditory cortex, the high-fidelity representations from earlier stages have been almost entirely reformatted to rate-based abstractions of the original signal (Wang, 2007). This shift can be attributed, in part, to the emergence of higher-order neurons that behave more like non-linear feature detectors than generalized spectral and temporal analyzers (Young et al., 2005; Wang, 2007; Nelson and Young, 2010; Schneider and Woolley, 2011; Atencio et al., 2012; Schinkel-Bielefeld et al., 2012; Yaron et al., 2012). Much can be learned about sensory encoding from the exhaustive characterization of neural tuning to isolated acoustic attributes. However, the approach has its limitations in terms of constructing a comprehensive and unifying theory of receptive field structure in higher-order auditory areas. The approach described here, in keeping with an emerging body of literature on sensory stimulus optimization, facilitates the adoption of unbiased stimulus selection procedures by relying on easily implemented and theoretically intuitive design principles.

In addition to a genetic algorithm (Bleeck et al., 2003; Yamane et al., 2008; Hung et al., 2012), other methods have been described for firing rate maximization, such as local hill-climbing or gradient search (O'Connor et al., 2005; Földiák, 2011; Koelling and Nykamp, 2012), as well as simplex search (Nelken et al., 1994). A potential benefit of the genetic algorithm is that, rather than updating the position of a single point, or simplex of points, in stimulus space, the algorithm used here allows the exploration of diverse regions of feature space by optimizing a group of breeder stimuli as well as incorporating random “mutations” into each generation. These features make the search strategy less likely to persevere in local firing rate maxima. In addition, some stimulus parameters may not be well defined along a gradient, and may be better suited to an evolutionary search or other similar strategies, such as particle swarm optimization (Kennedy and Eberhart, 1995).

The evolutionary procedure allows the experimenter to obtain information about stimulus selectivity across at least five separate acoustic feature dimensions, whereas traditional methods typically vary only across a frequency/intensity or spectrotemporal stimulus space and force other potentially relevant parameters to an arbitrary constant. In addition to variations of amplitude-modulated pure tones and band-limited noise to maximize firing rate, the evolutionary approach is flexible enough to accommo-

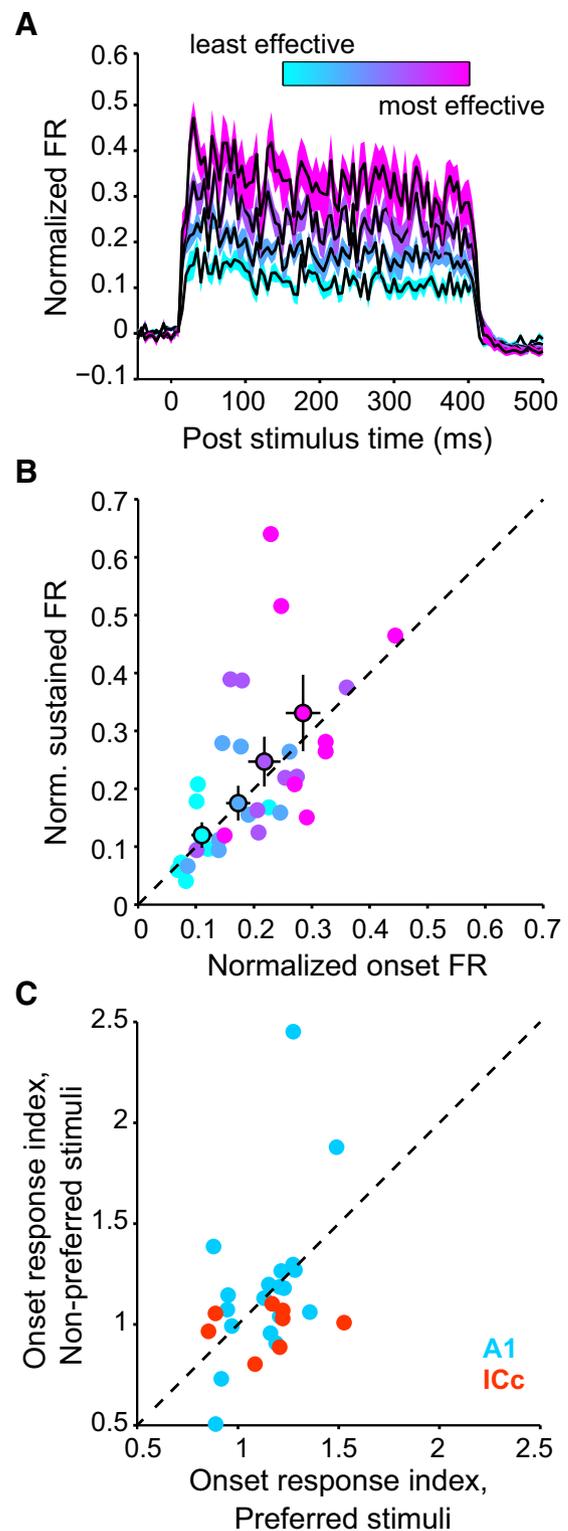


Figure 7. A shift from phasic to tonic firing is not seen in single units recorded from awake mouse ICc. **A**, For ICc neurons, mean normalized PSTHs to stimuli falling within each quartile of the response distribution (indicated by color corresponding with Fig. 6*G*). PSTHs are normalized to the peak 20 ms bin for that unit. **B**, Onset versus sustained firing rate (mean, normalized) to stimuli in each quartile of the response distribution in ICc. Mean \pm SEM along each axis indicated in black outlined circles. **C**, Onset Response Index (ratio of onset to sustained responses) plotted for A1 and ICc neurons, where the optimal stimulus was found via the evolutionary procedure and each stimulus parameter was varied individually along its entire sampling range. “Preferred” stimuli were those that fell within the top quartile of response magnitude; “nonpreferred” were those that fell within the bottom quartile but still exhibited responses above spontaneous rate.

date a discrete or continuous feature space in which complex stimuli, such as vocalizations, could be used to optimize not only firing rate but also any number of characteristics of the neural response. Further, the approach does not assume a linear stimulus–response function. These qualities highlight potential advantages of a genetic optimization algorithm over traditional techniques, such as the STRF (Aertsen and Johannesma, 1981; Klein et al., 2000), typically defined as the best linear fit (in a least-squares sense) between a spectrogram of the stimulus and the evoked spike rate.

Response maximization was explored in this study to describe the relationship between stimulus preference and firing rate in A1 single units. However, with the evolutionary optimization approach, the stimulus manifold is sampled sparsely and unevenly compared with STRF estimates derived from continuous, dynamic, and uncorrelated stimuli, such as ripple noise (Escabi and Schreiner, 2002). In addition, the online characterization used here did not generate a testable model of the receptive field organization, as has been described in a recent conference abstract (Feng et al., 2012) or other methods that maximize the mutual information between the stimulus and the response (Machens, 2002; Machens et al., 2005; Nelken et al., 2005) to characterize a maximally informative stimulus distribution (Sharpee et al., 2004). Along similar lines, other studies have used principles of information theory to adaptively optimize the expected information gain from the parameters of a receptive field model and the stimulus (Paninski, 2005; Paninski et al., 2007).

Although the model neurons on which the closed-loop procedure was initially tested exhibited a clearly defined maximum response, not all sensory neurons display this property. For example, neurons could be modeled with increasing responsiveness along a stimulus dimension such that the optimal stimulus can only be described by subjecting the response to a power constraint (Berkes and Wiskott, 2006). Indeed, some A1 neurons in our study exhibited monotonically increasing tuning functions across at least one stimulus parameter (Fig. 5*F*), and their true preferred stimulus features may not be completely described by the discrete stimulus space chosen for the study. Yet others may display plateaus or saddle shapes in their tuning functions that would result in multiple peaks (DiMattina and Zhang, 2013). Therefore, in practice, it is not necessarily useful to speak of only one “globally optimal” stimulus. Rather, a major benefit of online optimization is that it allows the experimenter to define a large and diverse stimulus parameter space without making assumptions as to which parameters may exert the most leverage over the neural response.

The procedure described here was able to rapidly (10 min for our paradigm, although the exact time will vary depending on the stimulus space and brain region studied) optimize 5-dimensional stimuli for cortical and midbrain neurons, which is well within the time frame of a typical single unit recording in an unanesthetized animal. However, it is likely that some experimenters may increase the dimensionality of their stimulus space or perform a more exhaustive search of the space, necessitating longer recording times. Therefore, the procedure could be a starting point for a more exhaustive, open-loop receptive field characterization or alternatively could be expanded to more fully delineate the topology of a higher dimensional feature space, rather than just its peak.

Through the use of adaptive, multidimensional stimulus search, we have shown that different acoustic properties have distinct amounts of leverage over the firing rate response to an optimized stimulus. Further, the leverage of some properties, such as intensity, can be related to the stage of the processing hierarchy from which neural responses were obtained (Fig. 6).

Side-by-side comparisons of tuning to multiple features allow the experimenter to see relationships, such as the interplay between tolerance and selectivity, which may underlie broader neural coding principles in a sensory system, such as the preservation of sparseness along a processing hierarchy. In the auditory system, the fact that both phasic and tonic responses have been observed to varying degrees in A1 has kept the field in disagreement regarding the relationship between neural preference and firing mode (DeWeese et al., 2003; Wang et al., 2005; Hromádka et al., 2008; Qin et al., 2009). However, the efficient and reliable approach described here provides reasonable evidence indicating that neural preference in mouse A1 is not necessarily marked by a shift from phasic to tonic firing (Fig. 6). The generality of this observation could be further reinforced by alternative single neuron measurement strategies, such as loose-patch recordings or two-photon fluorescence imaging, which may provide greater access to neuron types that are more easily overlooked with extracellular recording methods.

Sensory neuron firing rates are often used as a read-out for experimental manipulations, such as behavioral training, overexposure, or deprivation. With conventional methods, changes in feature selectivity to the manipulated parameters are studied in relative isolation; however, multidimensional search allows for feature selectivity to be observed within the context of a complex, and potentially more naturalistic, stimulus. Further, the ability to study neural responses outside of the typical constraints of stimulus selection creates opportunities to test principles of neural encoding, which may be consistent across sensory systems dealing with varying degrees of stimulus complexity. Heterogeneity in cortical responses has, for the most part, functioned as a hindrance to the understanding of sound representations. By working toward automated methods for auditory experiments, the complex and unpredictable nature of neural responses in cortex can instead be a valuable asset for delineating the neural mechanisms that support auditory perception.

References

- Aertsen AM, Johannesma PI (1981) The spectro-temporal receptive field: a functional characteristic of auditory neurons. *Biol Cybern* 42:133–143. [CrossRef Medline](#)
- Ankerst M, Keim DA, Kriegel HP (1996) Circle segments: a technique for visually exploring large multidimensional datasets. Presented at Proceedings of IEEE Visualization, Hot Topic Session, San Francisco, CA.
- Atencio CA, Sharpee TO, Schreiner CE (2012) Receptive field dimensionality increases from the auditory midbrain to cortex. *J Neurophysiol* 107:2594–2603. [CrossRef Medline](#)
- Barlow HB, ed (1961) *The coding of sensory messages*. Cambridge: Cambridge UP.
- Bartlett EL, Sadagopan S, Wang X (2011) Fine frequency tuning in monkey auditory cortex and thalamus. *J Neurophysiol* 106:849–859. [CrossRef Medline](#)
- Bellman R (1961) *Adaptive control processes: a guided tour*. Princeton, NJ: Princeton UP.
- Benda J, Gollisch T, Machens CK, Herz AV (2007) From response to stimulus: adaptive sampling in sensory physiology. *Curr Opin Neurobiol* 17:430–436. [CrossRef Medline](#)
- Berkes P, Wiskott L (2006) On the analysis and interpretation of inhomogeneous quadratic forms as receptive fields. *Neural Comput* 18:1868–1895. [CrossRef Medline](#)
- Bizley JK, Walker KM, Silverman BW, King AJ, Schnupp JW (2009) Interdependent encoding of pitch, timbre, and spatial location in auditory cortex. *J Neurosci* 29:2064–2075. [CrossRef Medline](#)
- Bleack S, Patterson RD, Winter IM (2003) Using genetic algorithms to find the most effective stimulus for sensory neurons. *J Neurosci Methods* 125:73–82. [CrossRef Medline](#)
- Bölinger D, Gollisch T (2012) Closed-loop measurements of iso-response

- stimuli reveal dynamic nonlinear stimulus integration in the retina. *Neuron* 73:333–346. [CrossRef Medline](#)
- Carruthers IM, Natan RG, Geffen MN (2013) Encoding of ultrasonic vocalizations in the auditory cortex. *J Neurophysiol* 109:1912–1927. [CrossRef Medline](#)
- David SV, Shamma SA (2013) Integration over multiple timescales in primary auditory cortex. *J Neurosci* 33:19154–19166. [CrossRef Medline](#)
- deCharms RC, Blake DT, Merzenich MM (1998) Optimizing sound features for cortical neurons. *Science* 280:1439–1443. [CrossRef Medline](#)
- DeWeese MR, Wehr M, Zador AM (2003) Binary spiking in auditory cortex. *J Neurosci* 23:7940–7949. [Medline](#)
- DiMattina C, Zhang K (2011) Active data collection for efficient estimation and comparison of nonlinear neural models. *Neural Comput* 23:2242–2288. [CrossRef Medline](#)
- DiMattina C, Zhang K (2013) Adaptive stimulus optimization for sensory systems neuroscience. *Front Neural Circuits* 7:101. [CrossRef Medline](#)
- Edin F, Machens CK, Schütze H, Herz AV (2004) Searching for optimal sensory signals: iterative stimulus reconstruction in closed-loop experiments. *J Comput Neurosci* 17:47–56. [CrossRef Medline](#)
- Escabi MA, Schreiner CE (2002) Nonlinear spectrotemporal sound analysis by neurons in the auditory midbrain. *J Neurosci* 22:4114–4131. [Medline](#)
- Feng L, Dekel E, Zhang K, Wang X (2012) Online adaptive stimulus design for studying non-linear spectral integration of neurons in auditory cortex. New Orleans: Society for Neuroscience Conference Abstract.
- Földiák P (2011) Neural control: closed-loop human brain reading. *Curr Biol* 21:R80–R81. [CrossRef Medline](#)
- Guo W, Chambers AR, Darrow KN, Hancock KE, Shinn-Cunningham BG, Polley DB (2012) Robustness of cortical topography across fields, laminae, anesthetic states, and neurophysiological signal types. *J Neurosci* 32:9159–9172. [CrossRef Medline](#)
- Hackett TA, Barkat TR, O'Brien BM, Hensch TK, Polley DB (2011) Linking topography to tonotopy in the mouse auditory thalamocortical circuit. *J Neurosci* 31:2983–2995. [CrossRef Medline](#)
- Hromádka T, Deweese MR, Zador AM (2008) Sparse representation of sounds in the unanesthetized auditory cortex. *PLoS Biol* 6:e16. [CrossRef Medline](#)
- Hung CC, Carlson ET, Connor CE (2012) Medial axis shape coding in macaque inferotemporal cortex. *Neuron* 74:1099–1113. [CrossRef Medline](#)
- Kennedy J, Eberhart R (1995) Particle swarm optimization. In: *Proceedings of the IEEE International Conference on Neural Networks*, Perth, Australia.
- Klein DJ, Depireux DA, Simon JZ, Shamma SA (2000) Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design. *J Comput Neurosci* 9:85–111. [CrossRef Medline](#)
- Koelling ME, Nykamp DQ (2012) Searching for optimal stimuli: ascending a neuron's response function. *J Comput Neurosci* 33:449–473. [CrossRef Medline](#)
- Lewi J, Butera R, Paninski L (2009) Sequential optimal design of neurophysiology experiments. *Neural Comput* 21:619–687. [CrossRef Medline](#)
- Lewi J, Schneider DM, Woolley SM, Paninski L (2011) Automating the design of informative sequences of sensory stimuli. *J Comput Neurosci* 30:181–200. [CrossRef Medline](#)
- Linden JF, Liu RC, Sahani M, Schreiner CE, Merzenich MM (2003) Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex. *J Neurophysiol* 90:2660–2675. [CrossRef Medline](#)
- Machens CK (2002) Adaptive sampling by information maximization. *Phys Rev Lett* 88:228104. [CrossRef Medline](#)
- Machens CK, Gollisch T, Kolesnikova O, Herz AV (2005) Testing the efficiency of sensory coding with optimal stimulus ensembles. *Neuron* 47:447–456. [CrossRef Medline](#)
- Merzenich MM, Knight PL, Roth GL (1975) Representation of cochlea within primary auditory-cortex in cat. *J Neurophysiol* 38:231–249. [Medline](#)
- Middlebrooks JC, Clock AE, Xu L, Green DM (1994) A panoramic code for sound location by cortical neurons. *Science* 264:842–844. [CrossRef Medline](#)
- Nelken I, Prut Y, Vaadia E, Abeles M (1994) In search of the best stimulus: an optimization procedure for finding efficient stimuli in the cat auditory-cortex. *Hearing Res* 72:237–253. [CrossRef Medline](#)
- Nelken I, Chechik G, Msrac-Flogel TD, King AJ, Schnupp JW (2005) Encoding stimulus information by spike numbers and mean response time in primary auditory cortex. *J Comput Neurosci* 19:199–221. [CrossRef Medline](#)
- Nelson PC, Young ED (2010) Neural correlates of context-dependent perceptual enhancement in the inferior colliculus. *J Neurosci* 30:6577–6587. [CrossRef Medline](#)
- O'Connor KN, Petkov CI, Sutter ML (2005) Adaptive stimulus optimization for auditory cortical neurons. *J Neurophysiol* 94:4051–4067. [CrossRef Medline](#)
- Olshausen BA, Field DJ (1996) Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* 381:607–609. [CrossRef Medline](#)
- Paninski L (2005) Asymptotic theory of information-theoretic experimental design. *Neural Comput* 17:1480–1507. [CrossRef Medline](#)
- Paninski L, Pillow J, Lewi J (2007) Statistical models for neural encoding, decoding, and optimal stimulus design. *Prog Brain Res* 165:493–507. [CrossRef Medline](#)
- Phillips DP, Irvine DR (1981) Responses of single neurons in physiologically defined primary auditory-cortex (Ai) of the cat: frequency tuning and responses to intensity. *J Neurophysiol* 45:48–58. [Medline](#)
- Qin L, Liu Y, Wang J, Li S, Sato Y (2009) Neural and behavioral discrimination of sound duration by cats. *J Neurosci* 29:15650–15659. [CrossRef Medline](#)
- Read HL, Winer JA, Schreiner CE (2001) Modular organization of intrinsic connections associated with spectral tuning in cat auditory cortex. *Proc Natl Acad Sci U S A* 98:8042–8047. [CrossRef Medline](#)
- Rust NC, DiCarlo JJ (2012) Balanced increases in selectivity and tolerance produce constant sparseness along the ventral visual stream. *J Neurosci* 32:10170–10182. [CrossRef Medline](#)
- Sadagopan S, Wang X (2008) Level invariant representation of sounds by populations of neurons in primary auditory cortex. *J Neurosci* 28:3415–3426. [CrossRef Medline](#)
- Sahani M, Linden JF (2003) Evidence optimization techniques for estimating stimulus-response functions. In: *Advances in neural information processing systems* (Becker S, Thrun S, Obermayer K, eds). Cambridge, MA: Massachusetts Institute of Technology.
- Sahani M, Williamson RS, Ahrens MB, Linden JF (2013) Probabilistic methods for linear and multilinear models. In: *Handbook of modern techniques in auditory cortex* (Depireux D, Elhilali M, eds). Hauppauge, NY: Nova.
- Schinkel-Bielefeld N, David SV, Shamma SA, Butts DA (2012) Inferring the role of inhibition in auditory processing of complex natural stimuli. *J Neurophysiol* 107:3296–3307. [CrossRef Medline](#)
- Schneider DM, Woolley SM (2011) Extra-classical tuning predicts stimulus-dependent receptive fields in auditory neurons. *J Neurosci* 31:11867–11878. [CrossRef Medline](#)
- Schreiner CE, Mendelson JR (1990) Functional topography of cat primary auditory cortex: distribution of integrated excitation. *J Neurophysiol* 64:1442–1459. [Medline](#)
- Sharpee T, Rust NC, Bialek W (2004) Analyzing neural responses to natural signals: maximally informative dimensions. *Neural Comput* 16:223–250. [CrossRef Medline](#)
- Wang X (2007) Neural coding strategies in auditory cortex. *Hear Res* 229:81–93. [CrossRef Medline](#)
- Wang X, Merzenich MM, Beitel R, Schreiner CE (1995) Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J Neurophysiol* 74:2685–2706. [Medline](#)
- Wang X, Lu T, Snider RK, Liang L (2005) Sustained firing in auditory cortex evoked by preferred stimuli. *Nature* 435:341–346. [CrossRef Medline](#)
- Willmore BD, Mazer JA, Gallant JL (2011) Sparse coding in striate and extrastriate visual cortex. *J Neurophysiol* 105:2907–2919. [CrossRef Medline](#)
- Wollberg Z, Newman JD (1972) Auditory cortex of squirrel monkey: response patterns of single cells to species-specific vocalizations. *Science* 175:212–214. [CrossRef Medline](#)
- Wu MC, David SV, Gallant JL (2006) Complete functional characterization of sensory neurons by system identification. *Annu Rev Neurosci* 29:477–505. [CrossRef Medline](#)
- Yamane Y, Carlson ET, Bowman KC, Wang Z, Connor CE (2008) A neural code for three-dimensional object shape in macaque inferotemporal cortex. *Nat Neurosci* 11:1352–1360. [CrossRef Medline](#)
- Yaron A, Hershenhoren I, Nelken I (2012) Sensitivity to complex statistical regularities in rat auditory cortex. *Neuron* 76:603–615. [CrossRef Medline](#)
- Yin P, Johnson JS, O'Connor KN, Sutter ML (2011) Coding of amplitude modulation in primary auditory cortex. *J Neurophysiol* 105:582–600. [CrossRef Medline](#)
- Young ED, Yu JJ, Reiss LA (2005) Non-linearities and the representation of auditory spectra. *Int Rev Neurobiol* 70:135–168. [CrossRef Medline](#)