Receptive field models fail to predict responses of V1 neurons to natural movies

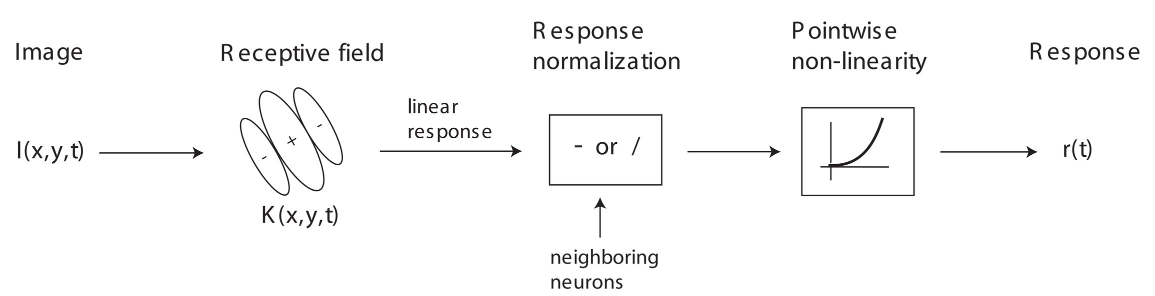
Bruno A. Olshausen*, Jonathan Baker, Shih-Cheng Yen, and Charles M. Gray
Department of Cell Biology and Neuroscience, Montana State University, Bozeman
*Redwood Neuroscience Institute and Center for Neuroscience, UC Davis

Introduction

Much effort has gone into characterizing the receptive field properties of visual cortical neurons. However, the extent to which receptive field models are capable of accounting for the responses of visual neurons to natural scenes is largely unexplored. In this study, we compare the activity of neurons in area V1 of the anaesthetized cat, recorded in response to a natural movie, to the predictions of a simple-cell receptive field model. We find that in most cases the receptive field model fails to adequately capture neural responses.

Model

A commonly accepted model of V1 simple cells is based on taking a linear weighted sum of image pixels over space and time, normalizing this value by the responses of other neurons, and passing the result through a point-wise nonlinearity.



Here we consider a somewhat simpler version without the response normalization step. The neuron's response *r* is modeled by computing the convolution of the image *I*(*x*,*y*,*t*) with a space-time kernel *K*(*x*,*y*,*t*) and passing the result through a point-wise non-linearity, plus an offset:

$r(t) = \alpha\, f(u(t) + \theta) + r_0$

$u(t) = \sum K(x,y,t) * I(x,y,t)$

$f(x) = x$^p if $x>0$, 0 otherwise

The receptive field kernel, *K*, was obtained via reverse correlation with an M-sequence. The kernel was then convolved with the movie to generate the intermediate response *u*. The parameters α, θ, r_0, and *p* where then adapted to minimize the squared error between the predicted response r(t) and the psth of the neuron recorded in response to the same movie.

Results

Results are shown in the boxes at right. For all plots the following legend applies:

Predicted response (red line)

Actual response (psth) (blue line)

Movies

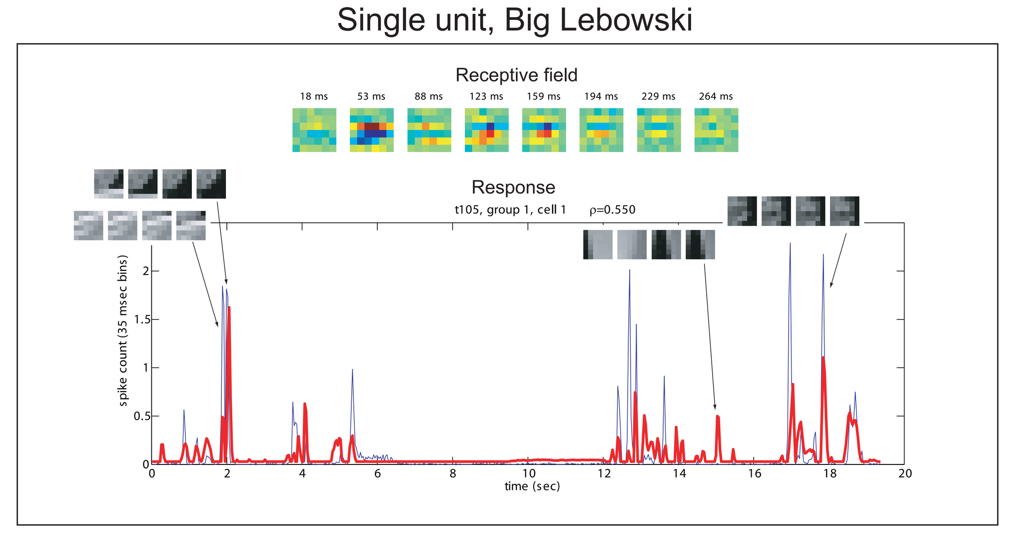
Animals

Big Lebowski

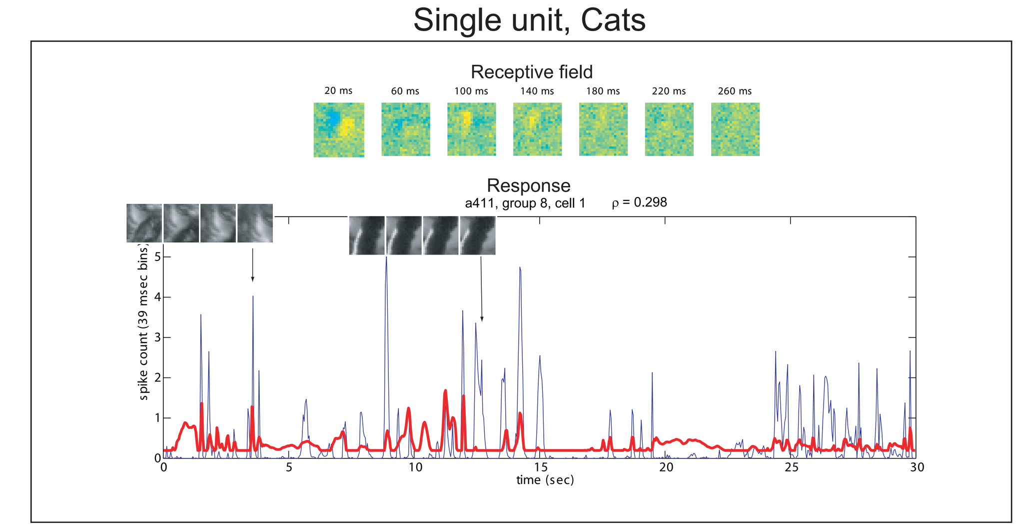
Cats

Everest

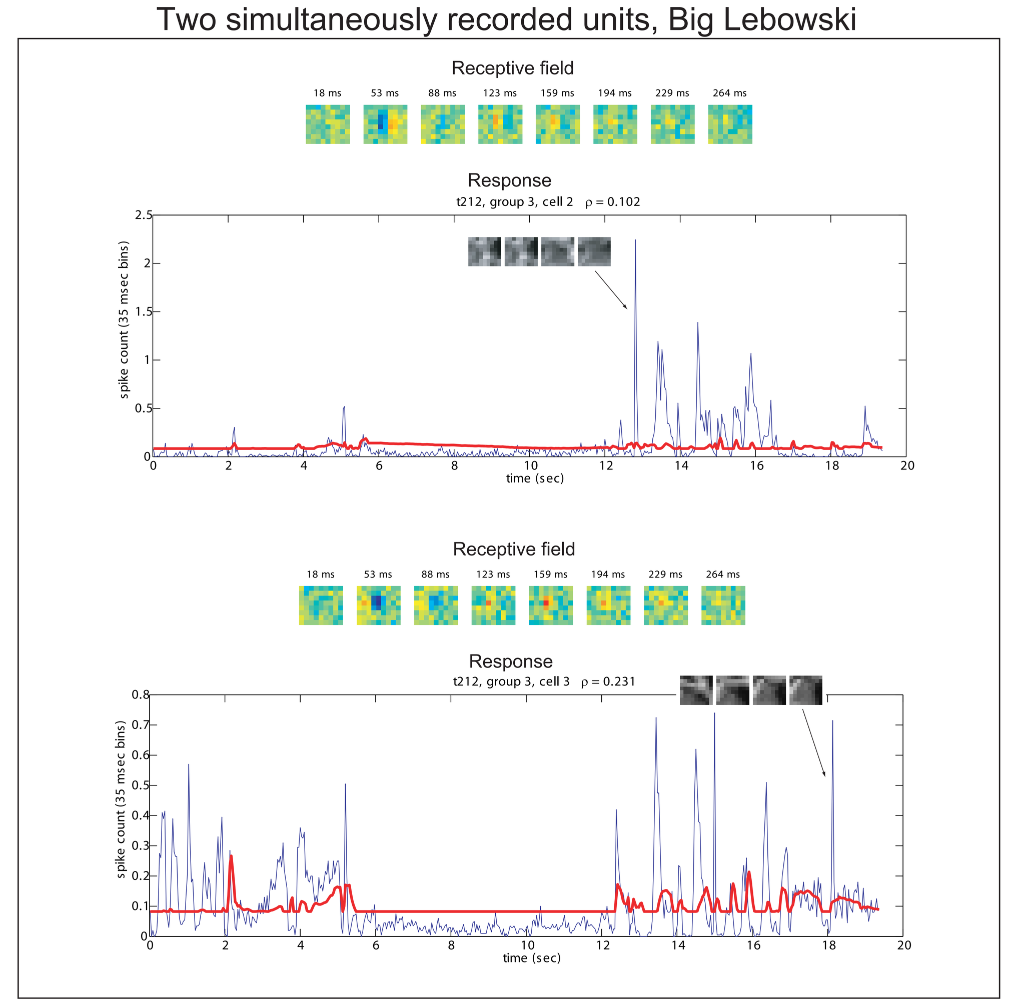
Single unit, Big Lebowski



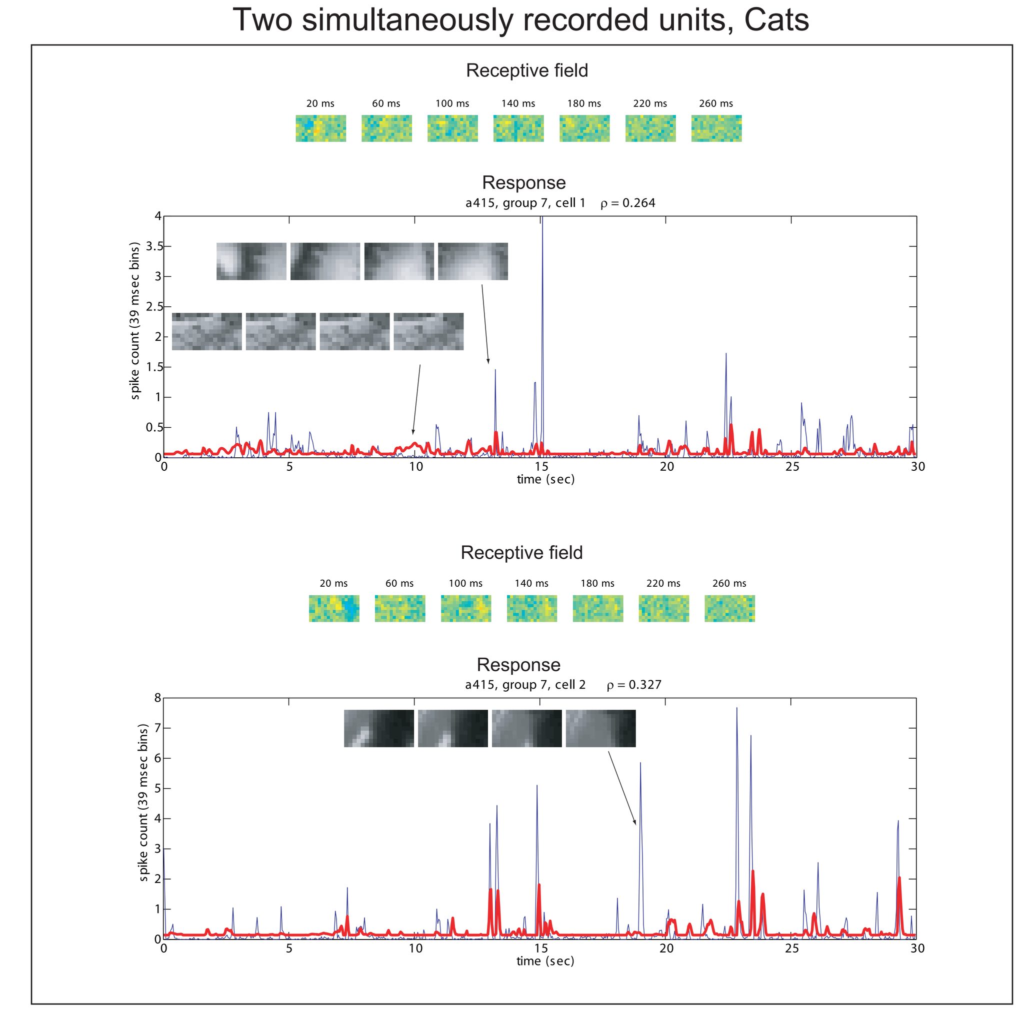
Single unit, Cats



Two simultaneously recorded units, Big Lebowski

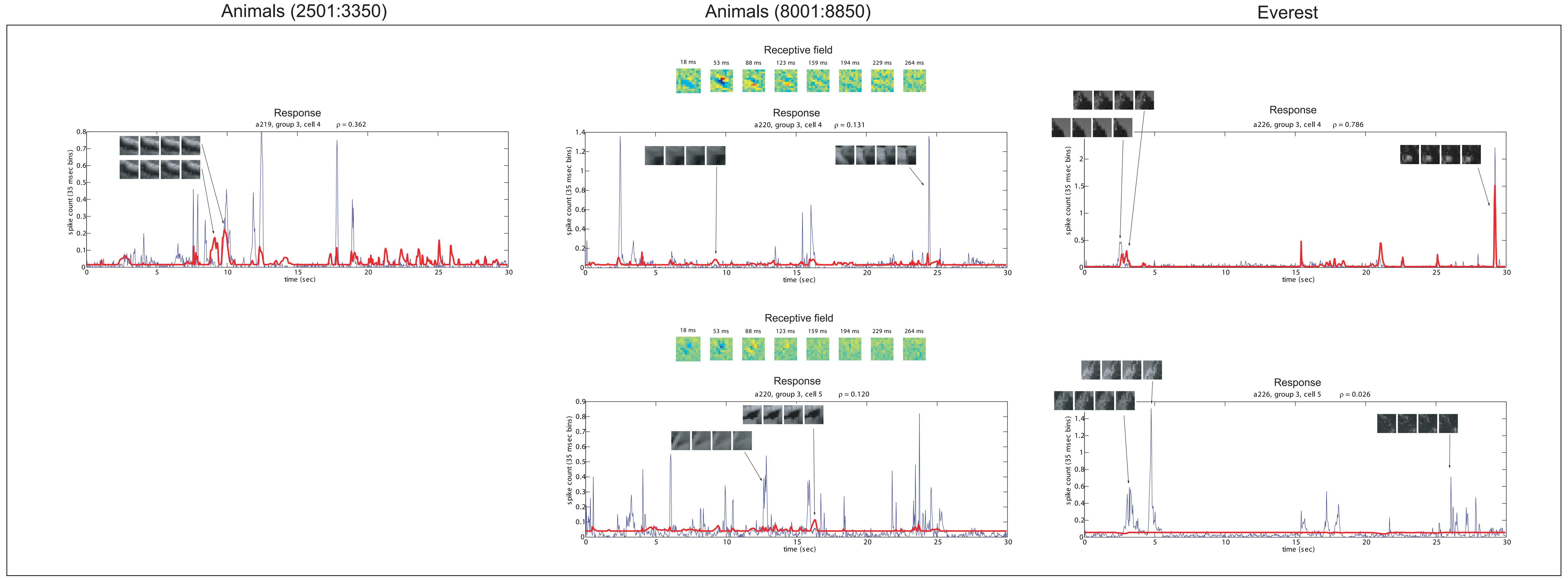


Two simultaneously recorded units, Cats

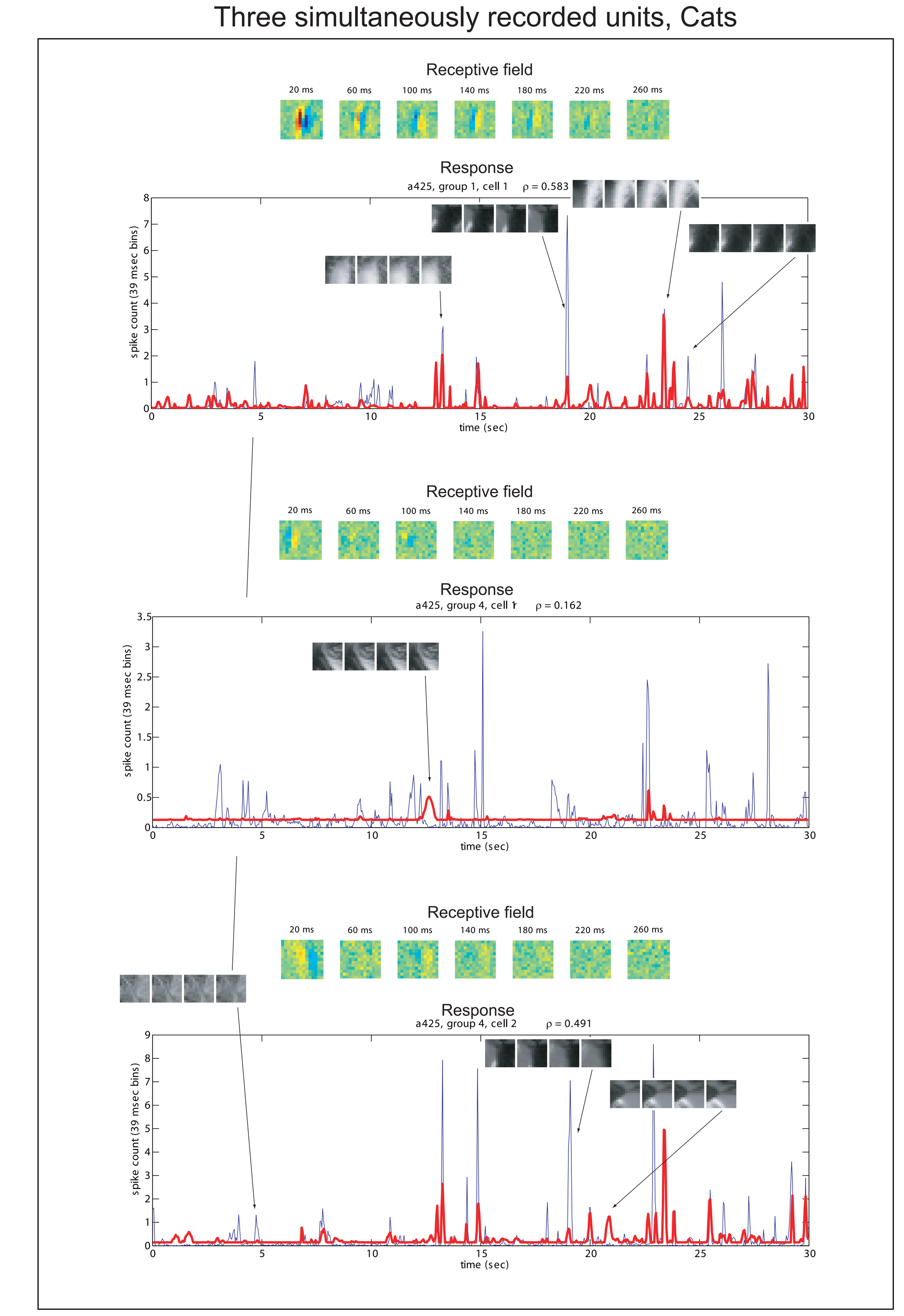


Two simultaneously recorded units

Animals (2501:3350) | Animals (8001:8850) | Everest

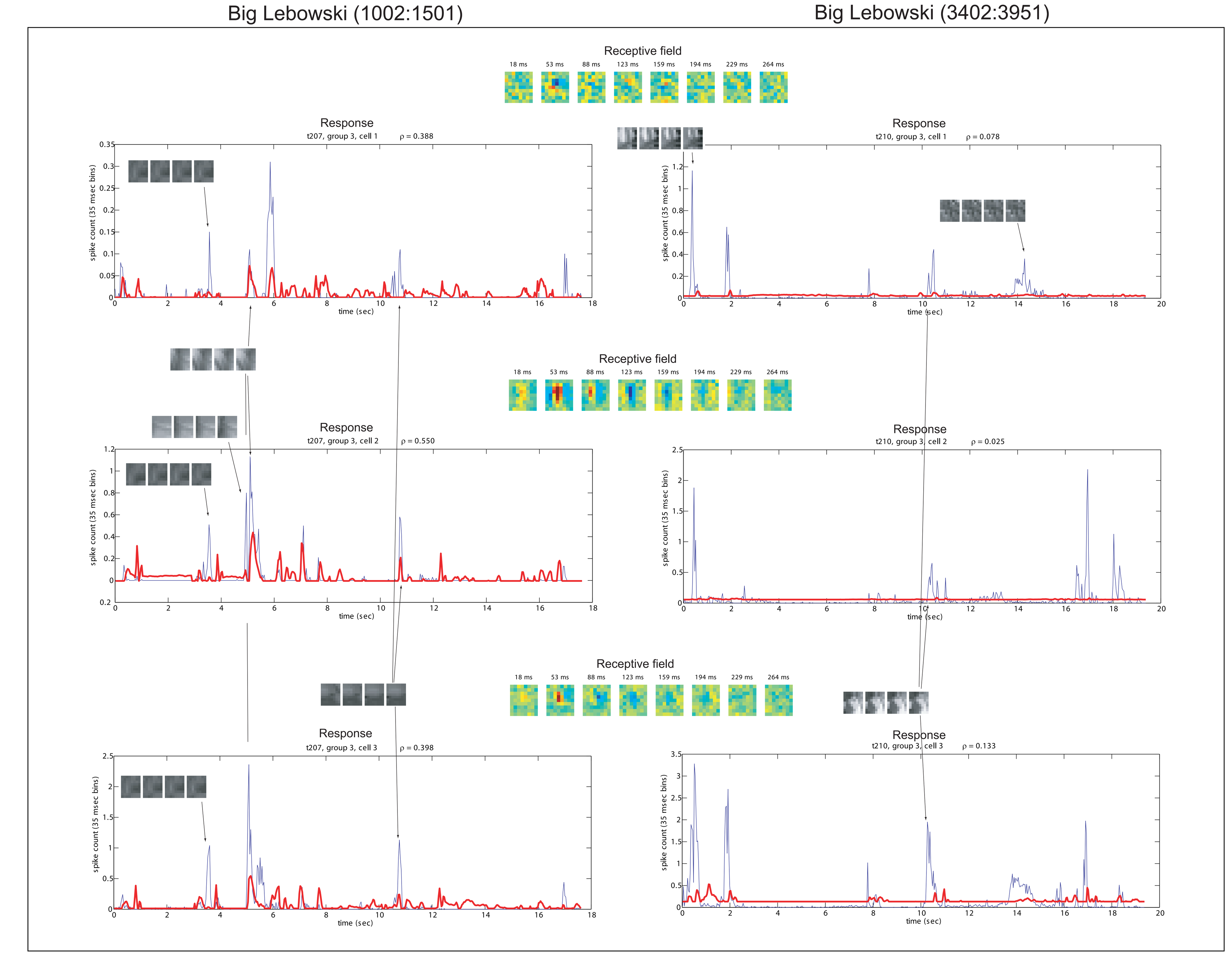


Three simultaneously recorded units, Cats



Three simultaneously recorded units

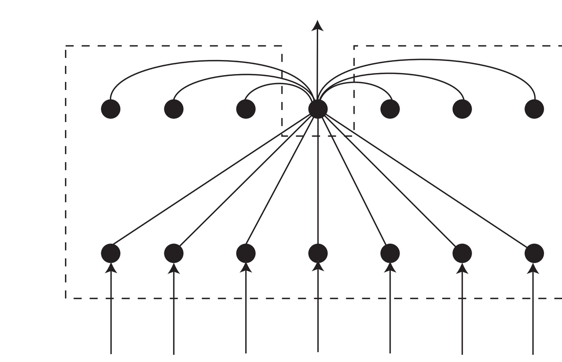
Big Lebowski (1002:1501) | Big Lebowski (3402:3951)



Discussion and Conclusions

For nearly all cells studied (for which we were able to obtain receptive field maps), the receptive field model failed to predict the actual response of the neuron. **Correlation coefficients hover well below 0.5**, meaning that the receptive field map *per se* is of limited value in helping us to understand how a neuron behaves under realistic conditions.

Obviously we will have to take into account other factors in order to explain how V1 neurons respond under natural conditions. In addition to their feedforward input from the LGN, neurons in layer 4 receive massive input from other cortical neurons:



Lumping together the effects of feedforward input from the LGN and recurrent input from cortex and calling it the "receptive field" is not a good idea. In order to properly understand what cortical neurons are doing, **it may well be necessary to model their response as a function of not only the input, but of other simultaneously recorded cortical neurons as well**.