

# A general framework explains receptive fields and plasticity in multiple sensory cortices

Andrew M. Saxe (asaxe@stanford.edu)  
Department of Electrical Engineering  
Stanford University, Stanford, CA 94305 USA

## Abstract

What are the computational principles that underly receptive fields in sensory cortices? We present a comprehensive, quantitative analysis of the fit between current theoretical models and experimental data across primary visual, auditory, and somatosensory cortex.

## Introduction

We use the word 'general' in three senses: first, we show that no specific algorithm has a decisive advantage in matching receptive field properties; rather, any instance of a broad class of sparsity inducing algorithms will perform comparably. Thus by 'general' we mean in part that there is a meta-algorithm, or set of principles, for generating plausible candidate algorithms: the method must attempt to preserve information about the input data, and encourage a sparse representation. Second, we show that the parameters governing the algorithmic implementations need not be chosen differently to accurately fit each modality; one set of parameters generalizes across modalities, and thus by 'general' we mean that these algorithms are not modality specific. Finally, we show that these algorithms also reproduce results from plasticity experiments. Hence these algorithms generalize to non-normal and even somewhat extreme manipulations of experience during rearing. Taken together, our results argue that one set of principles is shared across primary sensory cortices; that the parameters of the particular instantiation of these principles is shared; and that these principles are always active during rearing—be it normal rearing or altered rearing.

## Thoughts and speculation

Other possibilities: show that the recent PLOS alg for matching auditory data transfers to better match video temporal data in visual ctx, show that combining their temporal modification with the max superposition rule actually yields the overall most probable algorithm, even within each cortex, and joint parameter setting is the top performing one. that would be the strongest form. Might be a bit much and a bit longwinded for one paper... on the other hand it would have the virtue of proposing a new algorithm, with extremely careful/extensive quantitative and qualitative comparisons to experiment.

Have a downloadable version of the algorithms to allow experimentalists to make predictions. Have a downloadable version of the parameters to let modelers play around with it.

## Methods and Materials

Our results are divided into two main experiments. The first experiment consists of an extensive set of simulations that

quantitatively compare receptive fields measured in vivo to those generated by a variety of computational algorithms. We derive quantitative comparisons to Macaque primary visual cortex, Squirrel monkey primary auditory cortex, and Macaque primary somatosensory cortex. Each algorithm has a set of parameters governing its implementation. Using approximate Bayesian computation (ABC) (?), we estimate the posterior distribution of these parameters for each sensory cortex separately, and for all three cortices jointly. This enables a quantitative Bayesian comparison of the hypothesis that these parameters are modality dependent with the hypothesis that they generalize across modalities. The second experiment consists of a qualitative evaluation of the ability of these computational algorithms to replicate the receptive field plasticity observed in V1, A1, and S1 in response to altered rearing conditions.

## Learning algorithms

We use a

### ICA

### Sparse coding

### Sparse autoencoder

### Sparse restricted Boltzmann machine

### K-means clustering

### Binary sparse coding

### Max sparse coding

## Training data to simulate normal rearing conditions

### Vision

### Audition

### Somatosensation

## Comparisons to experimental data

### Vision

### Audition

### Somatosensation

## Estimating parameter posteriors using ABC

If the distance metric  $\rho(X, D)$  relies on summary statistics of  $X$  and  $D$ , ABC will still recover the true posterior in the limit  $\epsilon \rightarrow 0$ . Because computationally we must use  $\epsilon > 0$ , the estimated posterior will be approximate only. Additionally, in

our case we are relying on statistics of  $D$  reported in the literature, and there is no way to determine if they are sufficient—indeed that seems unlikely. Hence our posterior estimates will be affected by this additional level of approximation, in a way that is hard to predict. This is an unavoidable limitation at present. However, the statistics reported in the experimental literature were chosen with the aim to be informative, and hence it is assumed that they present a rich summary of the data.

## **Training data to simulate plasticity experiments**

**Vision**

**Audition**

**Somatosensation**

**Results**

**Discussion**