Modeling pre-attentive stereo grouping by intracortical interactions in early visual cortex.

Zhaoping Li, University College London

www.gatsby.ucl.ac.uk/~zhaoping

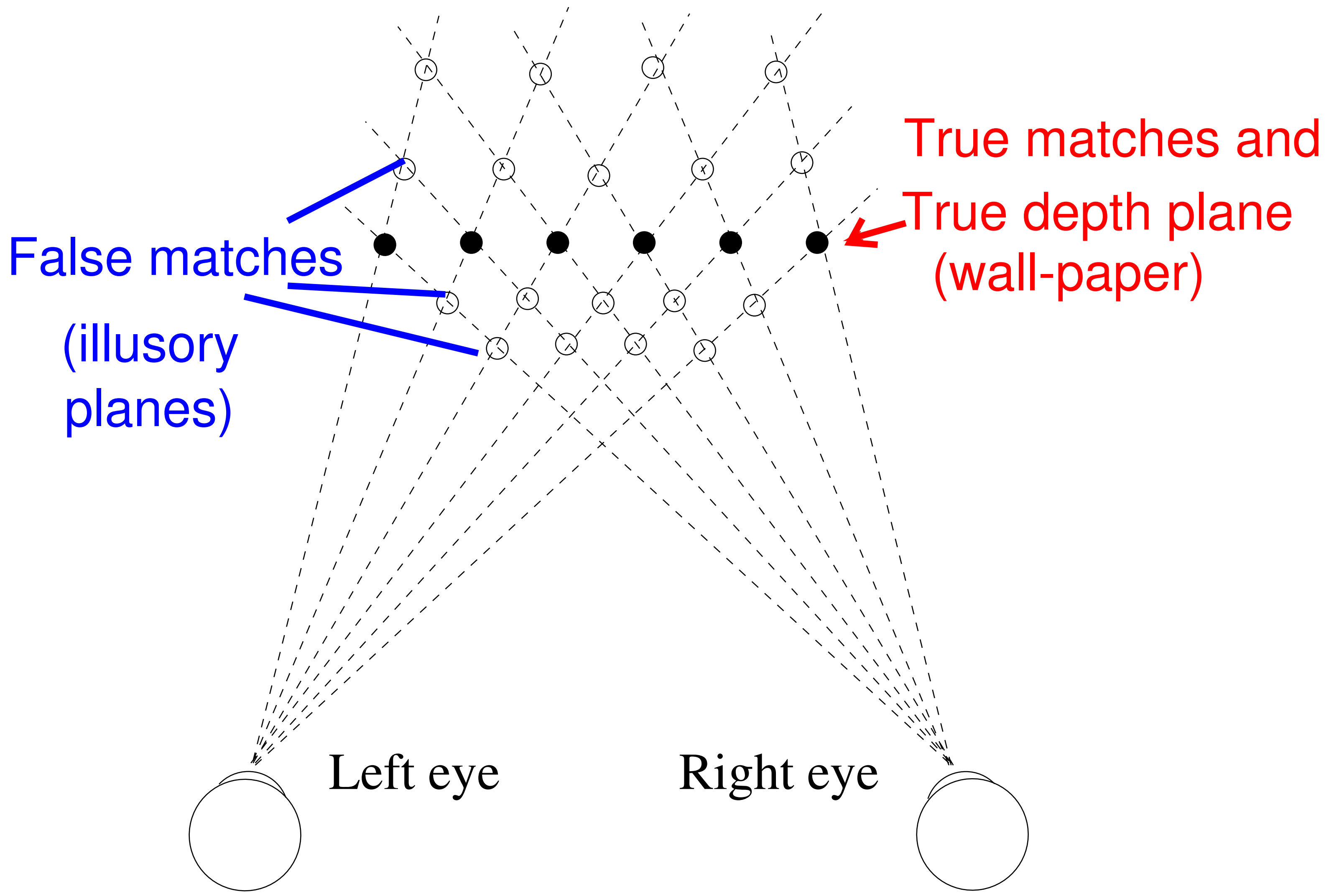
Purpose: to see if intracortical interactions can account for some stereo grouping phenomana in physiology and psychophysics

Stereo Grouping Phenomena modelled here:

- Enhanced V2 responses to stereo edges. (von der Heydt et al, 2000).
- Disparity capture (wall-paper effects) manifested by V2 responses (Bakin et al, 2000).
- Pop-out of a target of a unique depth from distractors of a different depth.
- Transparency.

The stereo matching problem

Example: the wall-paper effect



The visual input samples both the true and the false matches

V1 neurons respond to both the true and false matches.
(Cumming and Parker, 2000)

But the cortex has to compute the true matches only

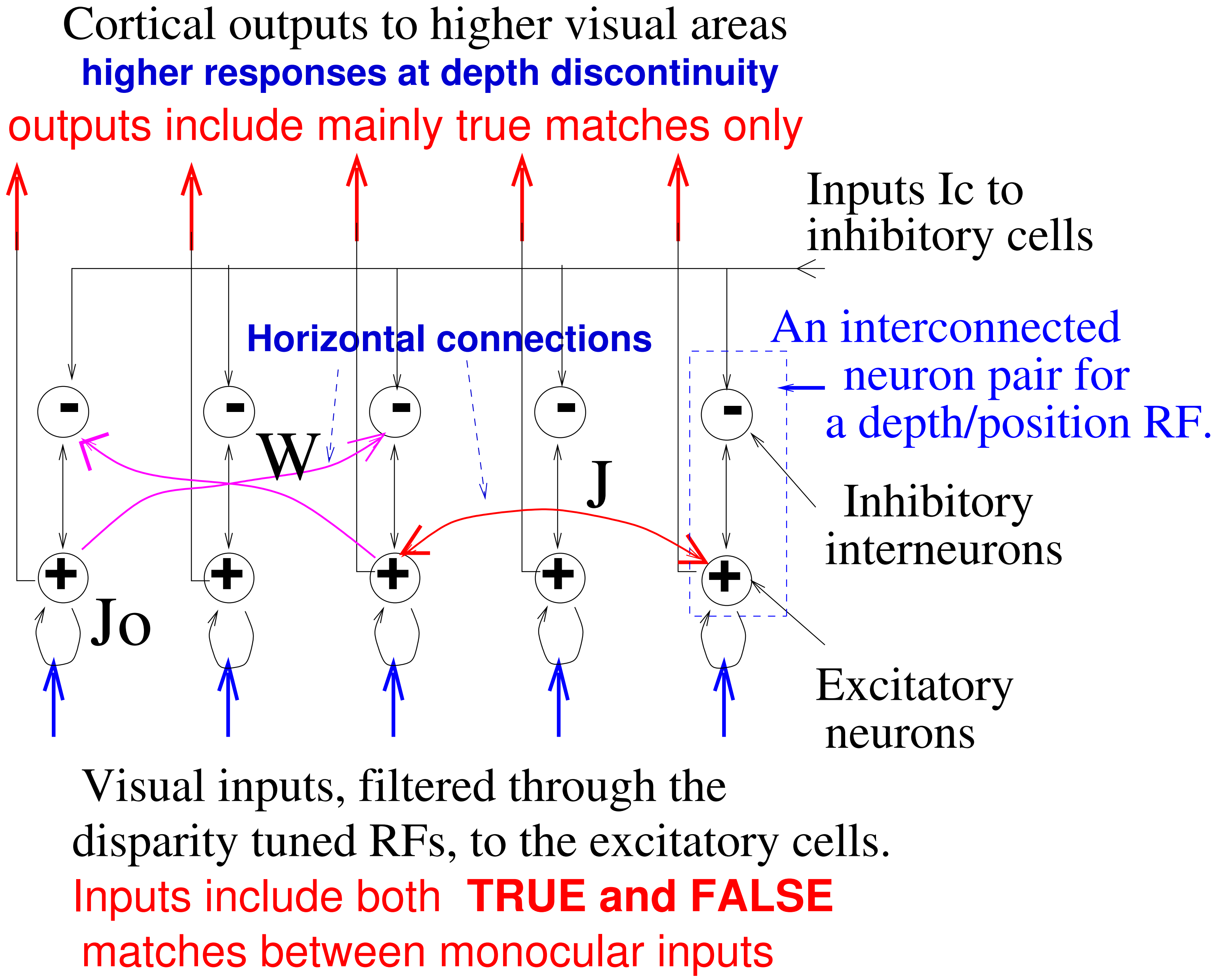
V2 neurons respond to only the true matches in the wall-paper effect (Bakin, Nakayama, and Gilbert, 2000 observed that a V2 neuron's response is tuned to the disparity value outside its receptive field, at the boundary of the depth plane (of the wall-paper-like grating.).

The stereo segmentation (grouping) problem

Detecting or highlighting discontinuities in depth, or depth edges, can serve segmentation, e.g., to segment two nearby planes of different depth, or to detect a target of a different depth — **pop-out**. Von der Heydt, Zhou, and Friedman, (2000) observed that **V2 cells respond more vigorously when their receptive fields are near a depth edge**.

Transparency: perceptually segregating two superimposed depth planes.

The model


Cortical outputs to higher visual areas
higher responses at depth discontinuity
outputs include mainly true matches only
Inputs Ic to
inhibitory cells
Horizontal connections
An interconnected
neuron pair for
a depth/position RF.
W
J
Jo
Inhibitory
interneurons
Excitatory
neurons
Visual inputs, filtered through the
disparity tuned RFs, to the excitatory cells.
Inputs include both TRUE and FALSE
matches between monocular inputs

Model features and elements

- The model aims to emulate intracortical computations in **V2**.

- **Each model unit is binocular and disparity tuned**. The excitatory cells model cortical pyramidal cells. Each excitatory cell couples with a local interneuron to form a disparity tuned unit with a finite (small) size receptive field. **Other input dimensions (e.g., orientation, color, motion) are omitted**.

- Different cells are tuned to different depths, the receptive fields (RFs) of all model cells **sample 3D visual space** (2D frontol-parallel and 1D depth).

- **Both true and false matches provide input to the model** (pyramidal) cells.

- **Long but finite range horizontal connections** mediate monosynaptic excitation and disynaptic inhibtion between nearby pyramidal cells. Horizontal connections tend to link cells tuned to similar depth.

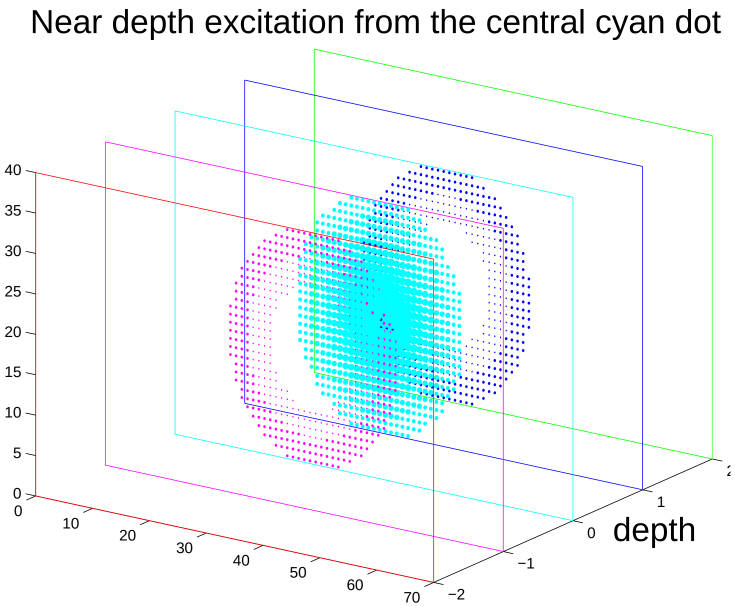
- **Cells responding to the same monocular location but different depths inhibit each other**.

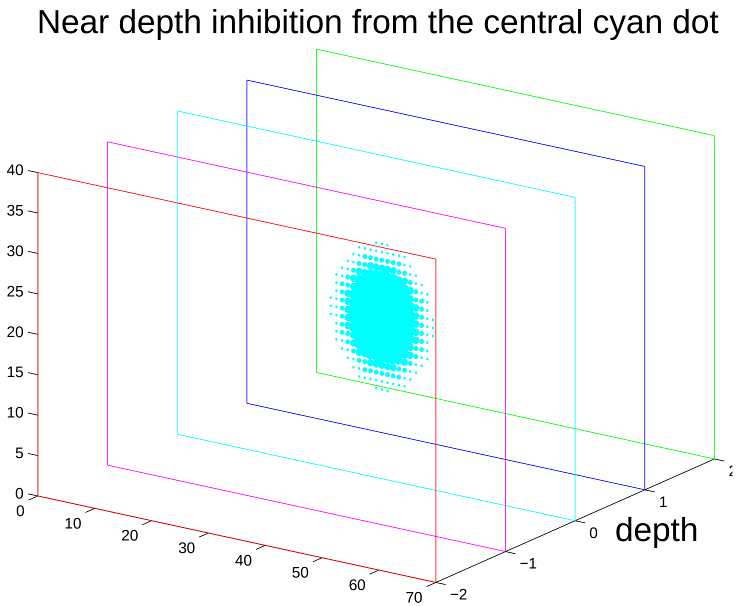
- Horizontal connections mediate contextual influences such that, after initial transients, **(1) the model cells respond significantly only to the true matches, and, (2) responses to depth discontinuities (depth edges or pop-out targets) are relatively higher.**

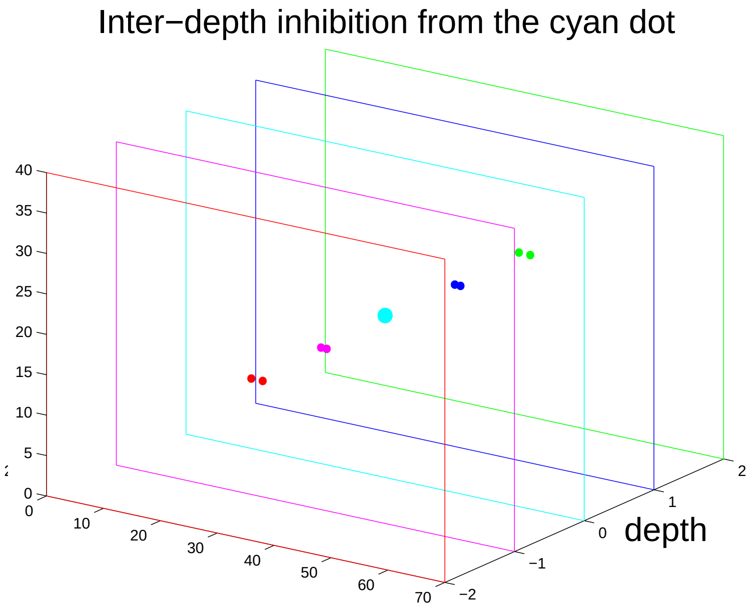
The model horizontal connection pattern

There are 3 components in the connections

Near depth excitation **Near depth inhibition** **Inter-depth suppression**







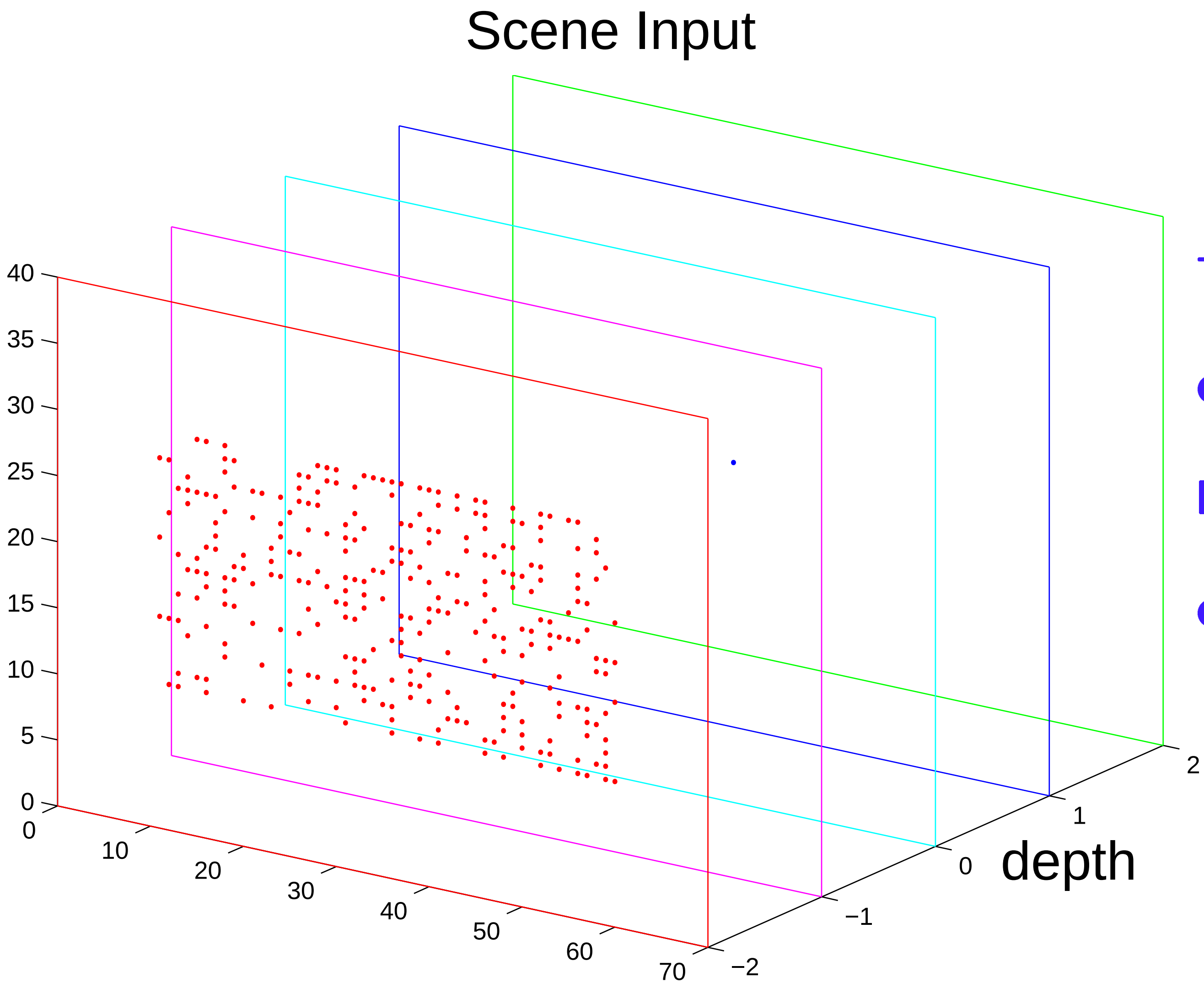
Model cells sample the visual space of 5 depth planes (5x70x40 3-d locations).
Different depth planes are color coded
Each dot on a depth plane represents an RF center.
Dot size codes interaction (or response) strength

The equations of motion:

$$\begin{aligned}
\dot{x}_{id} &= -\alpha_x x_{id} - g_y(y_{id}) + J_o g_x(x_{id}) + \sum_{j,d' \neq i,d} J_{id,jd'} g_x(x_{jd'}) + I_{id} + I_o \\
\dot{y}_{id} &= -\alpha_y y_{id} + g_x(x_{id}) + \sum_{j,d' \neq i,d} W_{id,jd'} g_x(x_{jd'}) + I_c
\end{aligned}$$

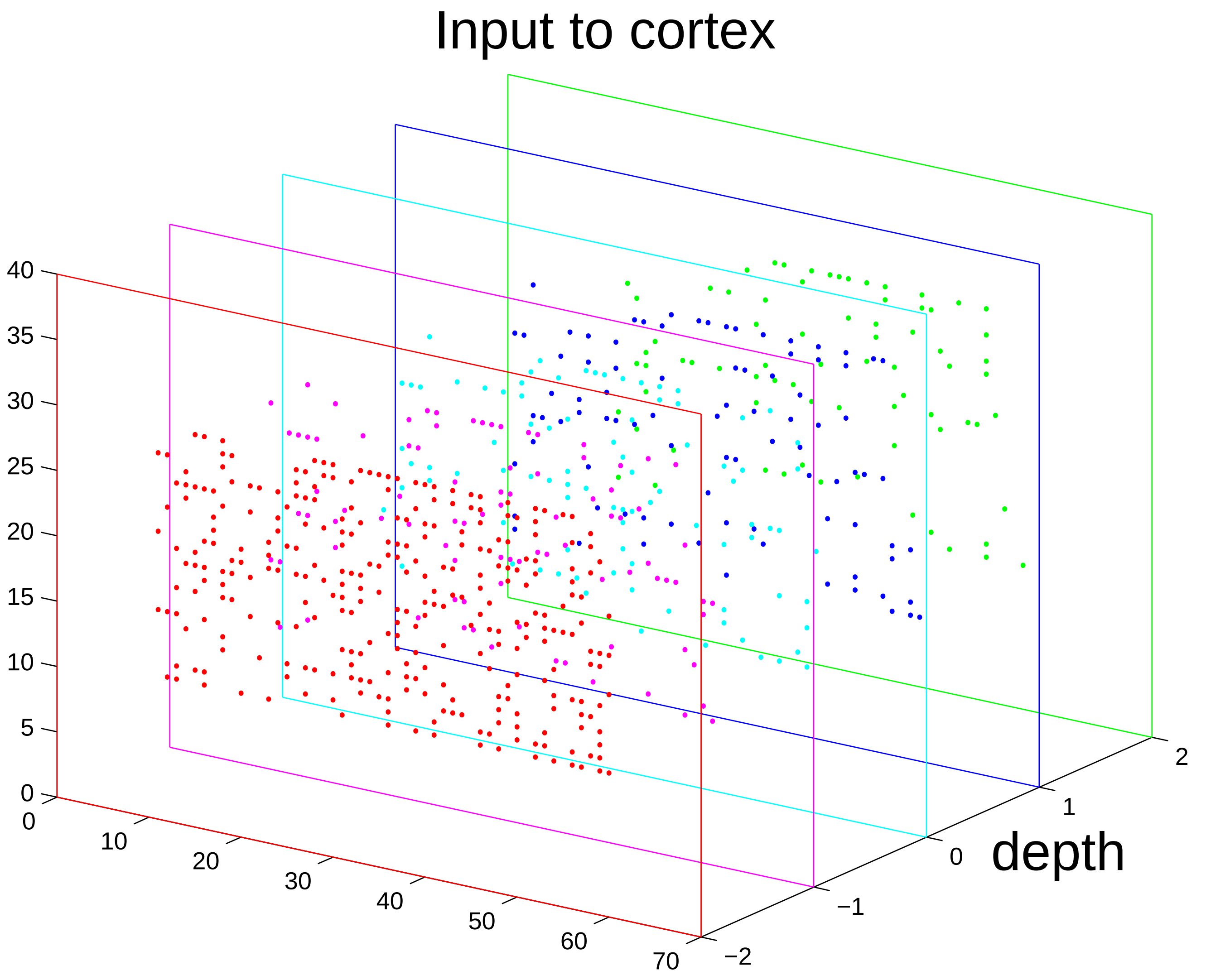
x or y: membrane potential for excitatory or inhibitory cells, i, d index for frontal parallel location i and depth d, J, W, horizontal connection matrix, g sigmoid-like activation functions, I_{id} visual inputs, etc.

Model computation illustrated by Popout



The scene consists of one depth plane and a single lone target dot at a different (blue) depth

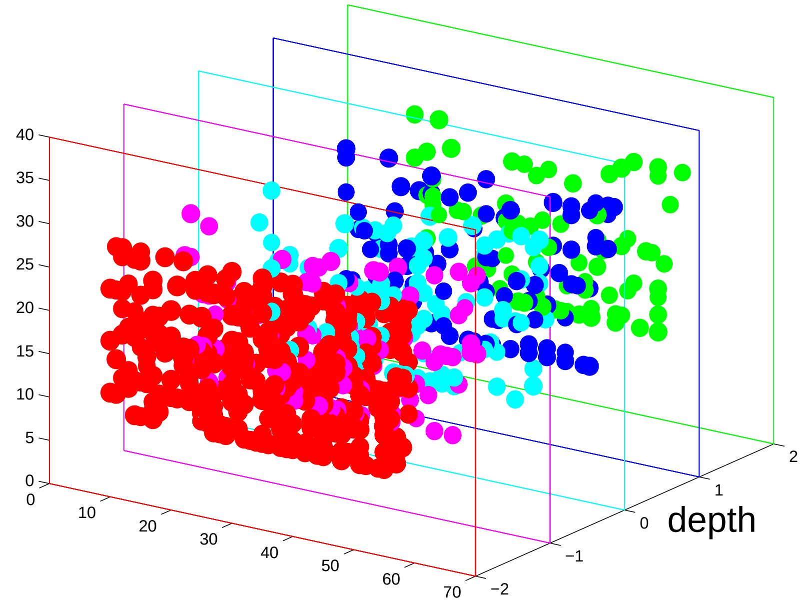
Input to the model



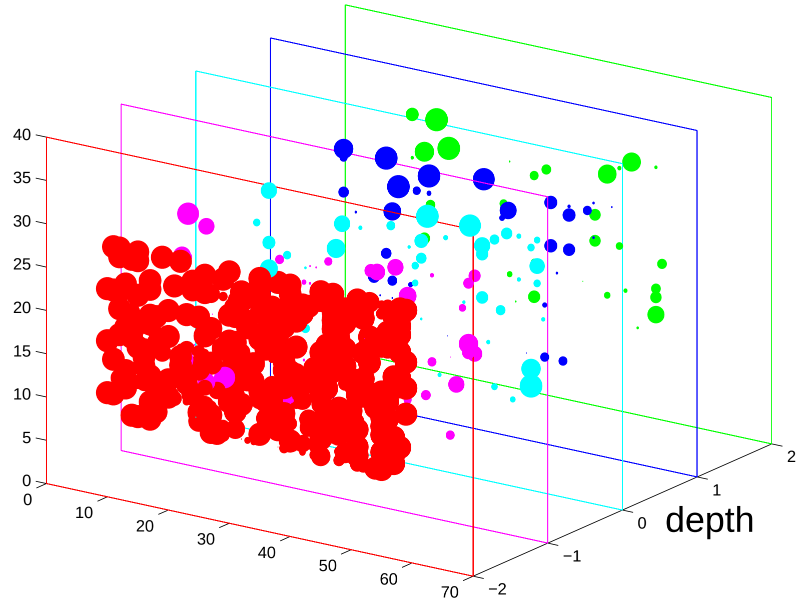
contains true and false matches

Evolution of activity with time

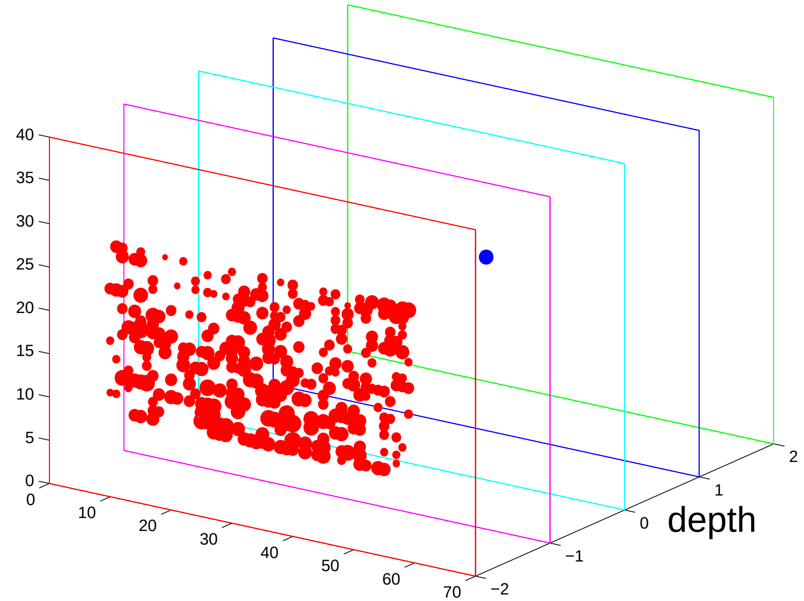
cortical activity at 0.60 membrane time constants



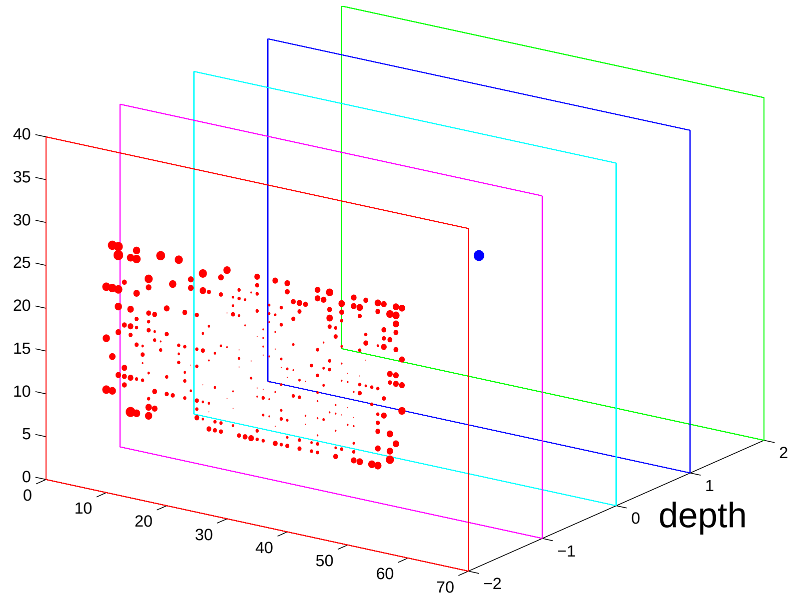
cortical activity at 1.40 membrane time constants



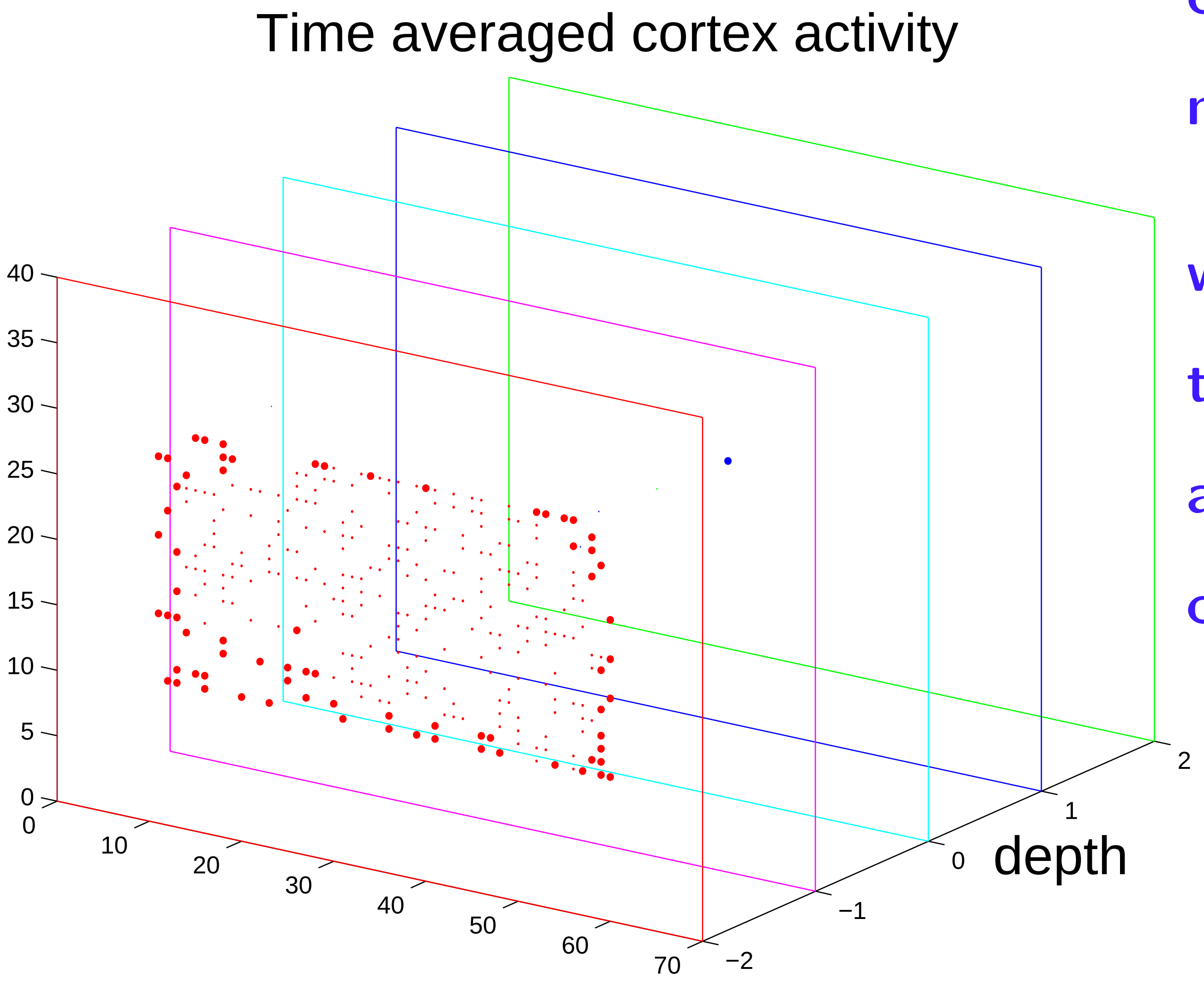
cortical activity at 6.00 membrane time constants



cortical activity at 12.80 membrane time constants



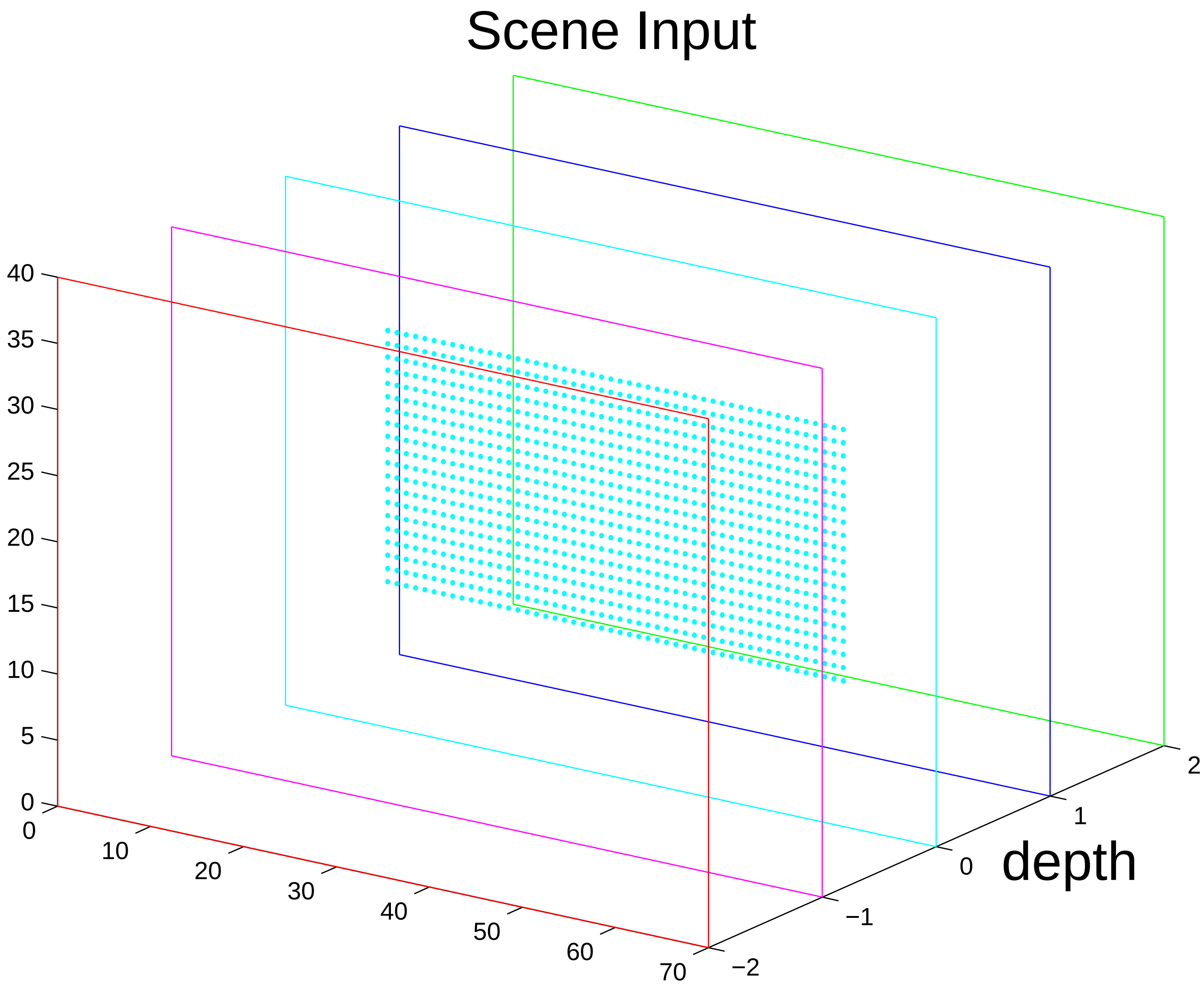
After initial transients, time averaged model response



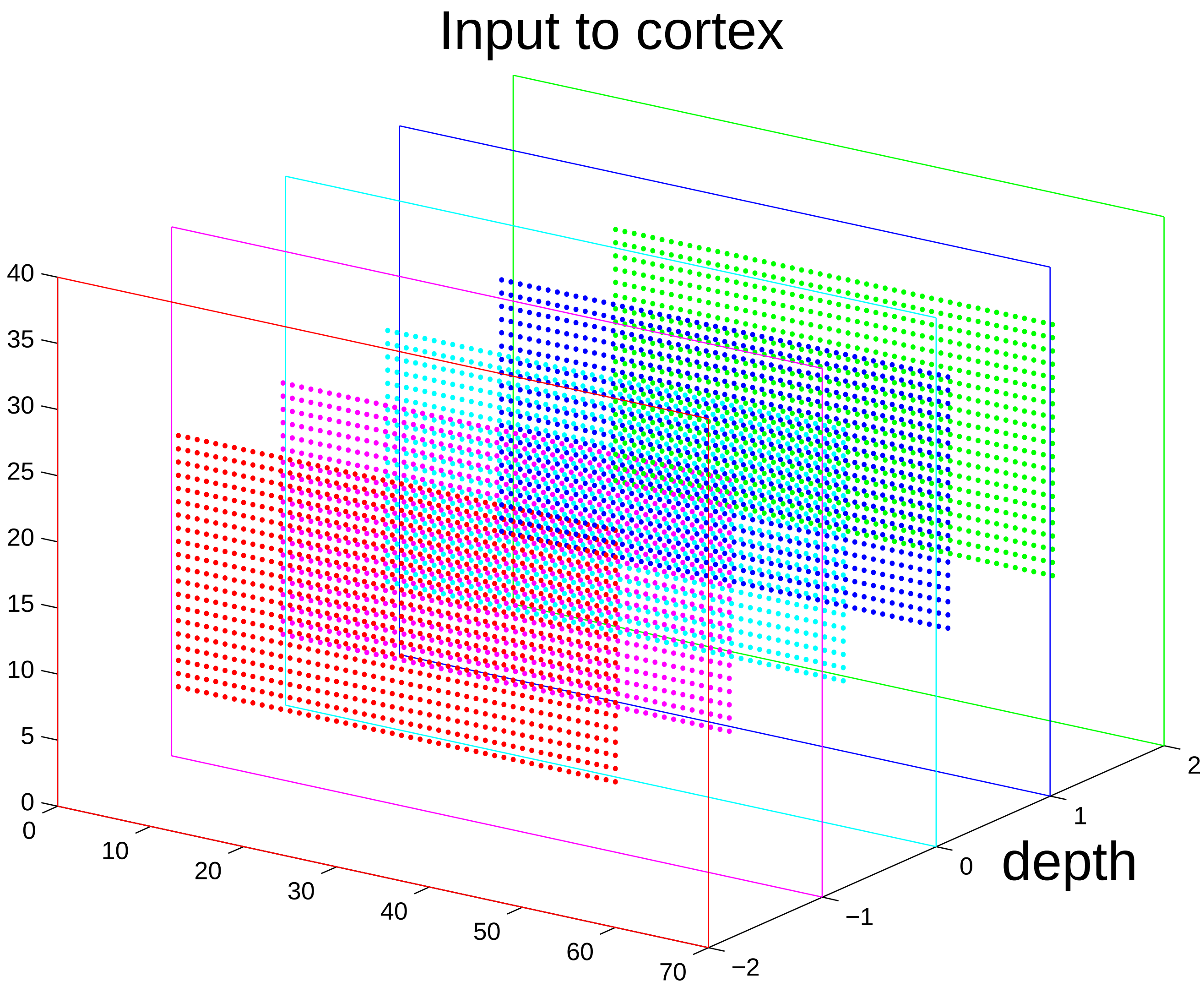
contains mainly true matches

with higher responses to the lone target — popout and to the borders of the depth plane

Model computation for Disparity Capture

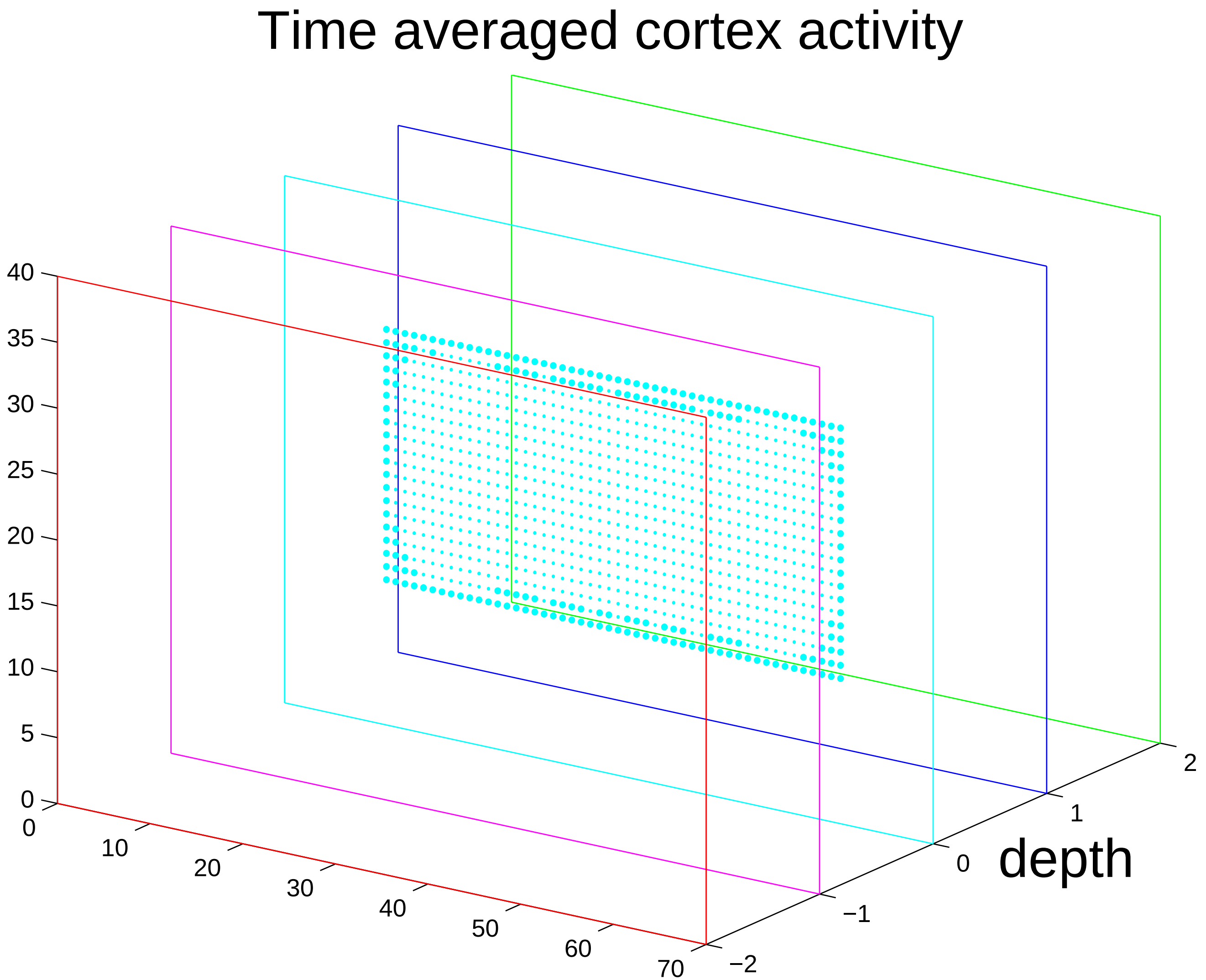


Input to the model



contains true and illusory depth planes from true and false matches

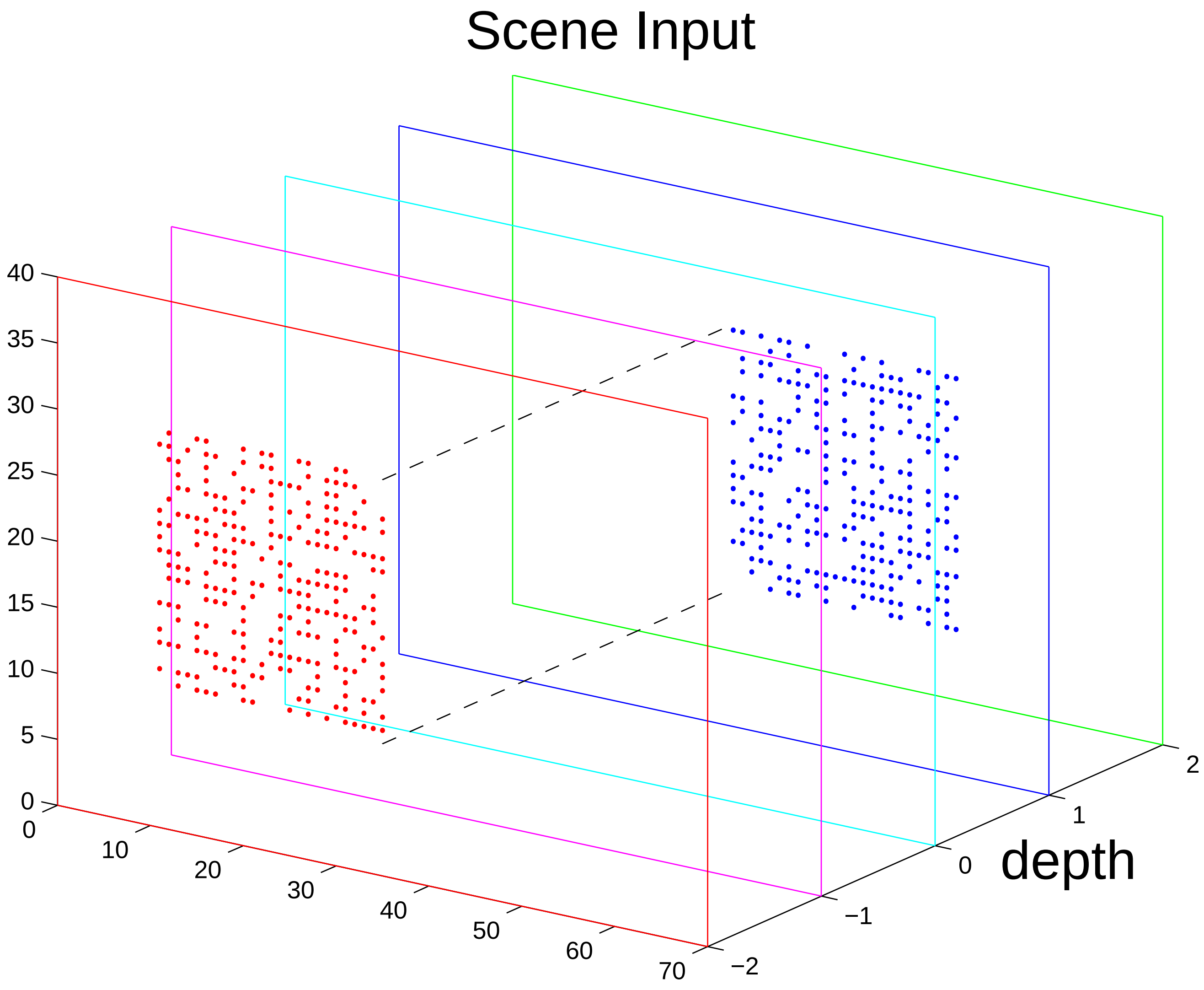
After initial transients, time averaged model response



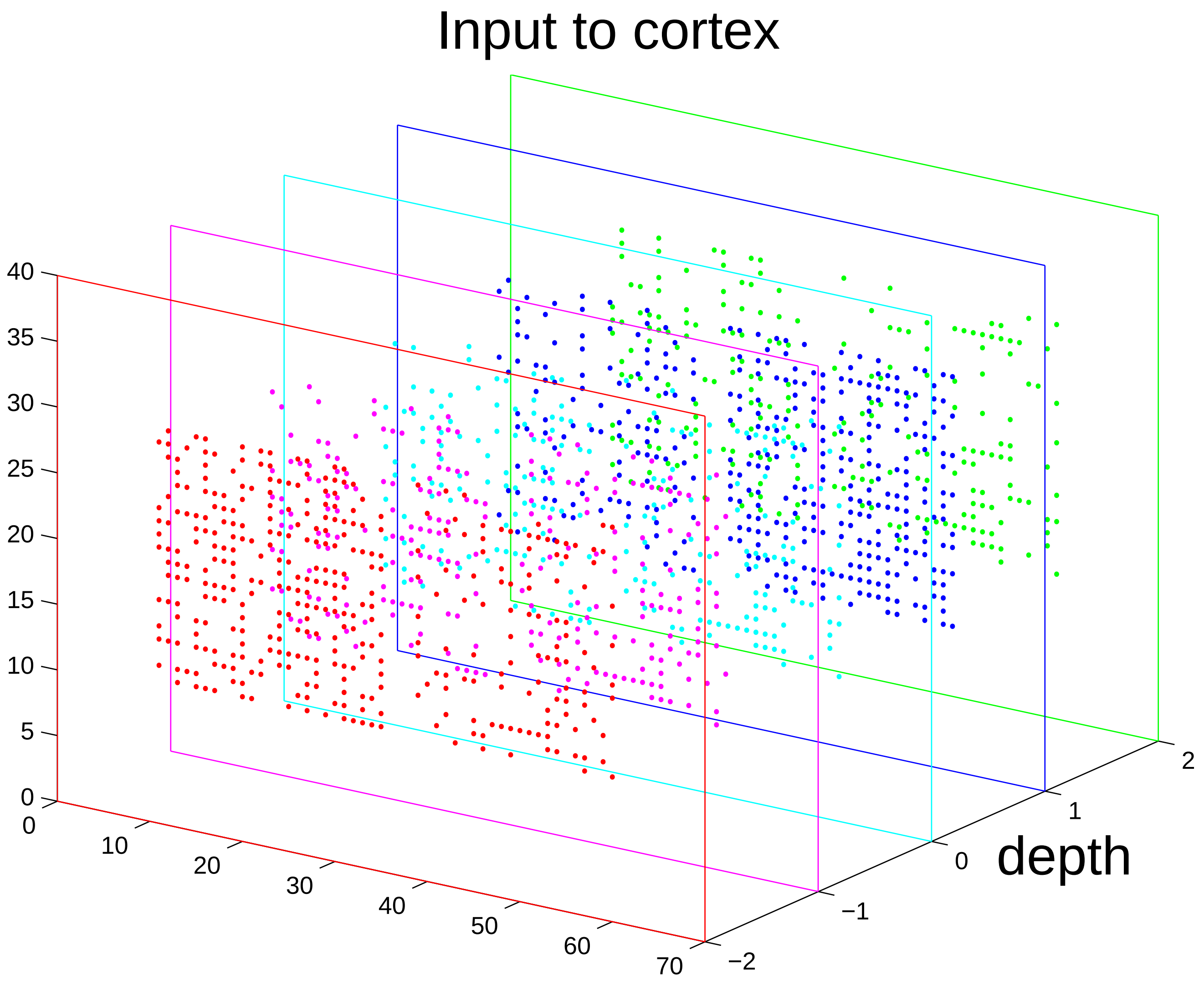
Higher responses to the borders of the plane

Contextual influences enable disparity information at the boundaries to propagate into the center of the plane

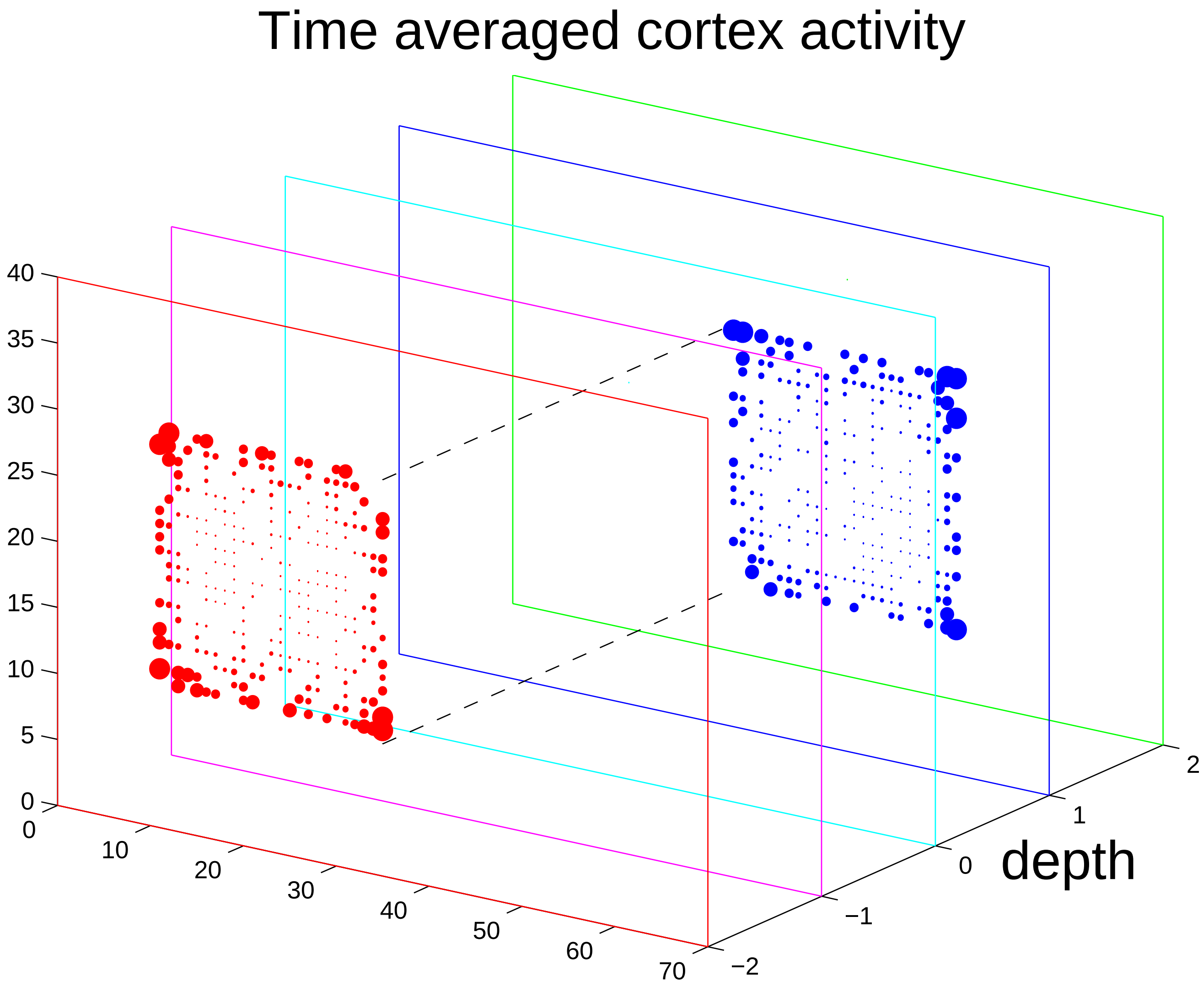
Model computation for Depth Discontinuity



Input to the model

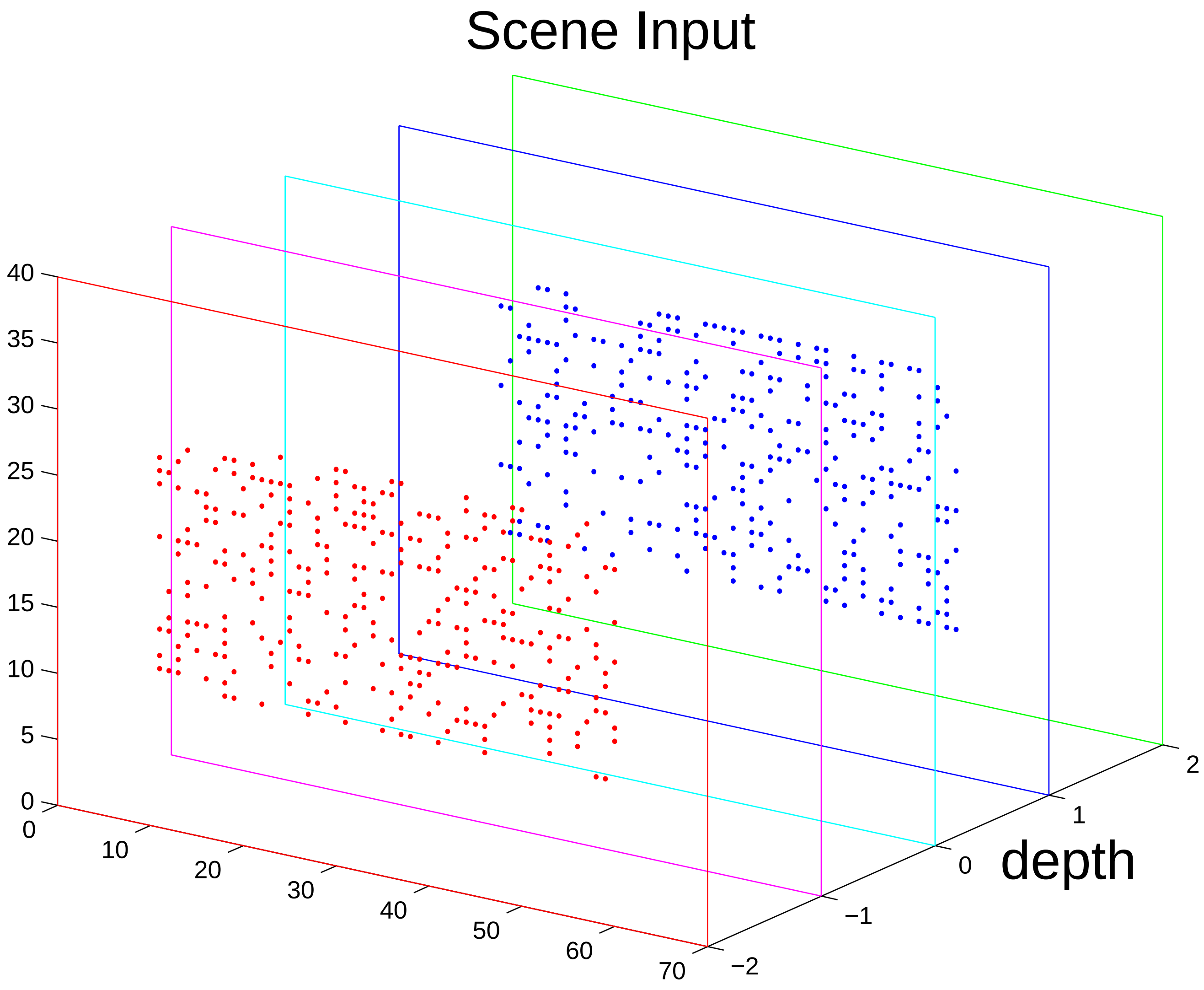


After initial transients, time averaged model response

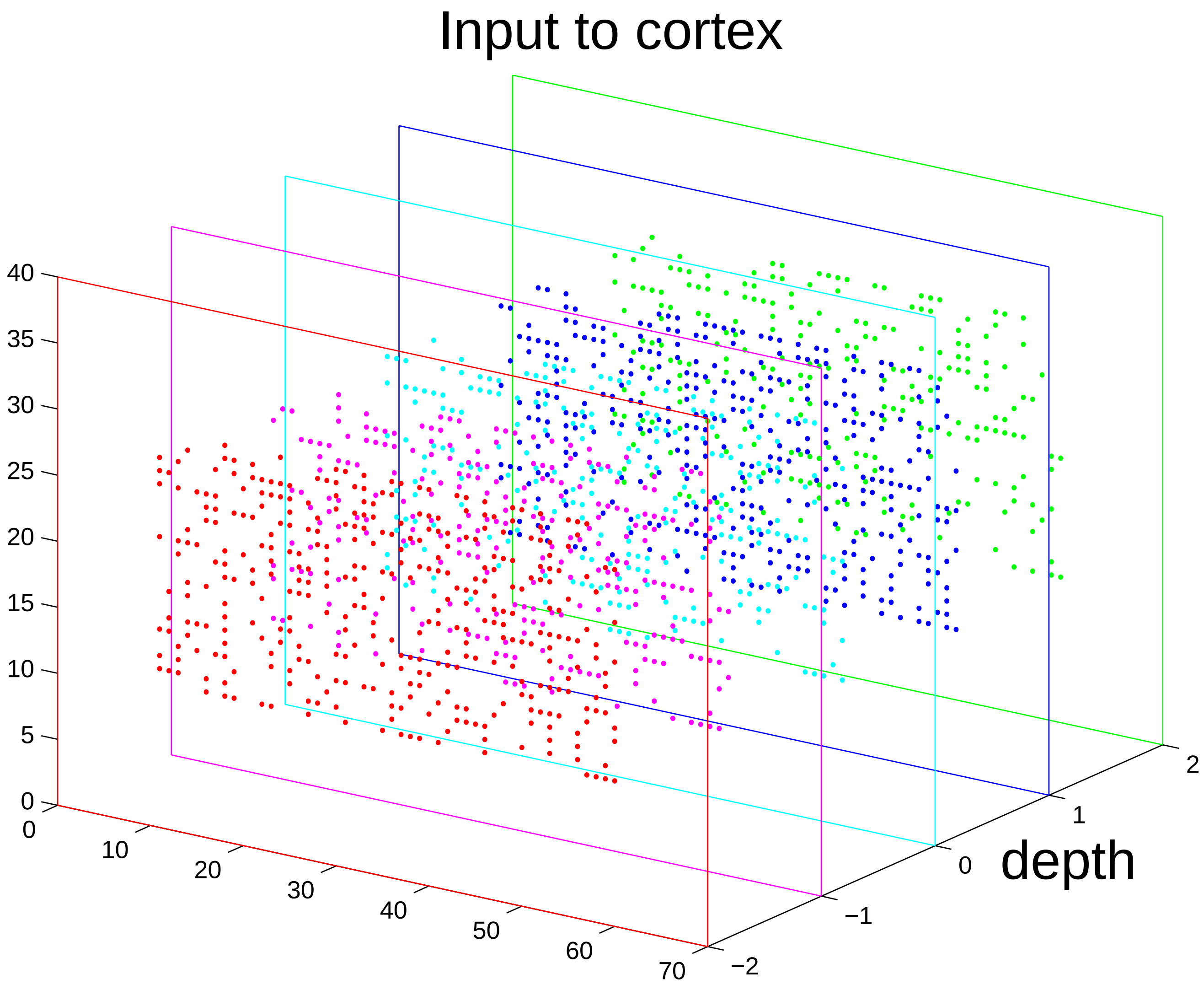


Higher responses to the depth discontunity and boundaries

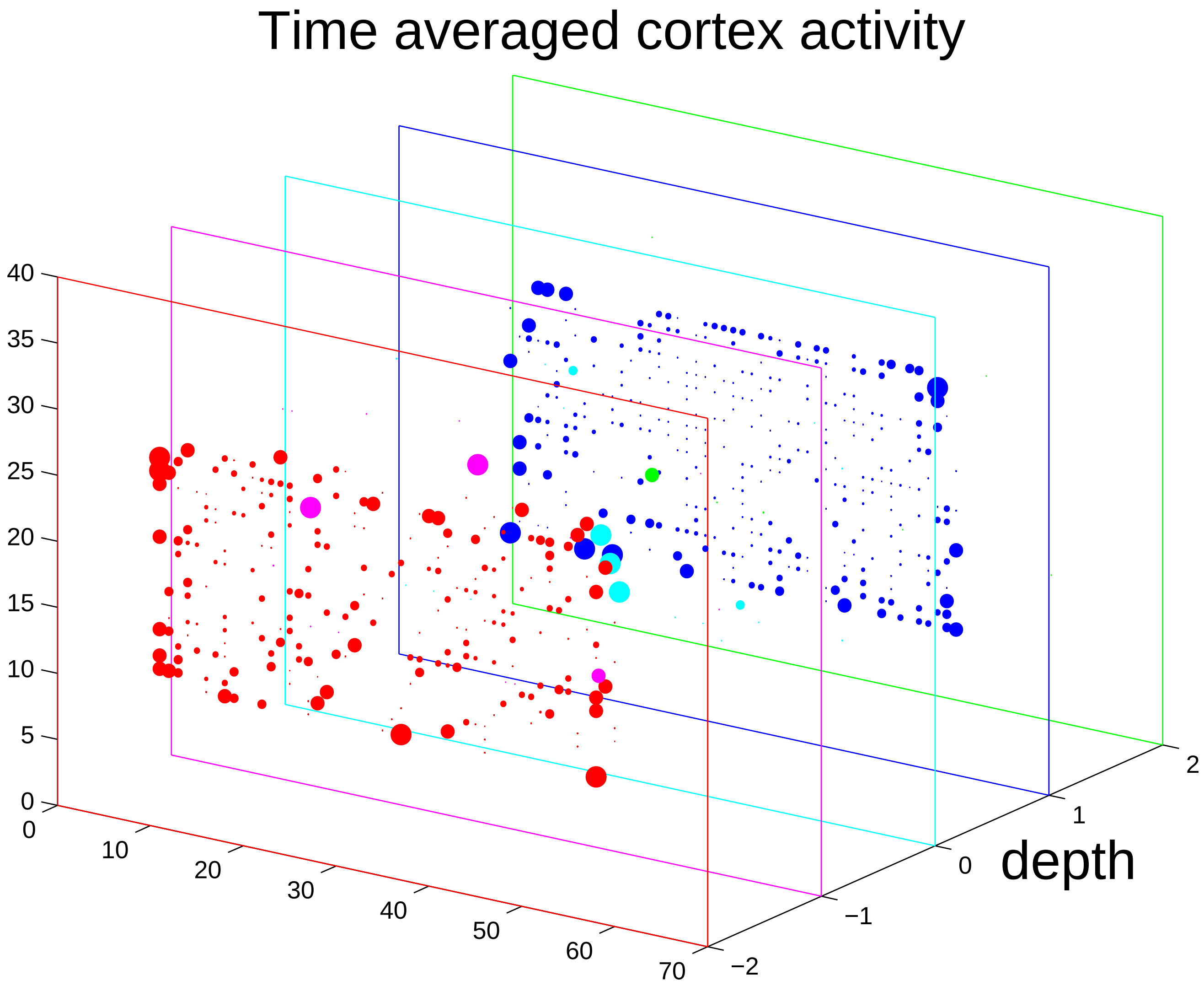
Model computation on Transparency



Input to the model



After initial transients, time averaged model response



Depth segregation

Higher responses to the plane boundaries

plus ghost dots

Summary and Discussion

- Aim to capture both physiology and psychophysics of stereo grouping.

- Suggest contextual influences in early cortex play important roles.

- Relating to previous models: (1) Use cooperative algorithms like previous models (e.g., Marr and Poggio); (2) popout and depth edge highlights modelled for the first time; (3) more physiologically realistic mechanisms for transparency than previous models (e.g., Prazdny, 1985, Pollard et al 1985, Nishihara, 1987, Qian and Sejnowski, 1989, Marshall et al 1996).

Thanks to Peter Dayan, Yury Petrov, and Ariella Popple for discussions and comments, Gatsby Foundation for support.
Comments and feedbacks welcome: z.li@ucl.ac.uk
www.gatsby.ucl.ac.uk/~zhaoping

Due to unexpected difficulties to obtain a U.S. visa for travel, the author is very sorry not to be able to attend this VSS meeting. Many thanks to Ariella Popple for her help to present this poster.