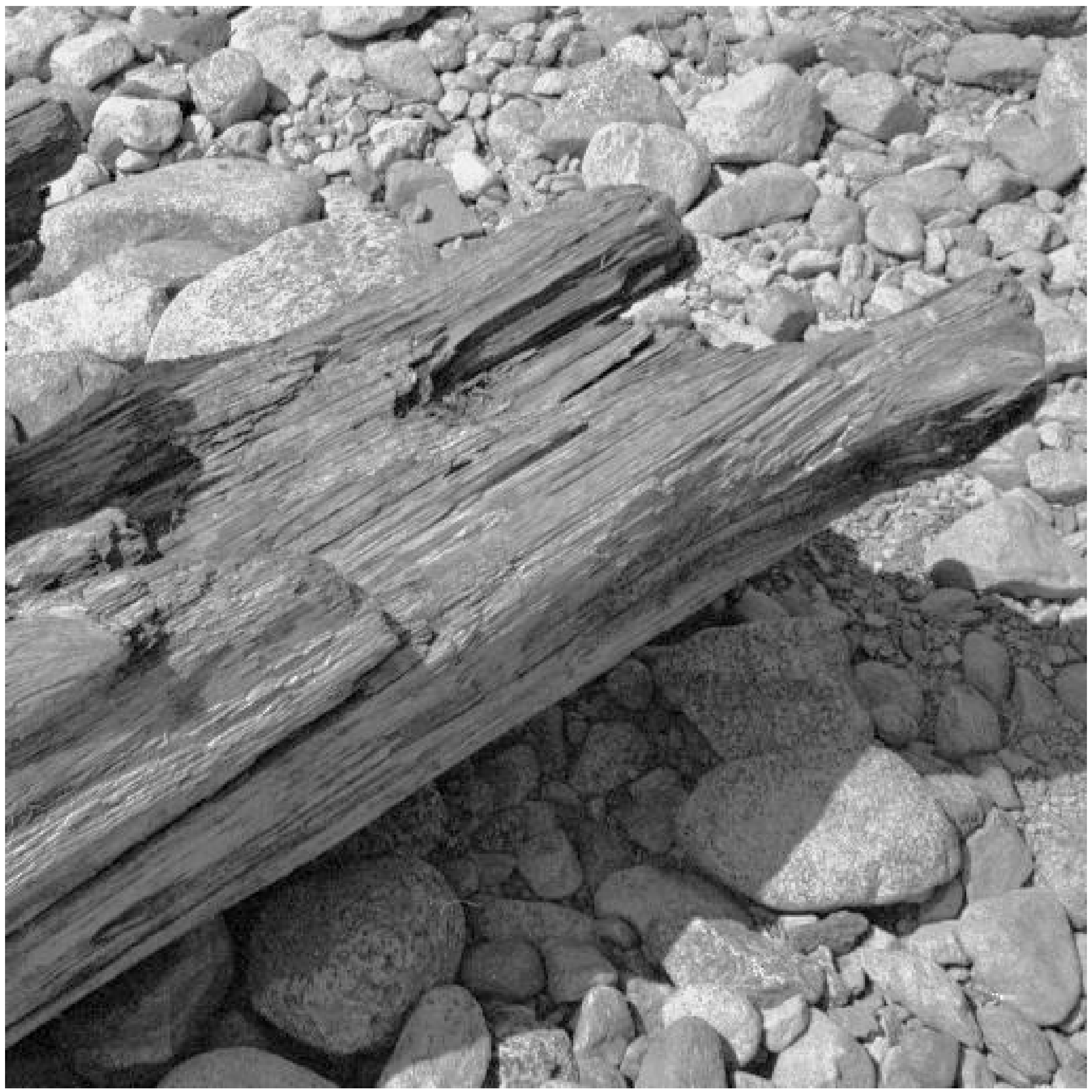
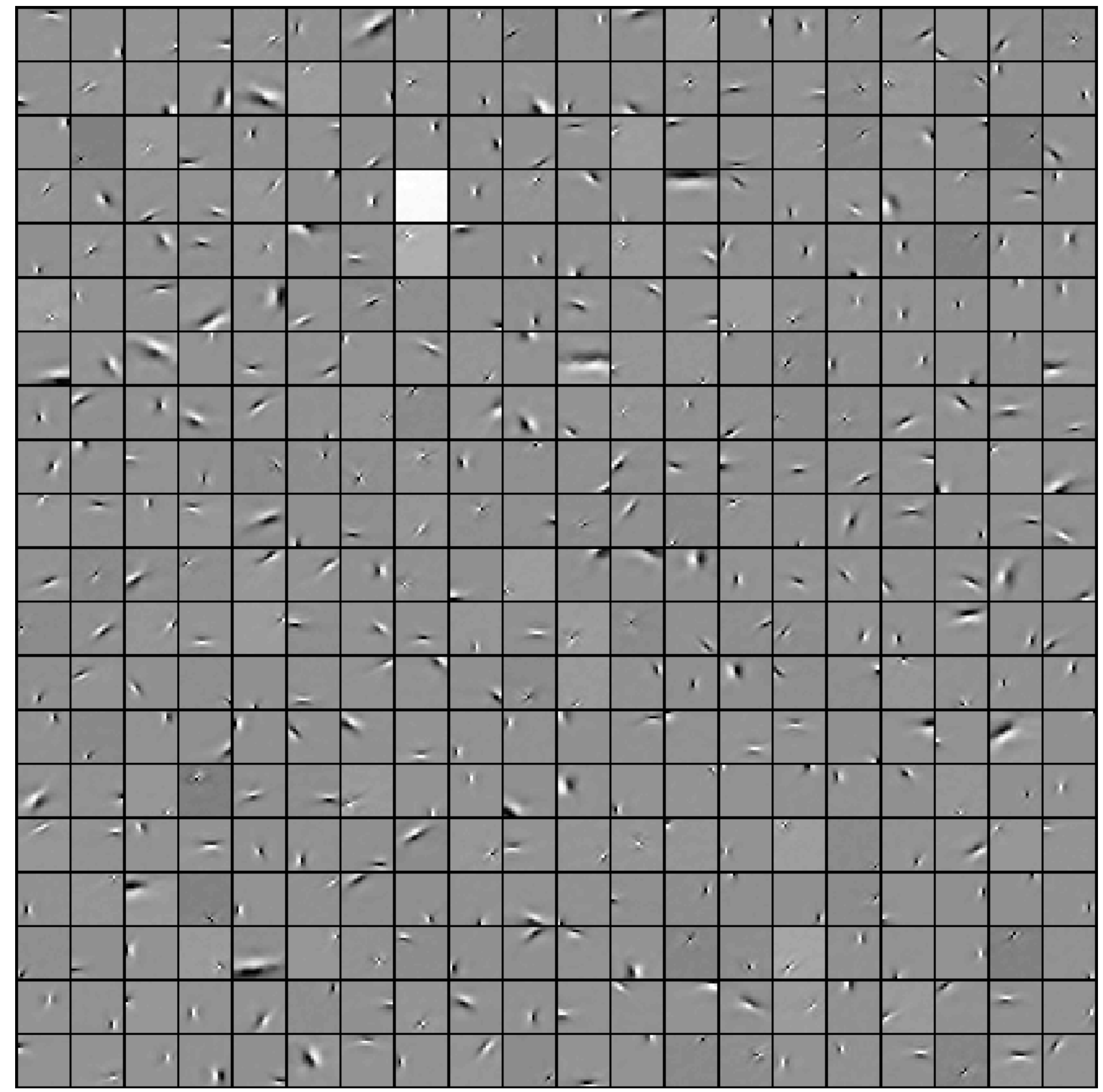
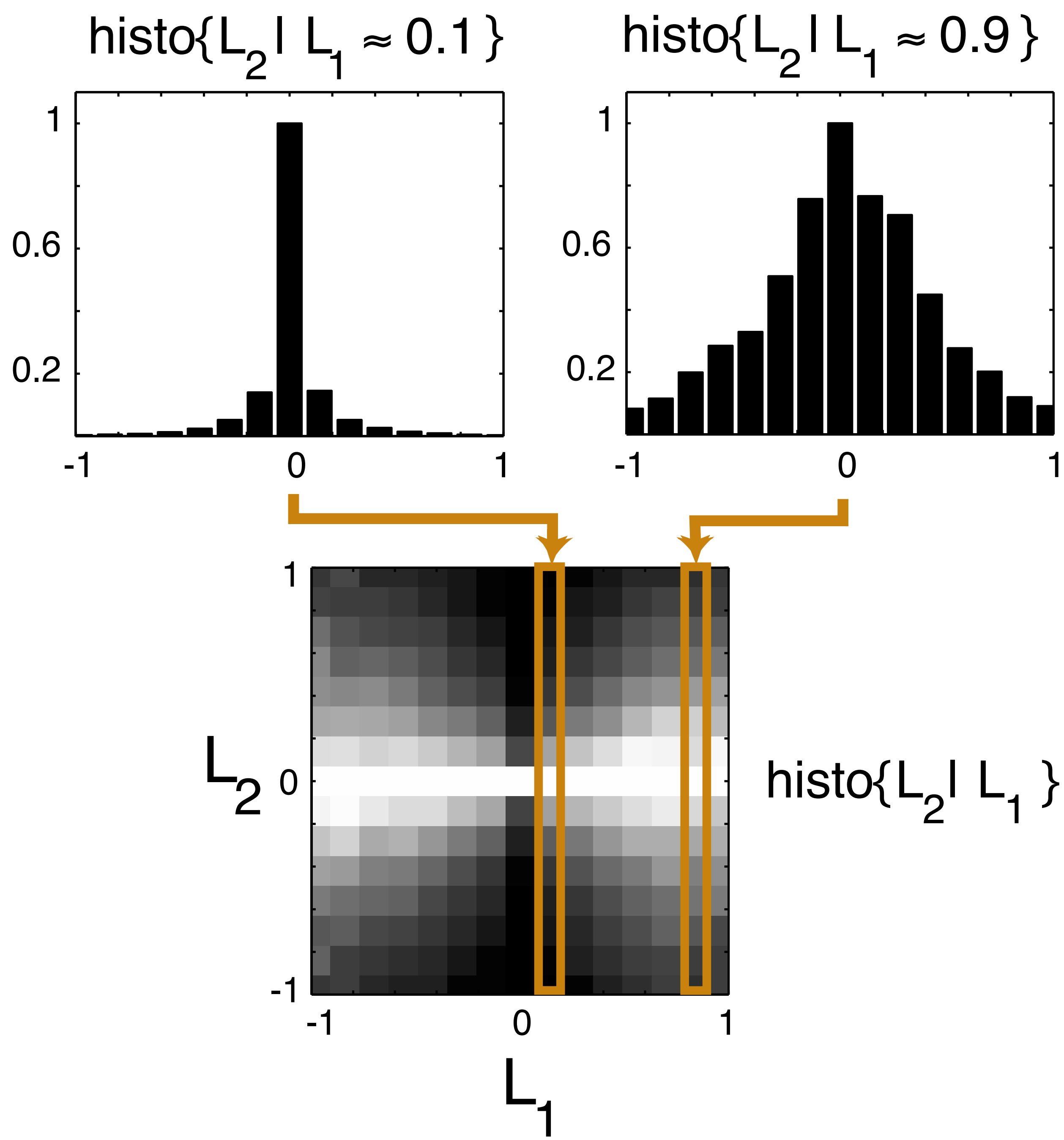
Part 4: Higher-order Image Structure

What happens next?





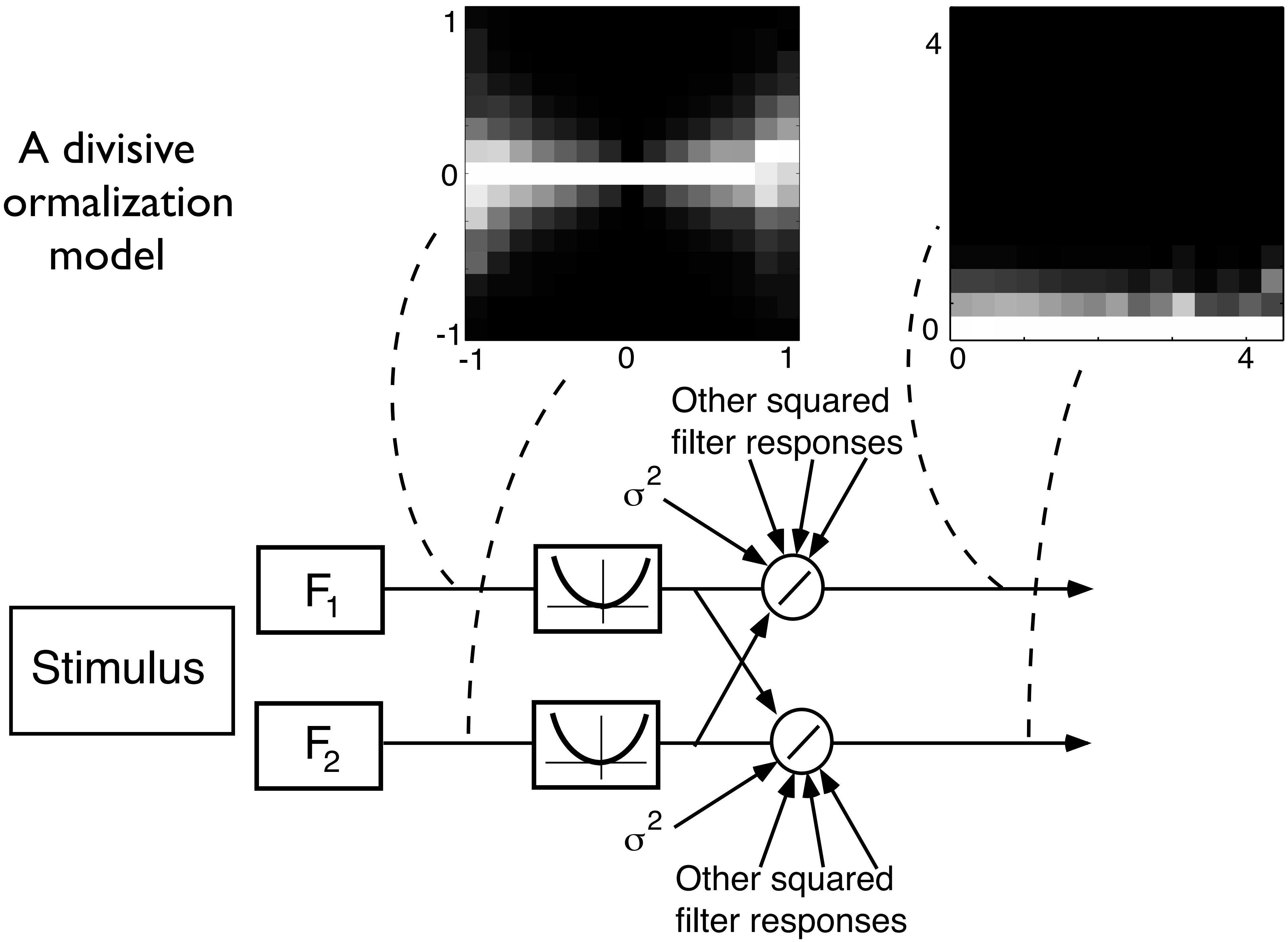
Joint statistics of filter outputs show magnitude dependence



from Schwartz and Simoncelli (2001)

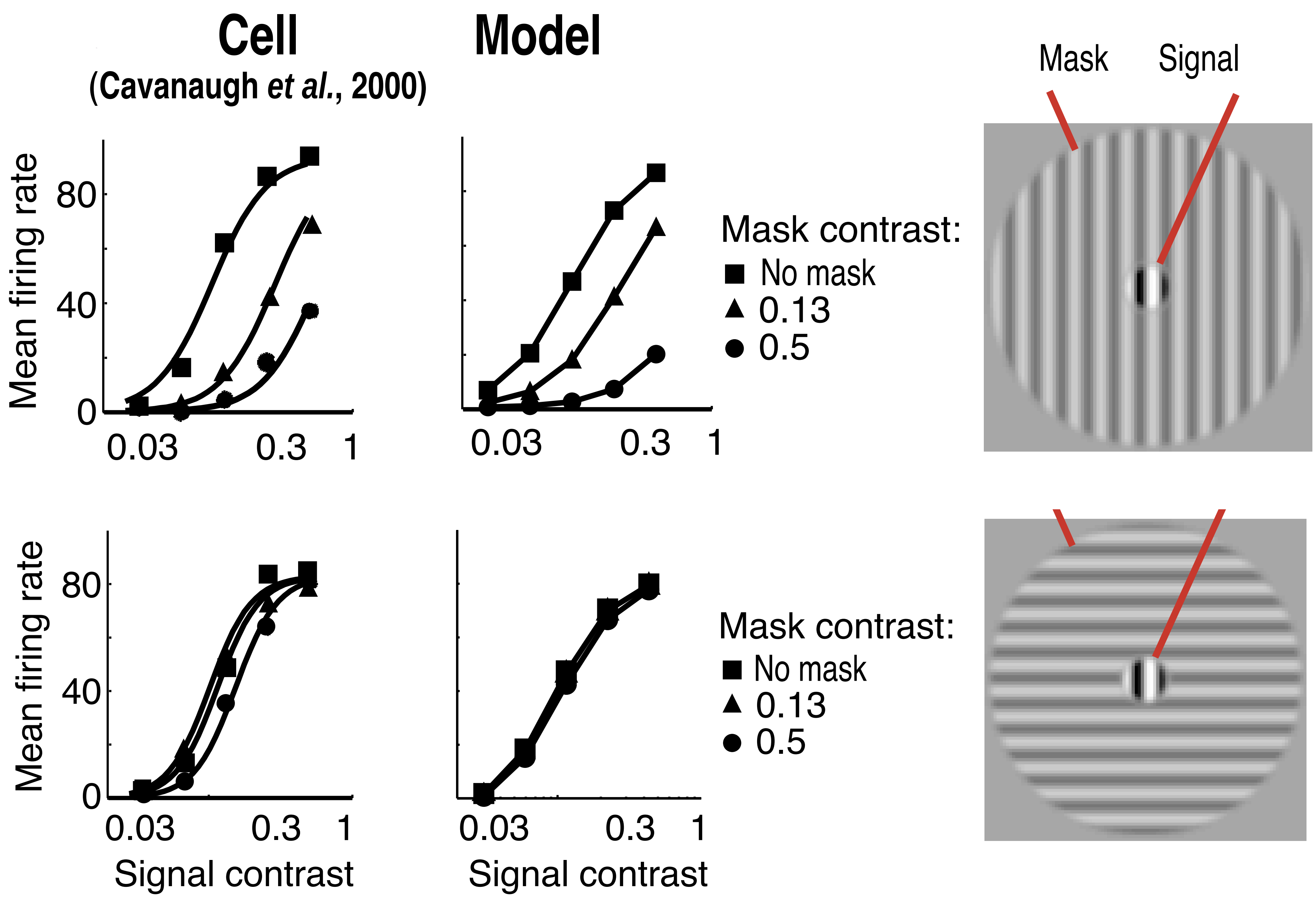
Using sensory gain control to reduce redundancy

A divisive normalization model



from Schwartz and Simoncelli (2001)

Model accounts for several non-linear response properties



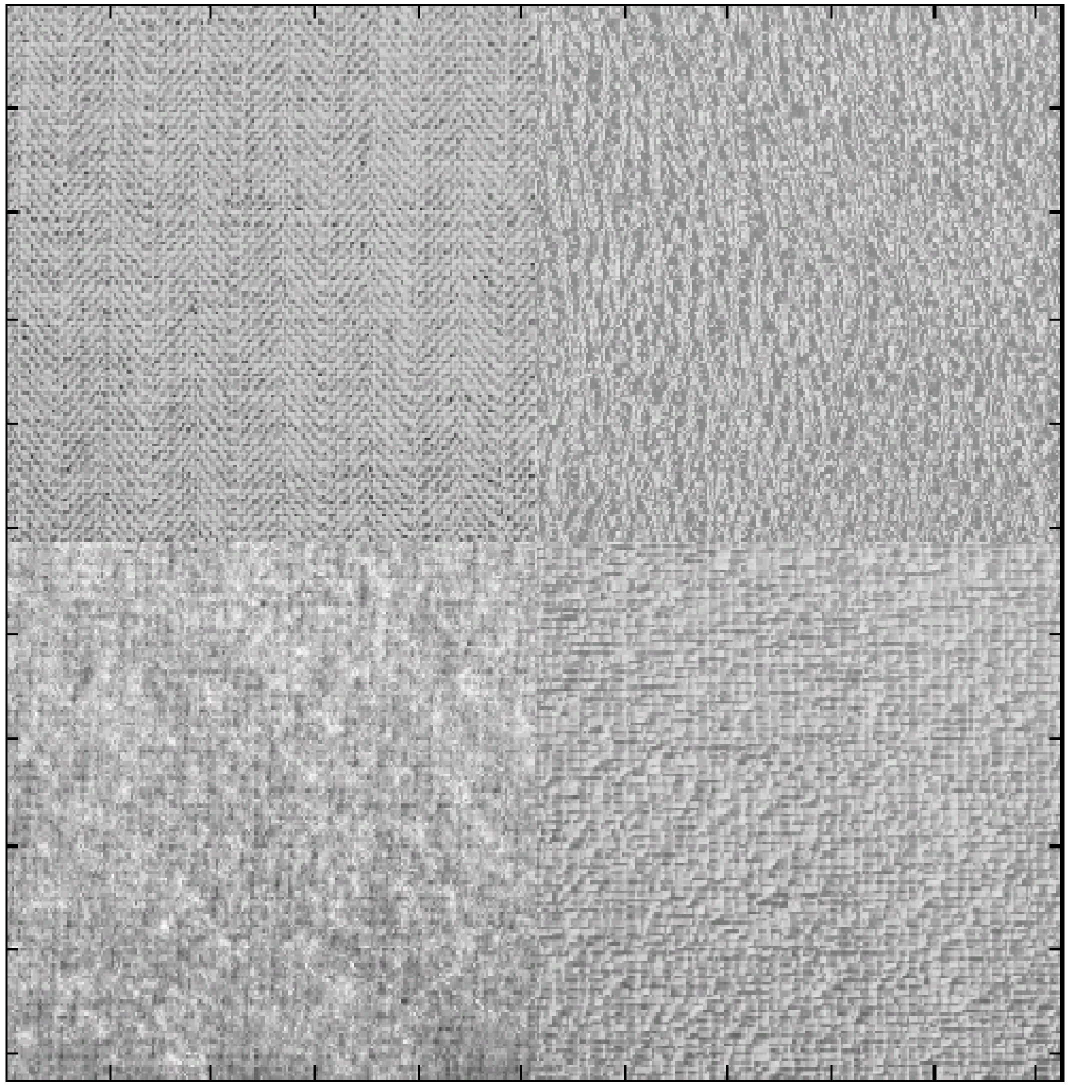
What about image structure?

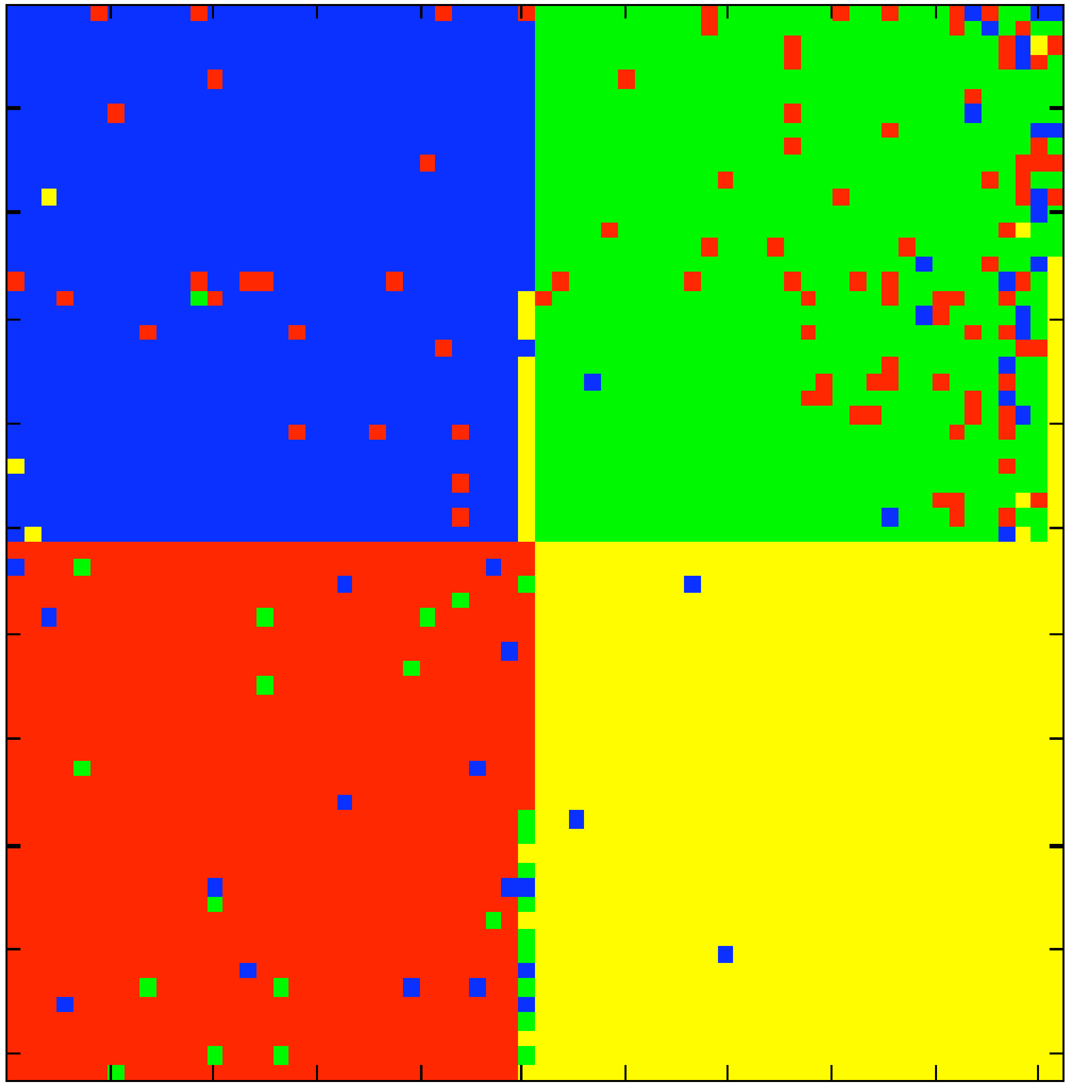
- Bottom-up approaches focus on the non-linearity.
- Our aim here is to focus on the *computational* problem:

  How do we learn the intrinsic dimensions of natural image structure?

- Idea: characterize how the local image *distribution* changes.

ICA mixtures for similar images



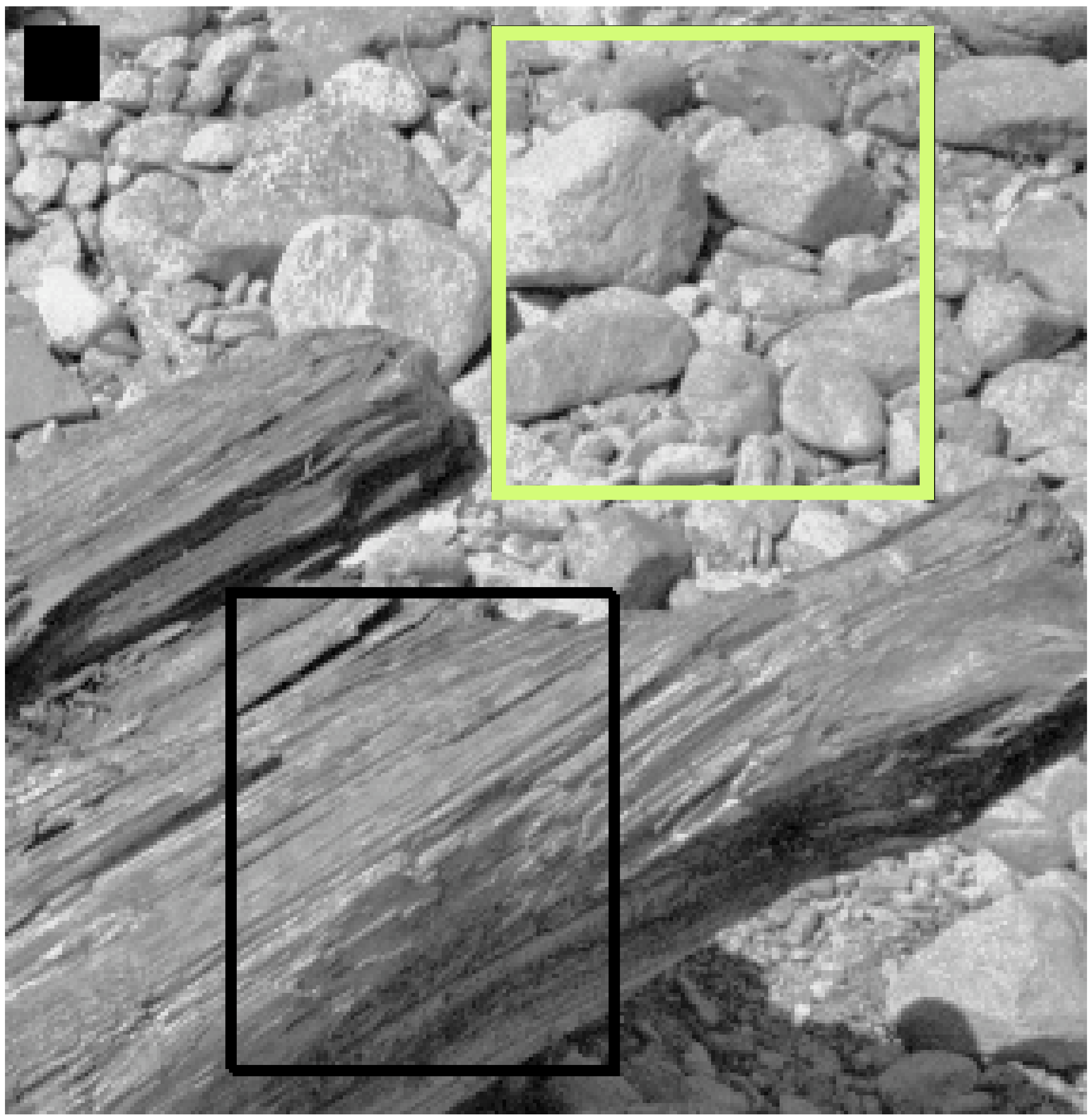


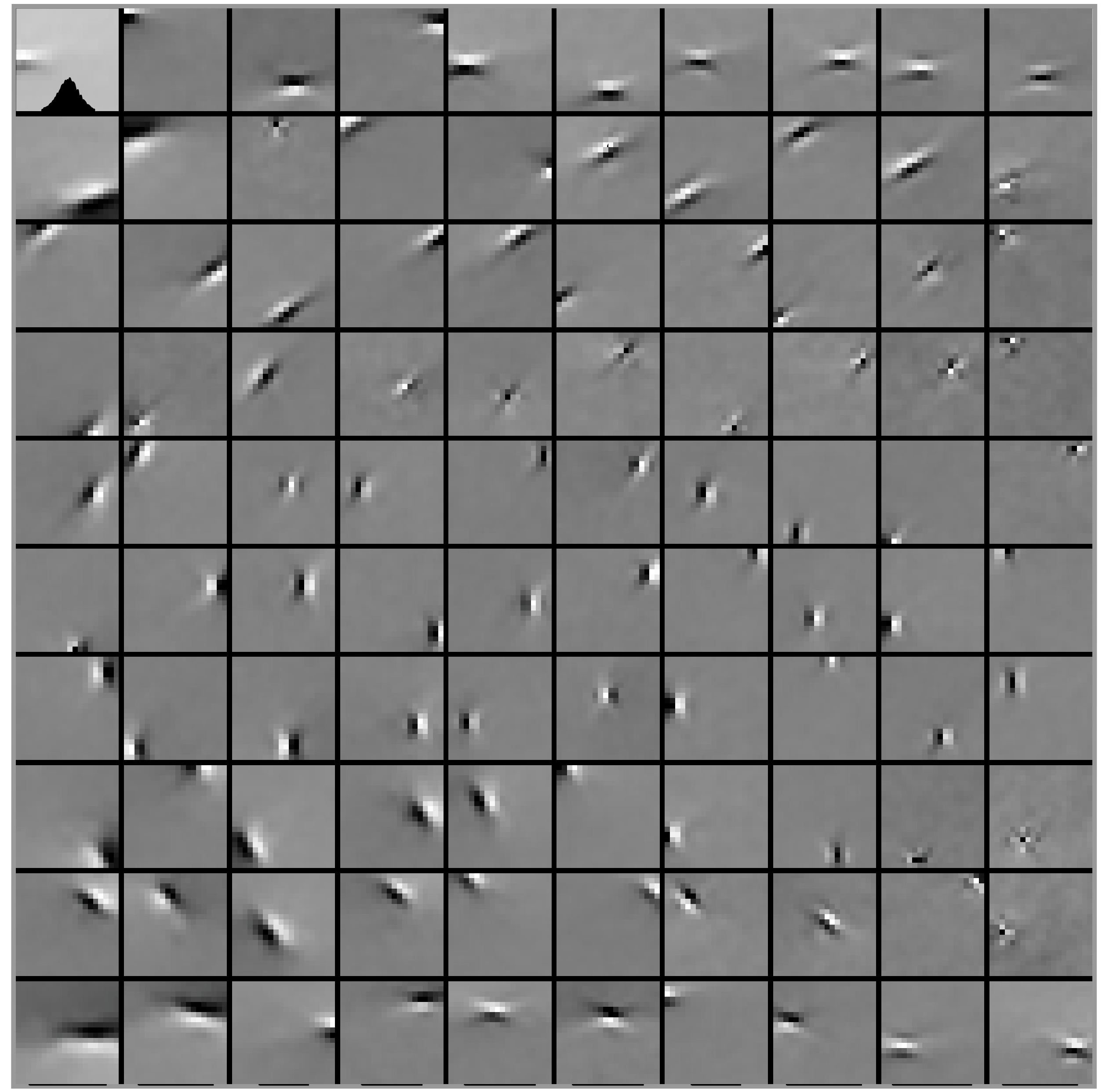
from Lee, Lewicki, and Sejnowski 2000

Limitations:

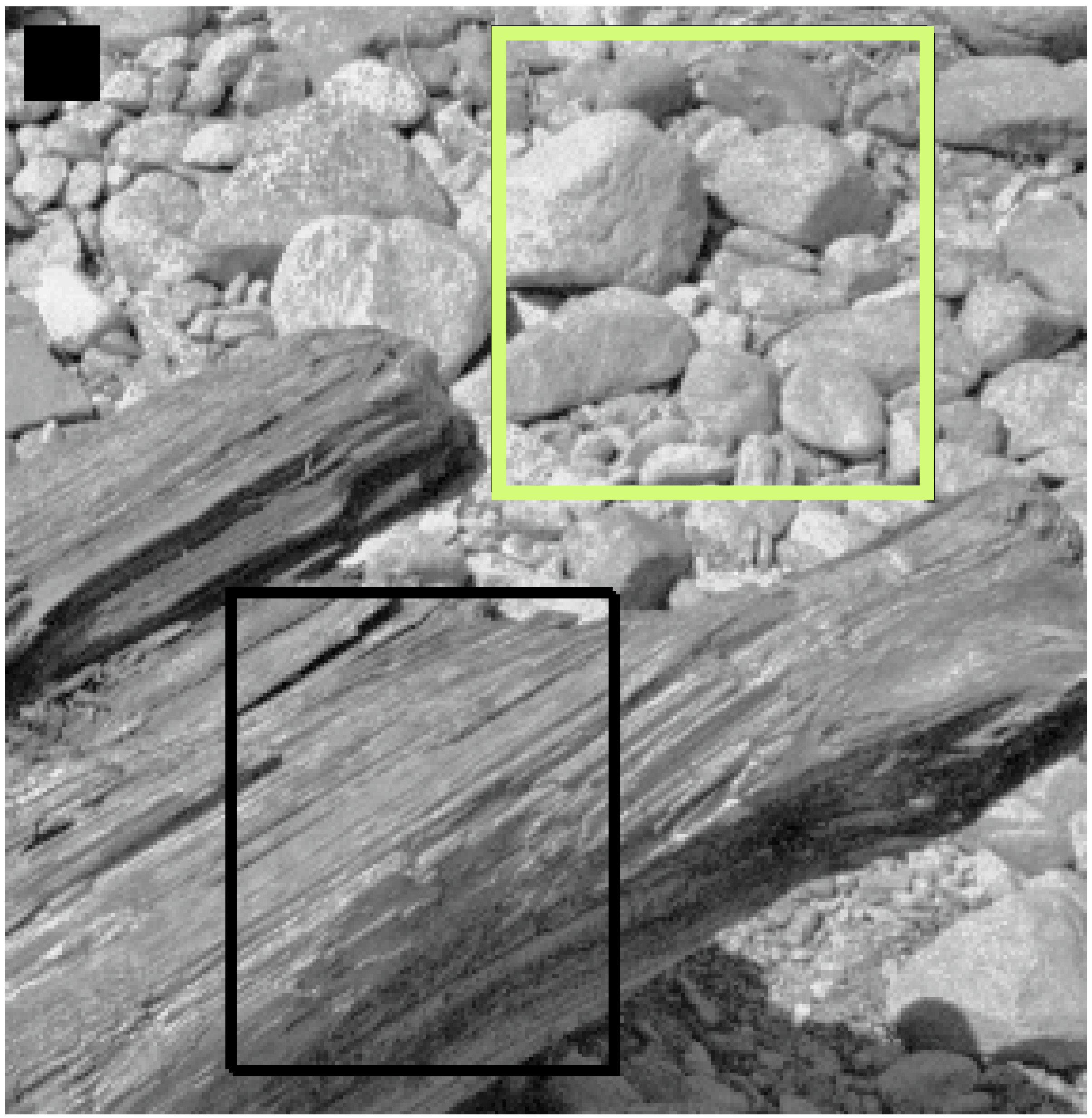
- can only have a small number of classes
- representations are not shared
- cannot learn *intrinsic* dimensions

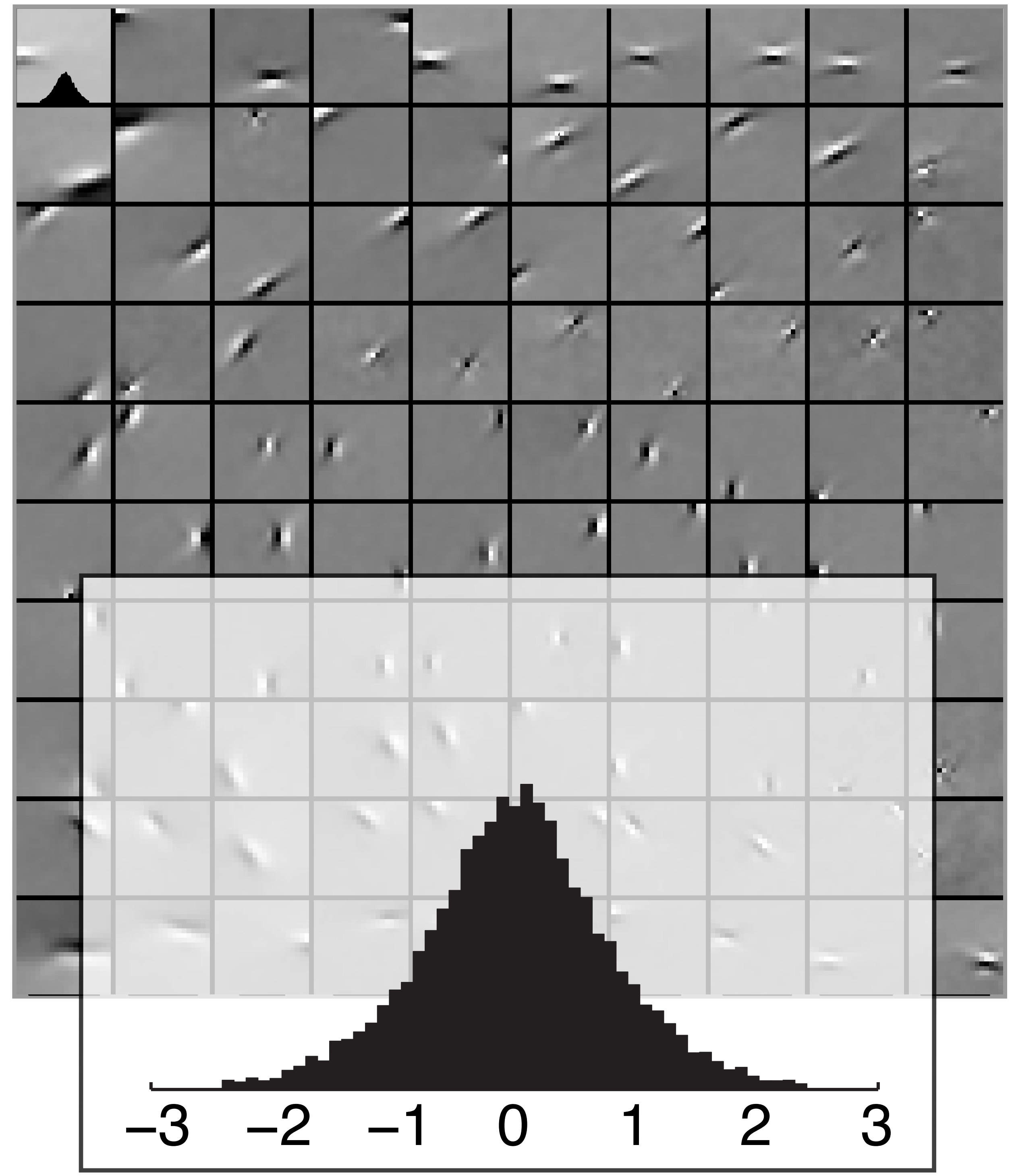
Characterizing different natural image densities



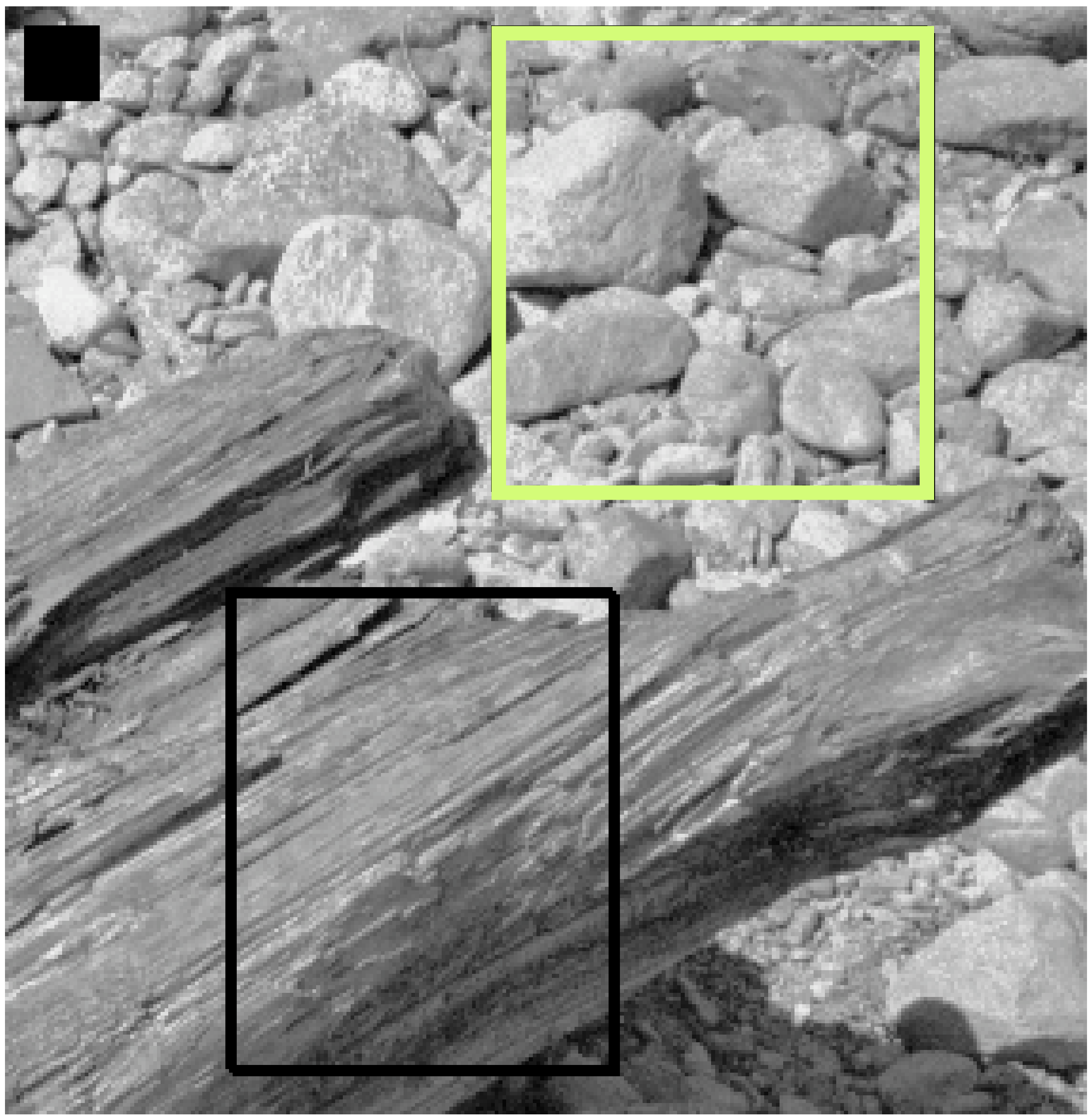


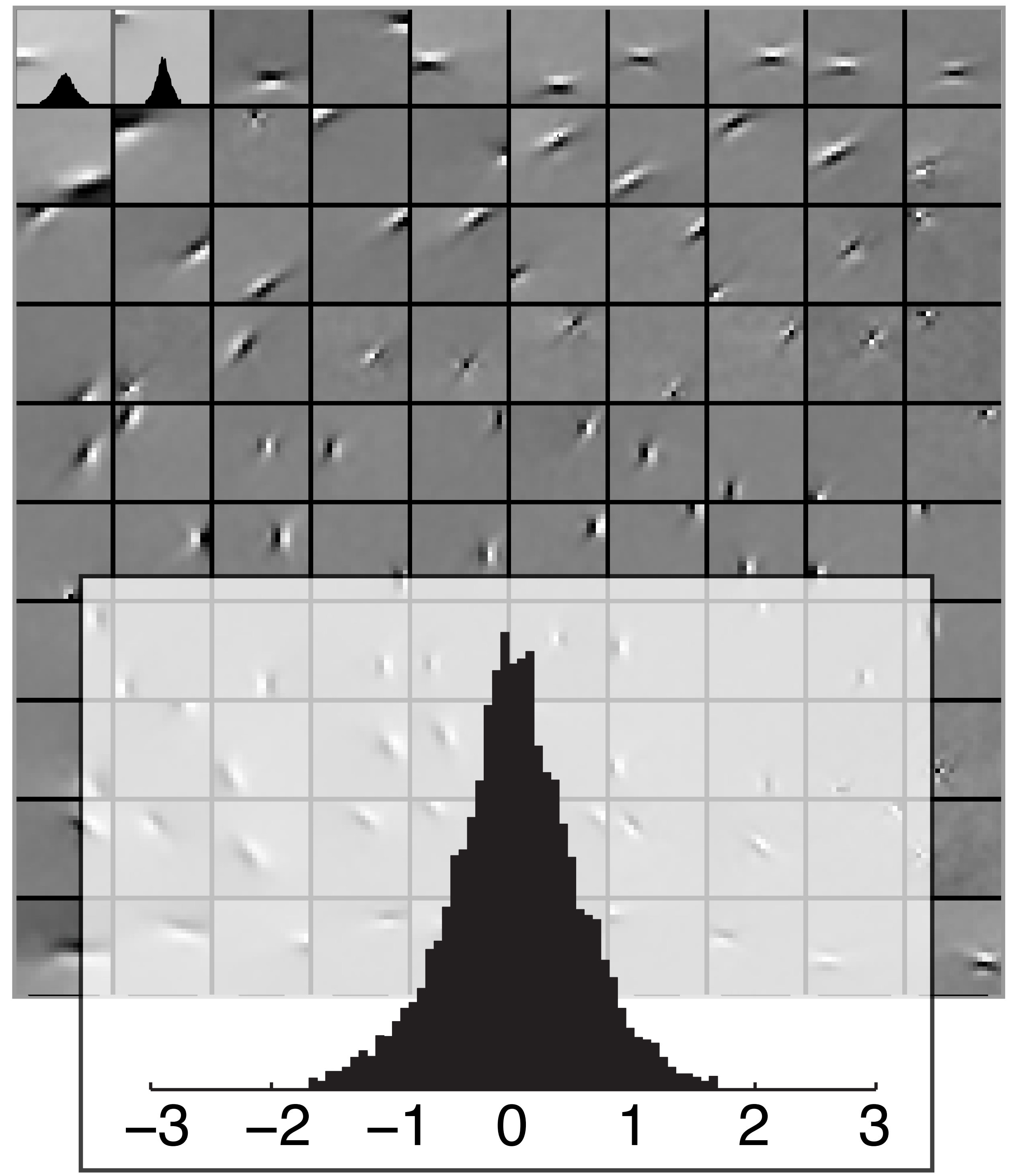
Characterizing different natural image densities



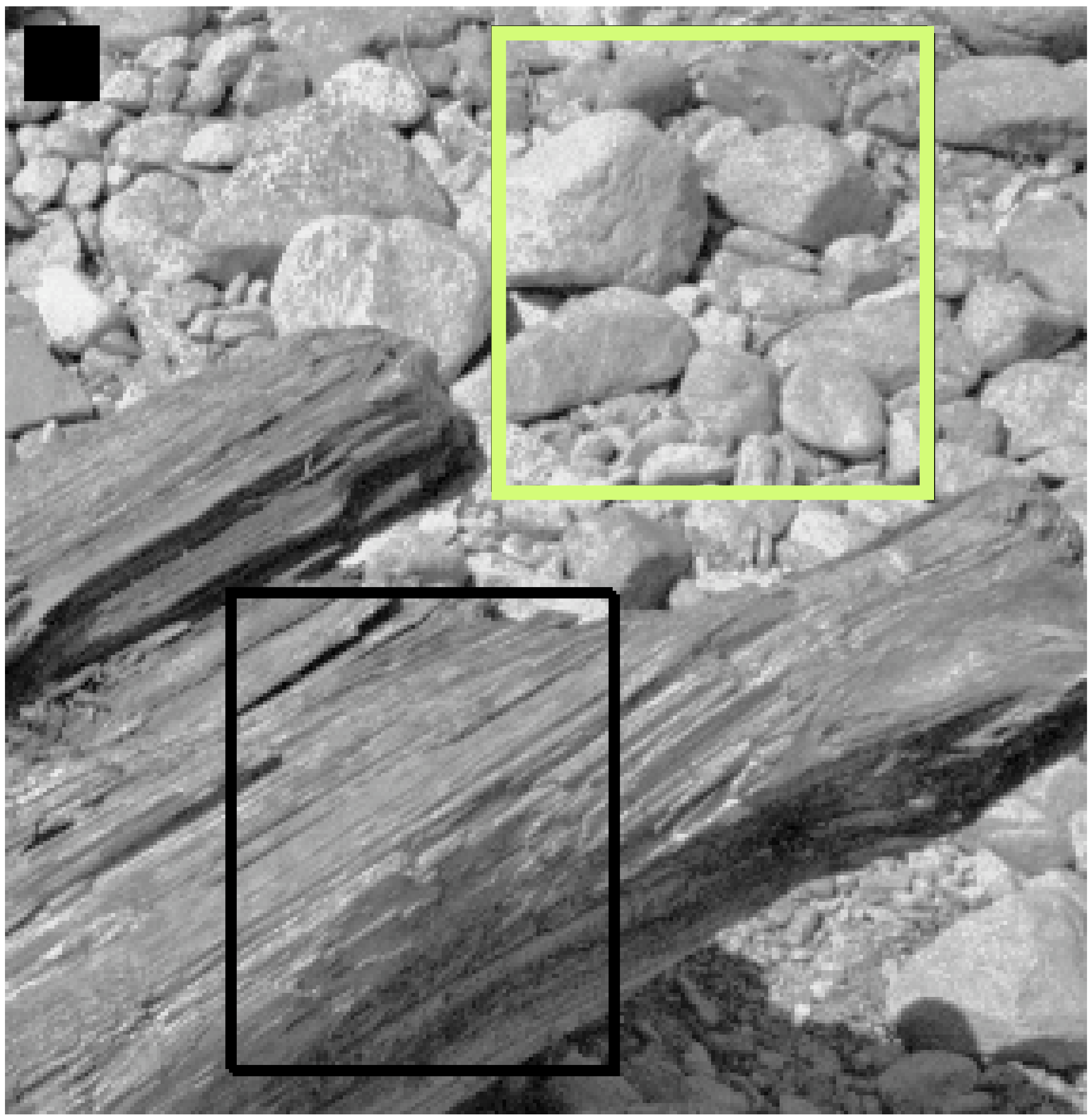


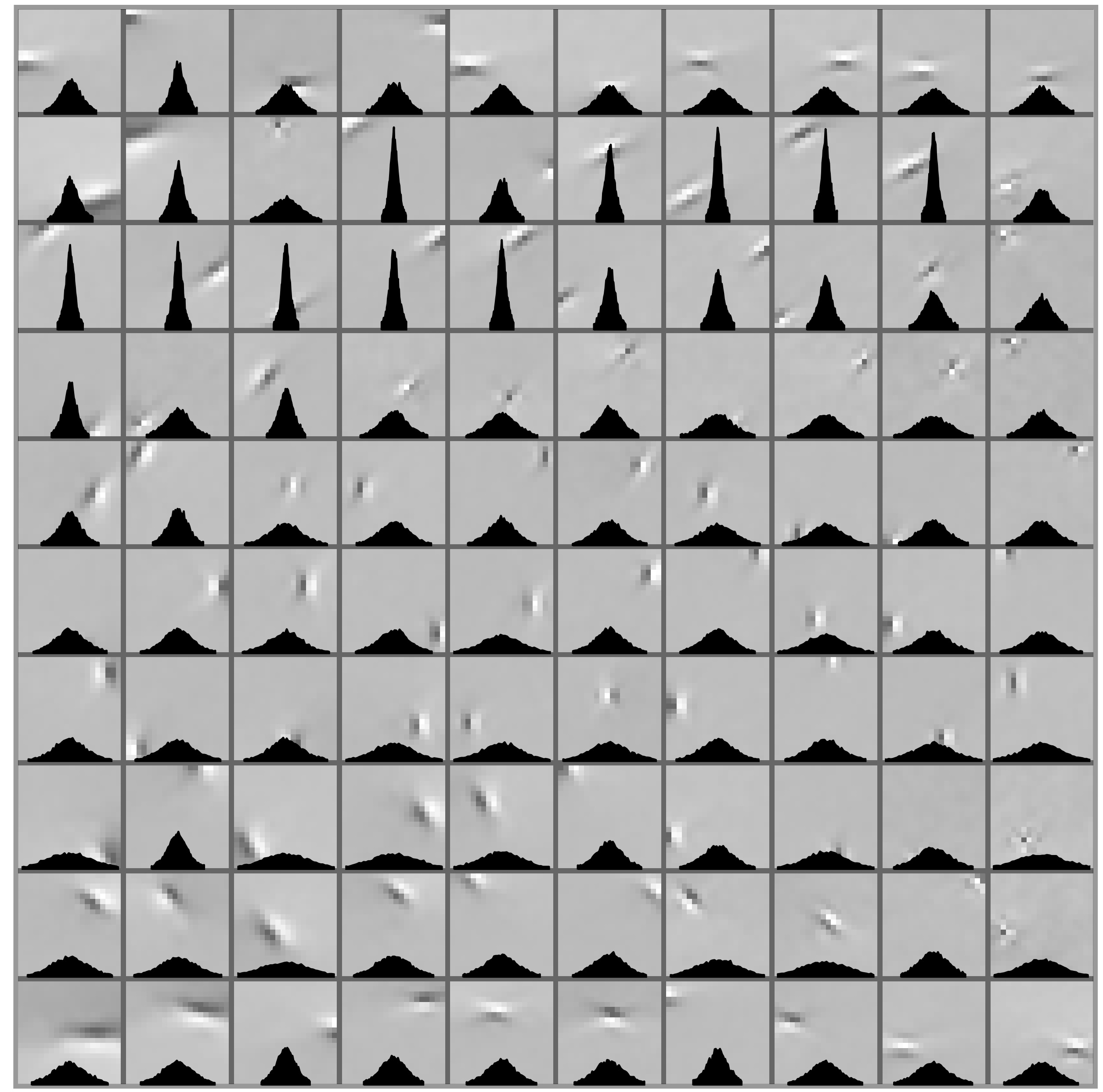
Characterizing different natural image densities



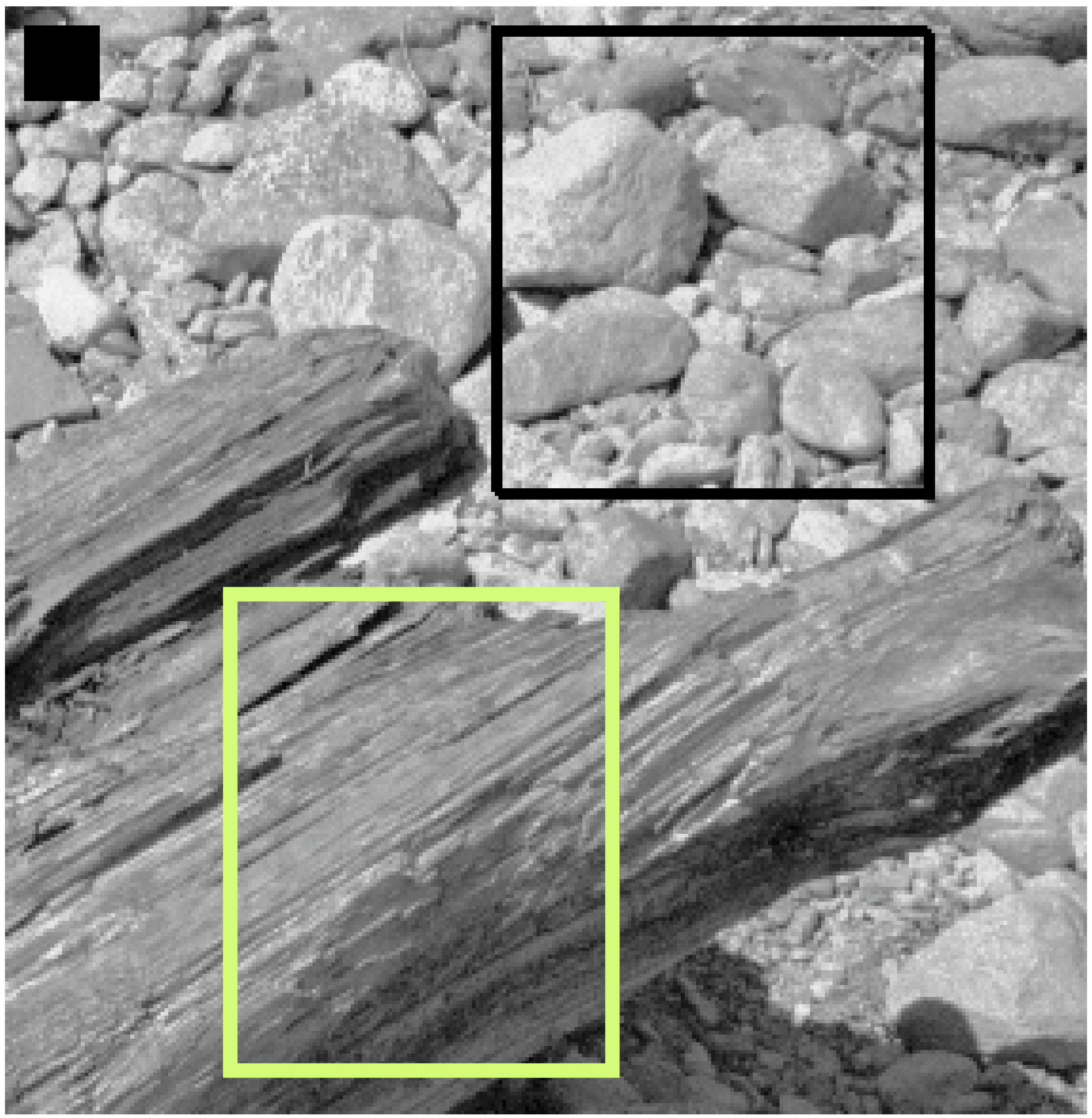


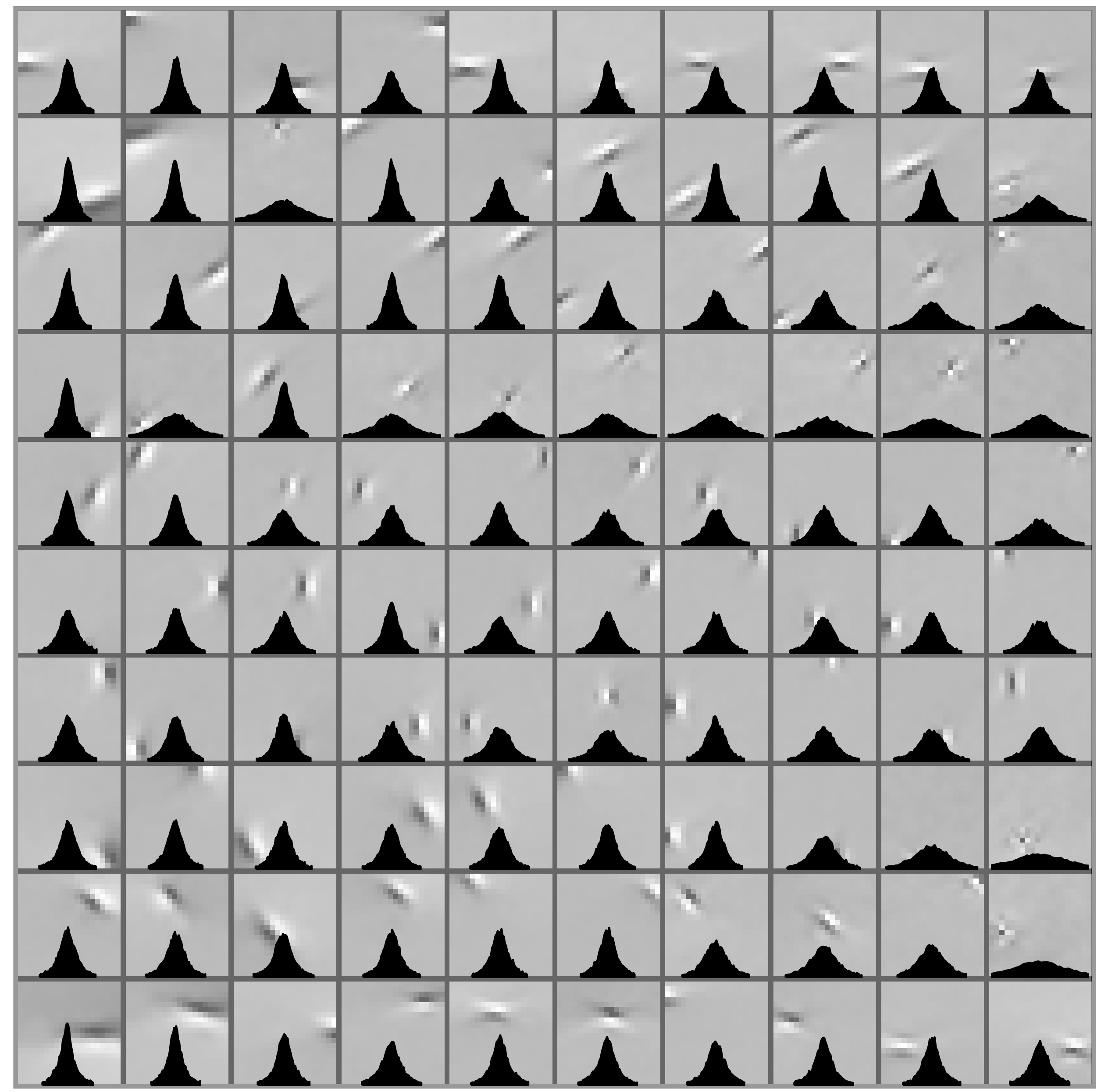
Characterizing different natural image densities





Characterizing different natural image *distributions*

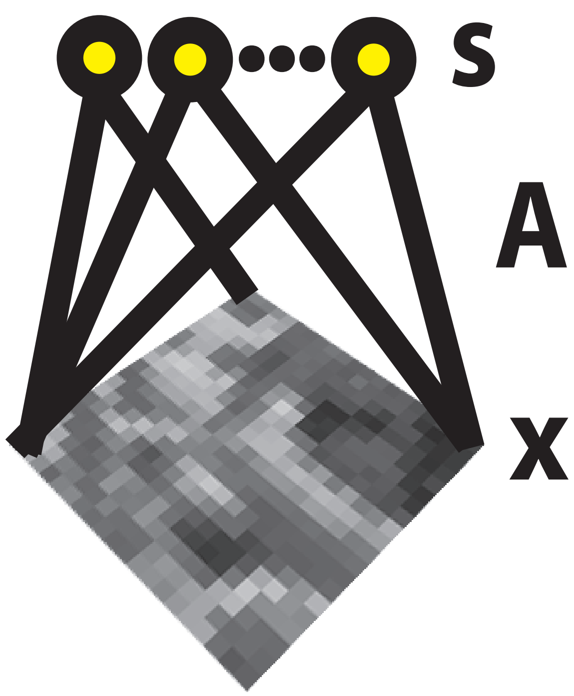


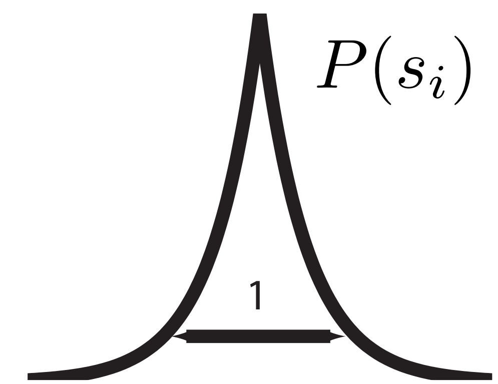


Generalizing the standard ICA model

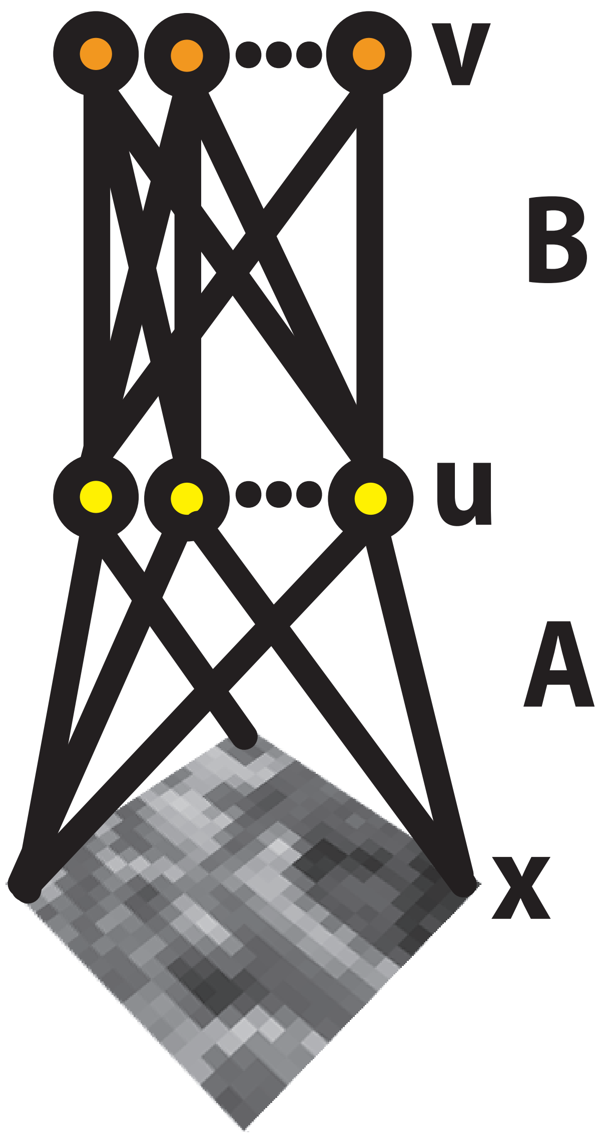
$$P(\mathbf{s}) = \prod_i P(s_i)$$

$$P(s_i) \propto \exp\left[-\left|\frac{s_i}{\lambda_i}\right|^{q_i}\right]$$





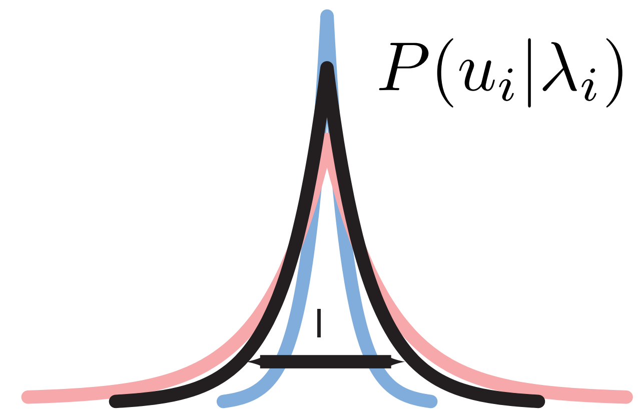
Generalizing the standard ICA model

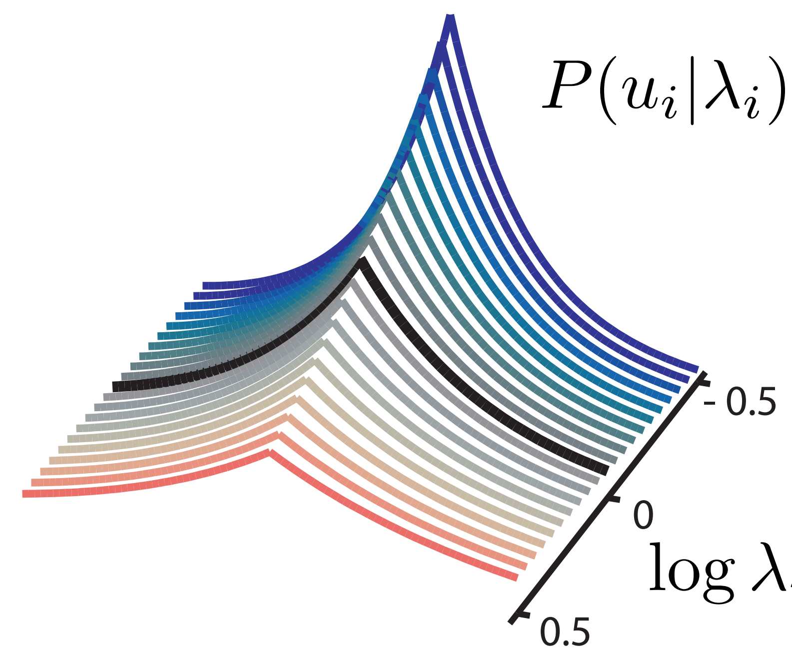


$$-\log p(\boldsymbol{u}|\mathbf{B},\boldsymbol{v}) \propto \sum_i \left|\frac{u_i}{\exp([\mathbf{B}\boldsymbol{v}]_i)}\right|^{q_i}$$

$$P(u_i|\lambda_i) \propto \exp\left[-\left|\frac{u_i}{\lambda_i}\right|^{q_i}\right]$$

$$\log \lambda_i = [\mathbf{B}\boldsymbol{v}]_i$$





Independent density components

$$-\log p(\boldsymbol{u}|\mathbf{B}, \boldsymbol{v}) \propto \sum_i \left| \frac{u_i}{\exp([\mathbf{B}\boldsymbol{v}]_i)} \right|^{q_i}$$

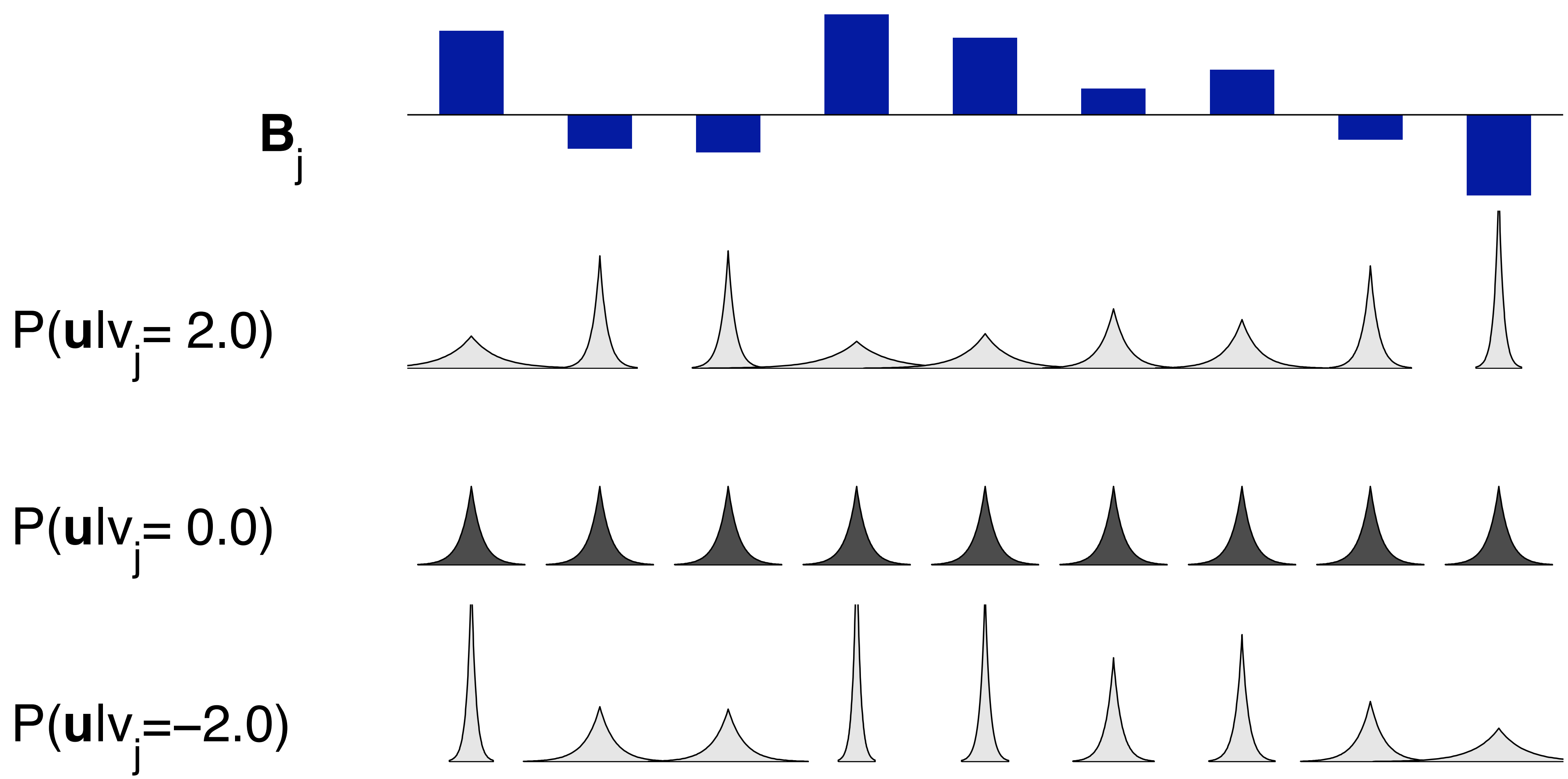
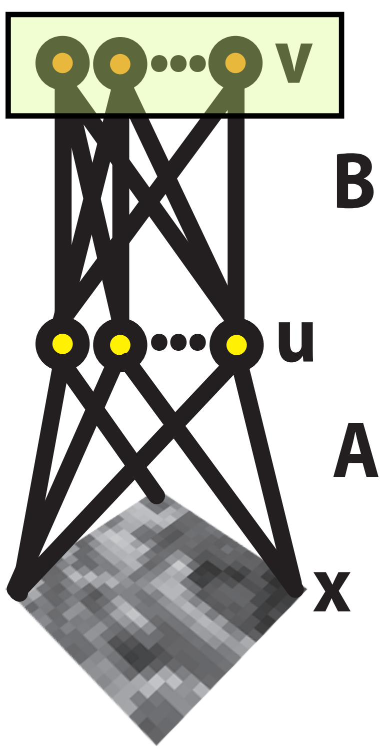


Illustration of inference in the model using synthesized data



$$P(v_i) \propto \exp(-|v_i|^{q_i})$$

$$\log \lambda_i = [\mathbf{B}\boldsymbol{v}]_i$$

$$P(u_i) \propto \exp\left[-\left|\frac{u_i}{\lambda_i}\right|^{q_i}\right]$$

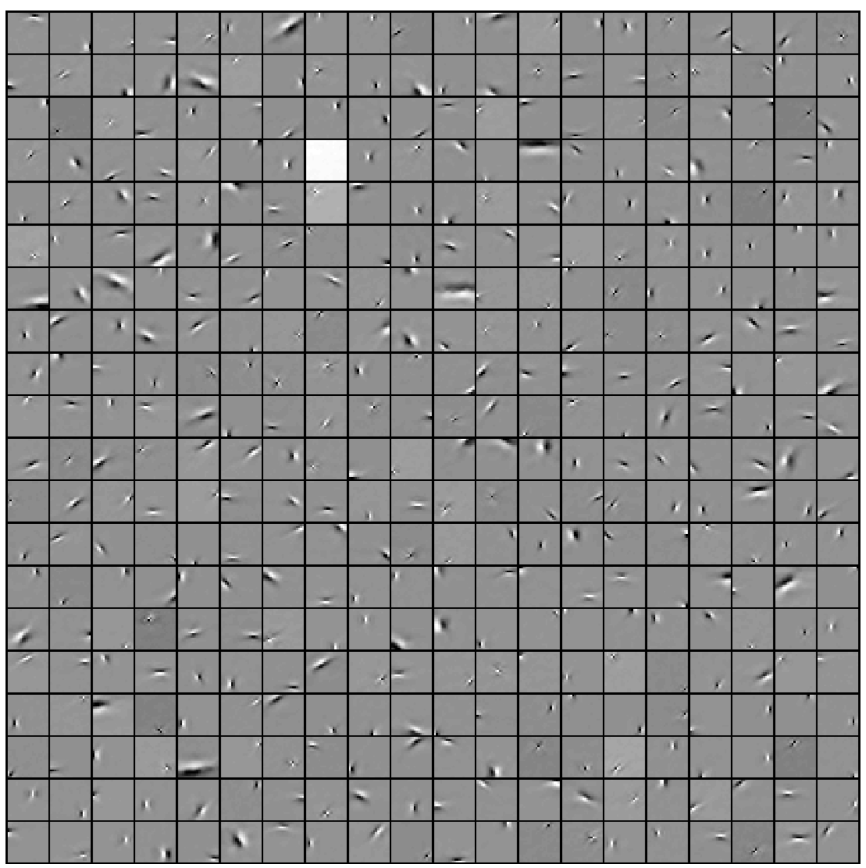
$$\hat{\boldsymbol{v}} = \arg\max_{v} P(\boldsymbol{v}|\mathbf{B}, \boldsymbol{u})$$

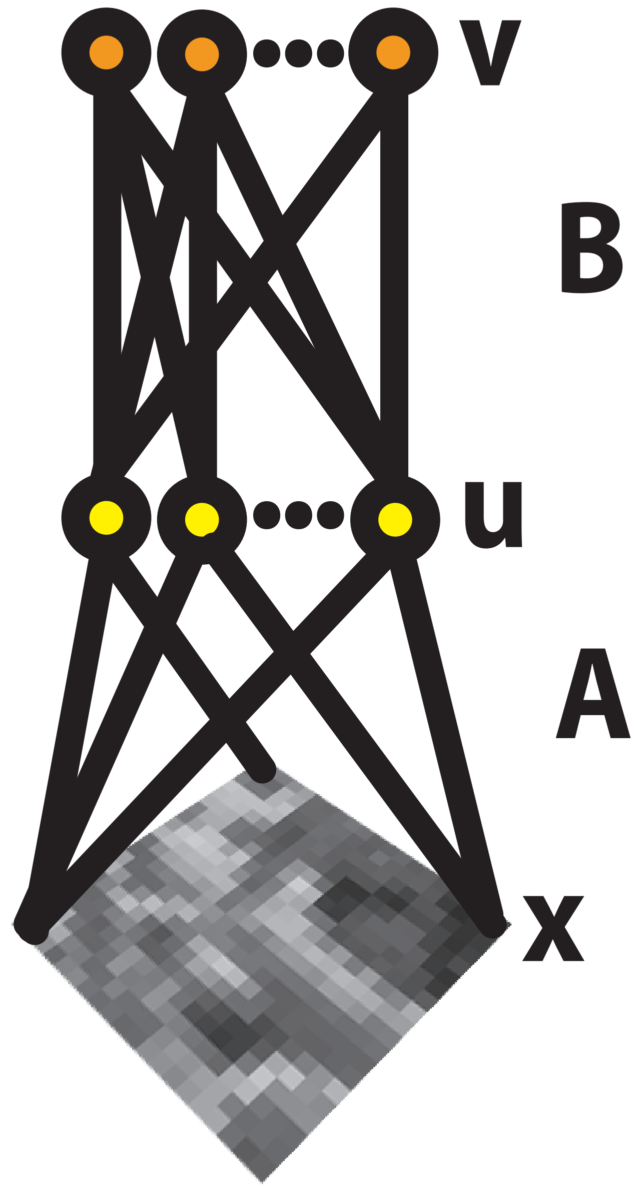
Synthesis → Inference →

$\mathbf{v} \sim p(\mathbf{v})$ → $\boldsymbol{\lambda} = c\exp(\mathbf{Bv})$ → $\mathbf{u} \sim p(\mathbf{u}|\boldsymbol{\lambda})$ → $\hat{\boldsymbol{\lambda}}$ → $\hat{\mathbf{v}}$

Learning higher-order structure of natural images

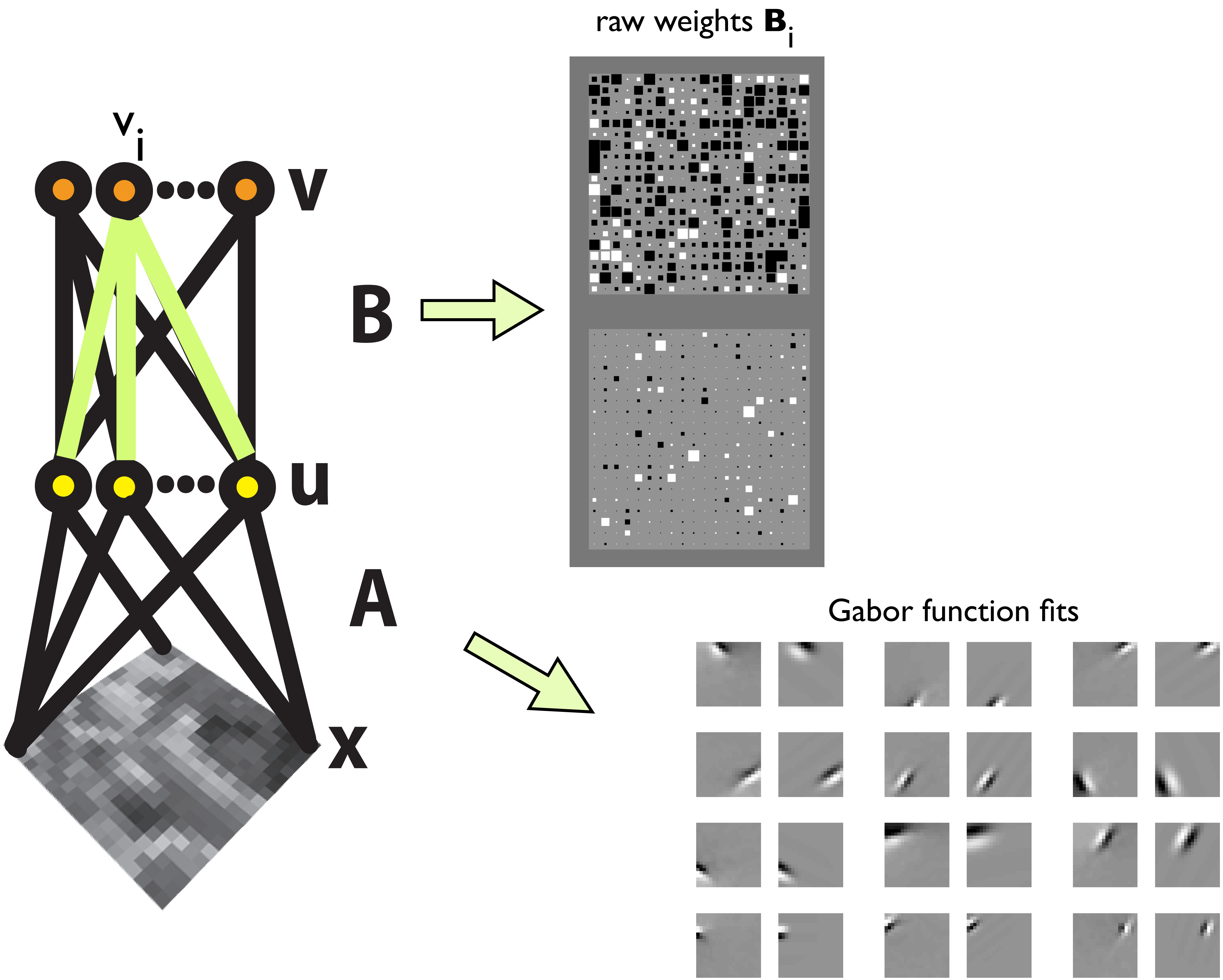
- **A** is learned with ICA
- **B** is learned by maximizing the posterior distribution



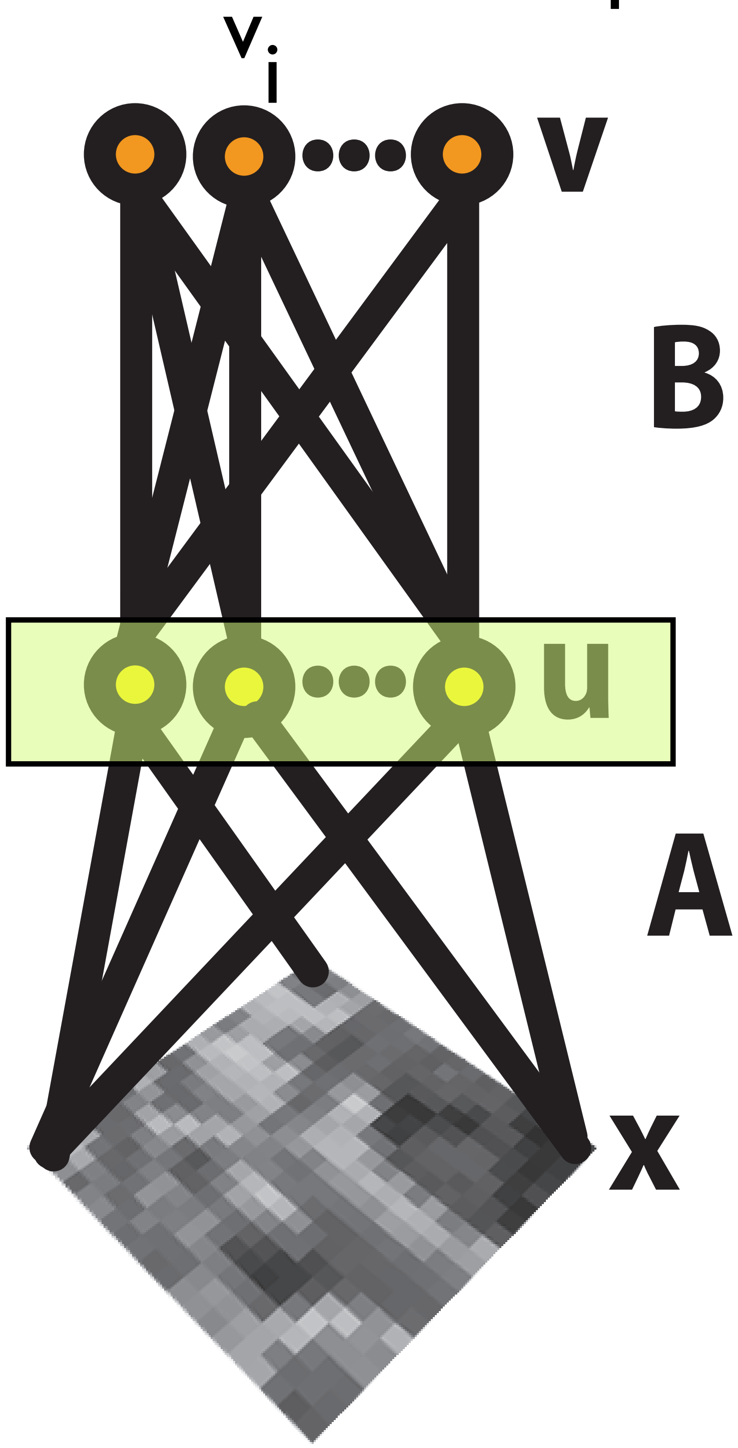


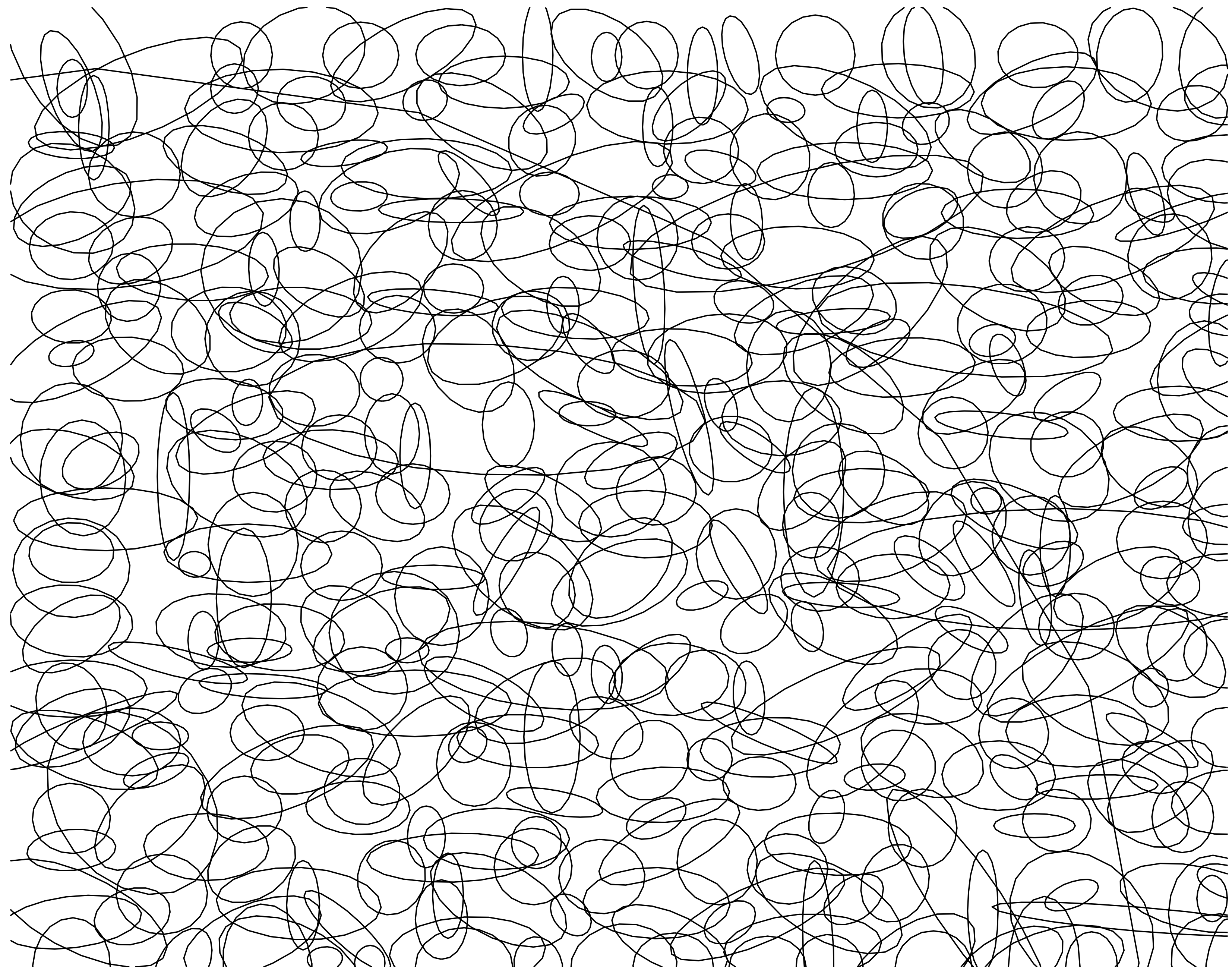
Train on natural images

Interpreting higher-order density components

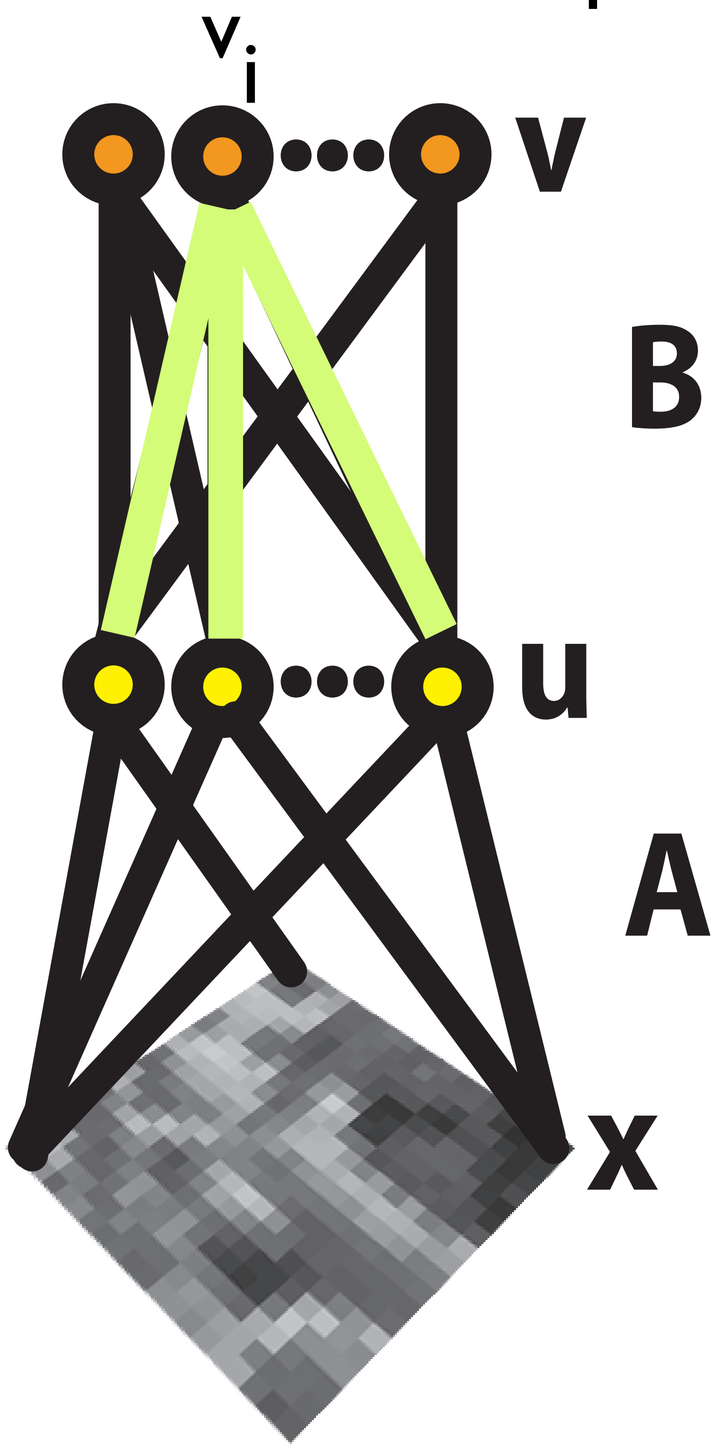


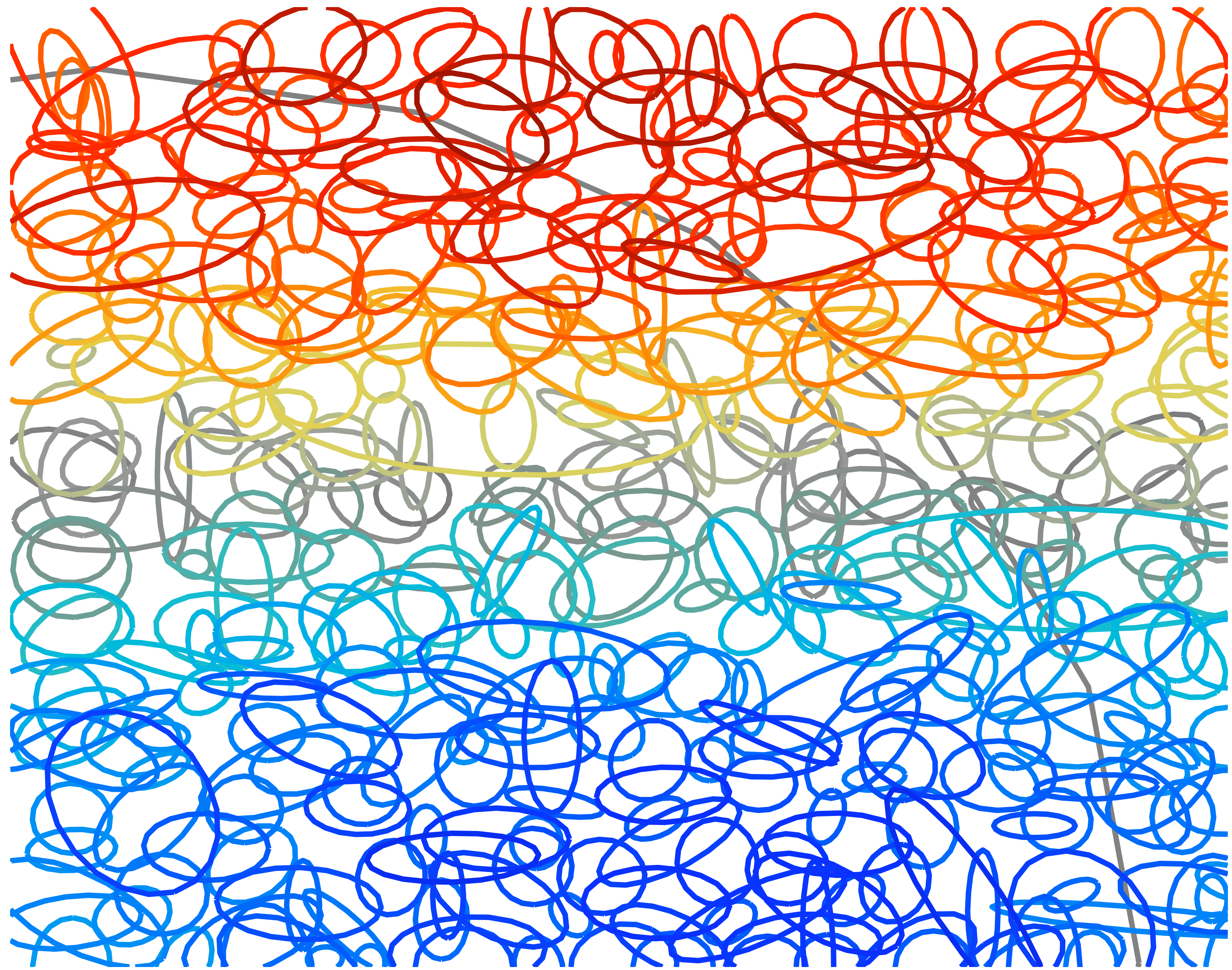
Interpreting higher-order density components

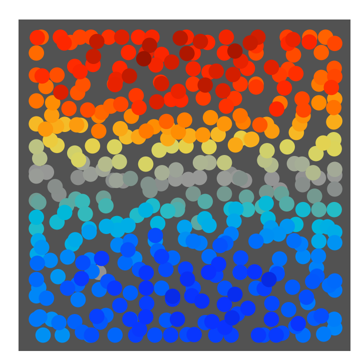




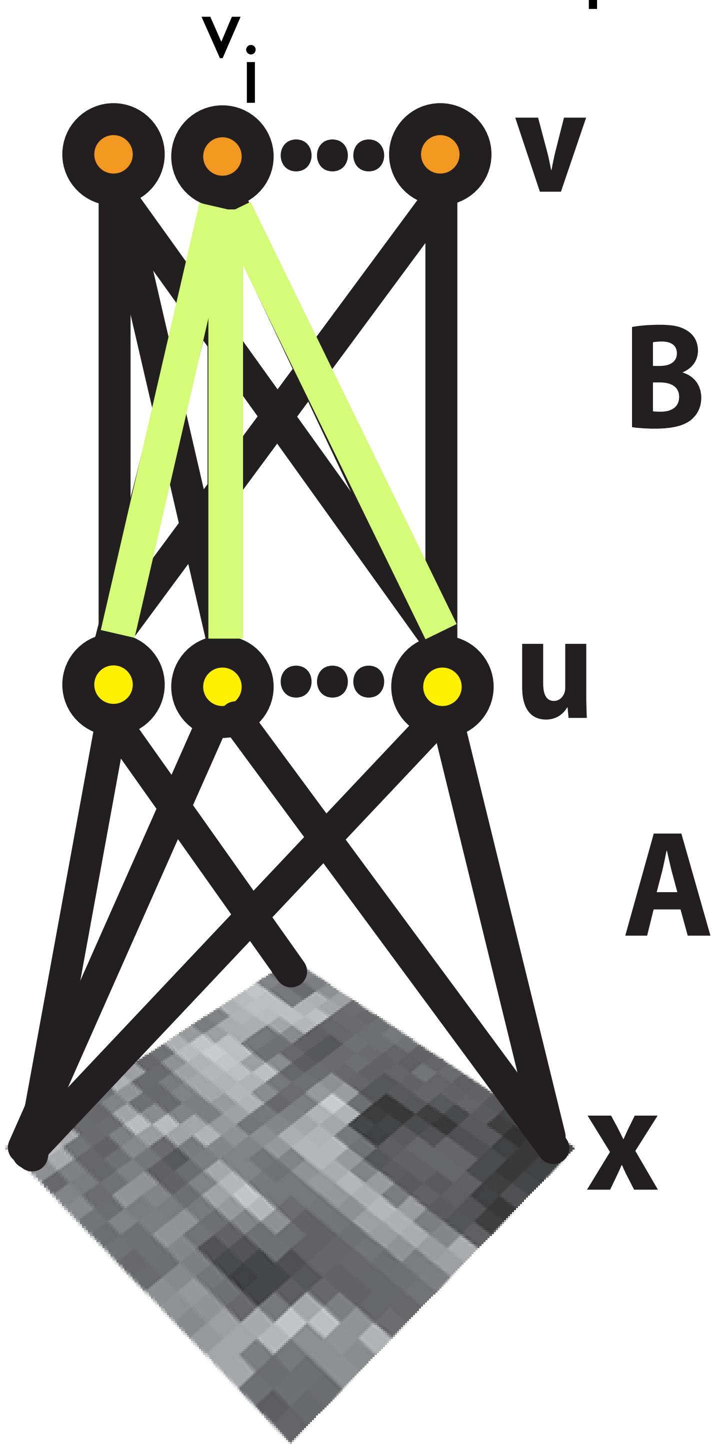
Interpreting higher-order density components

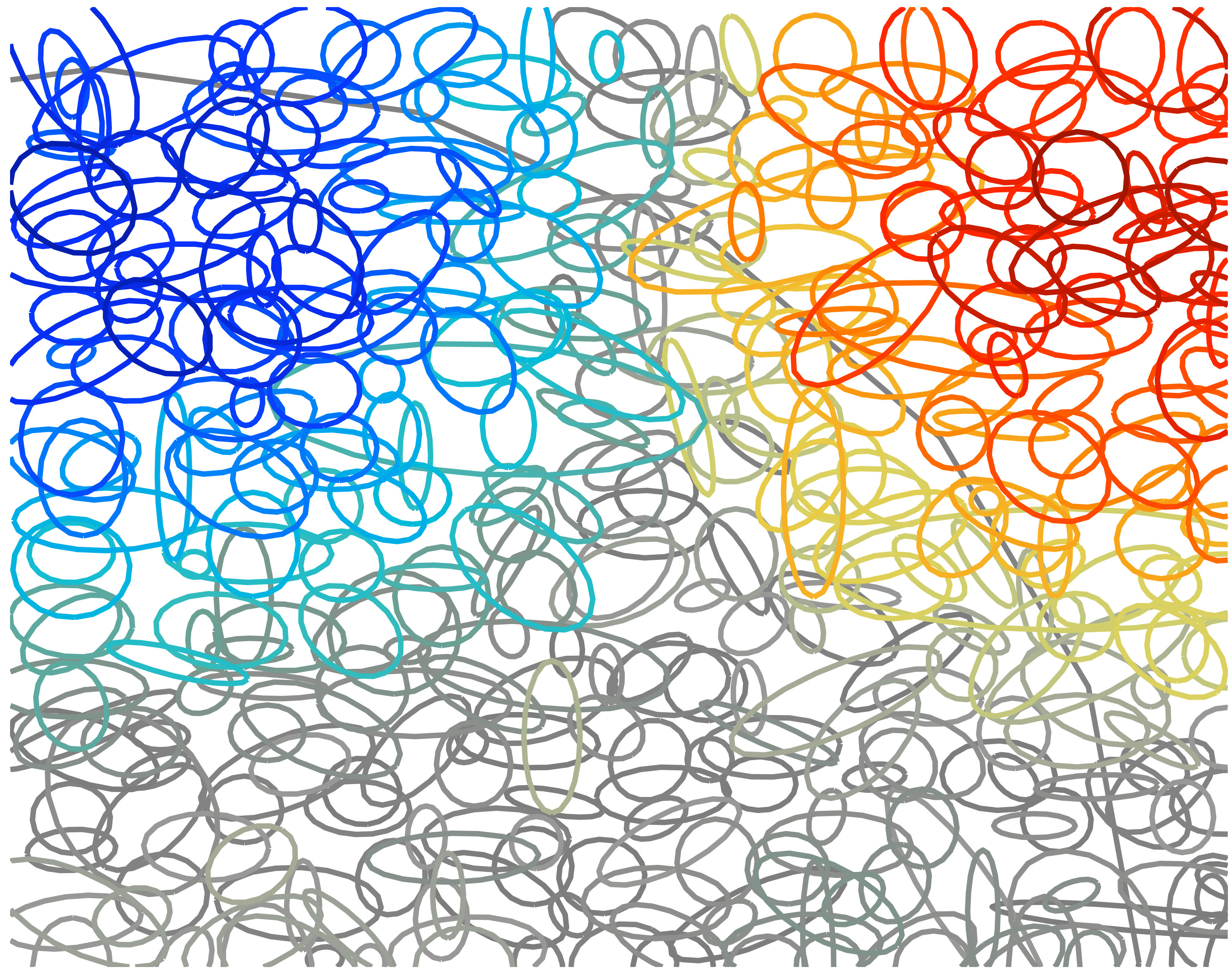


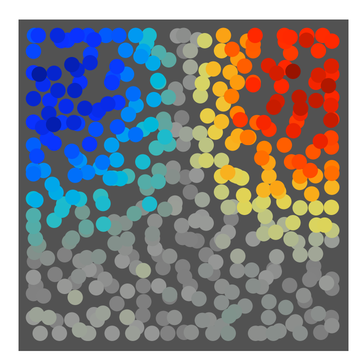




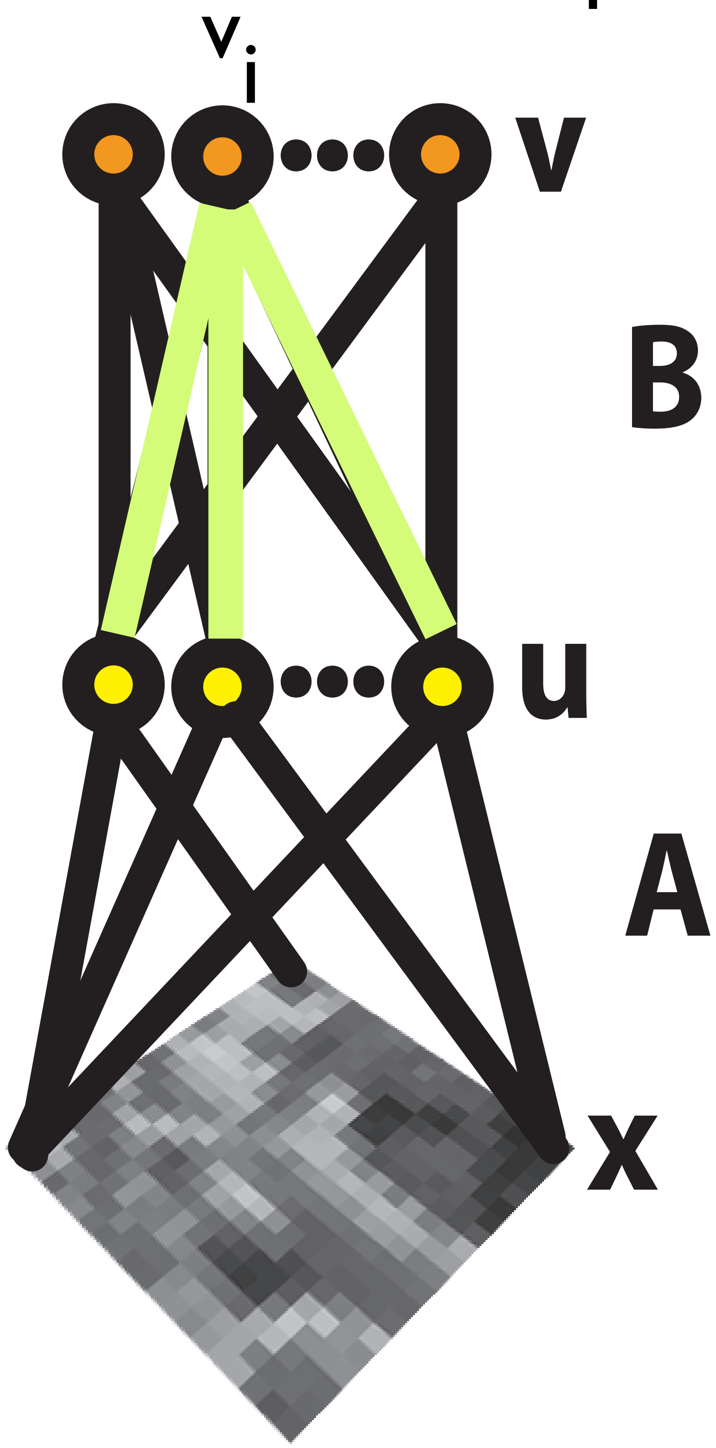
Interpreting higher-order density components

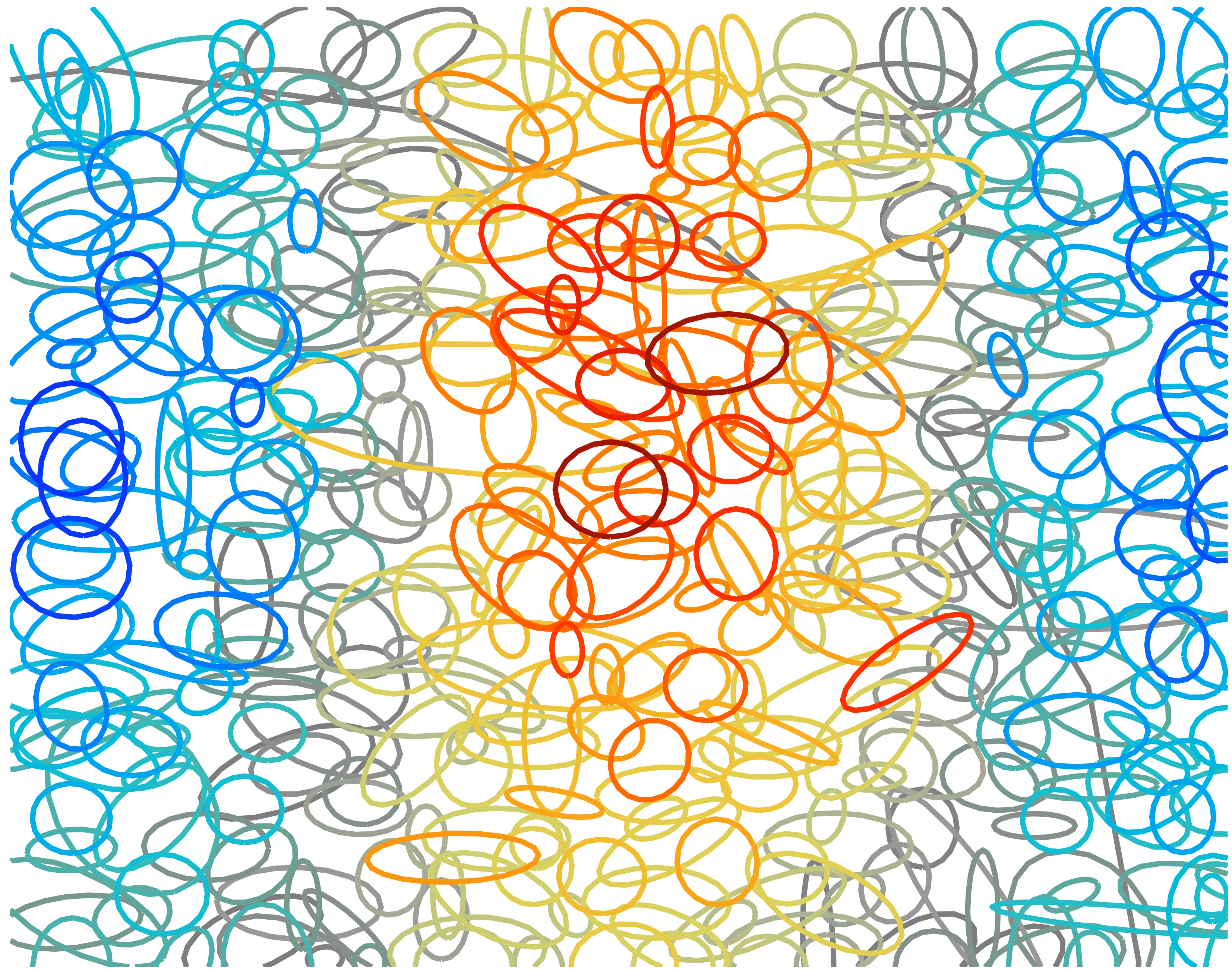


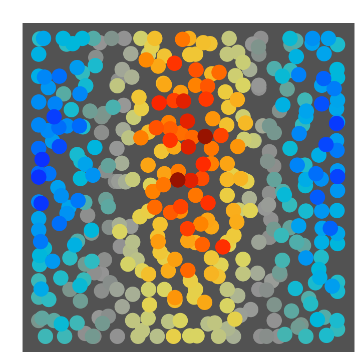




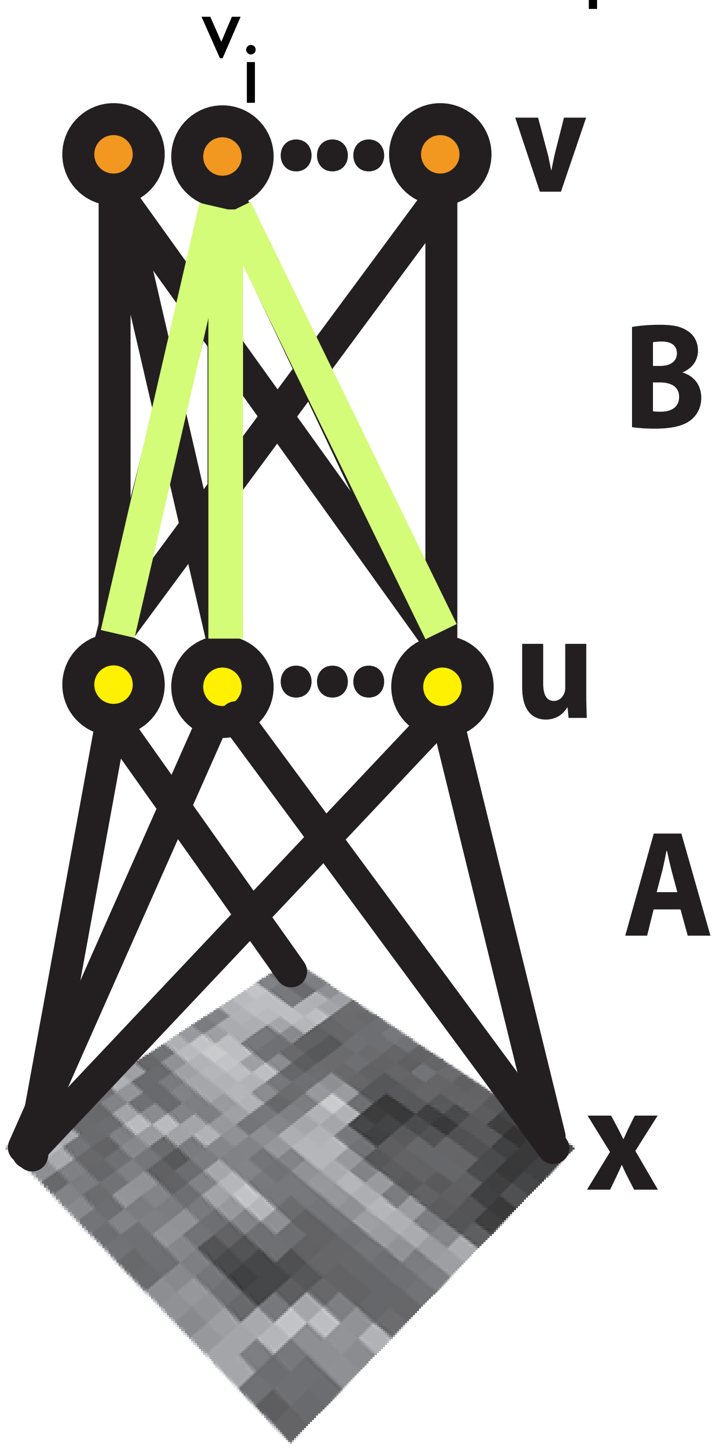
Interpreting higher-order density components

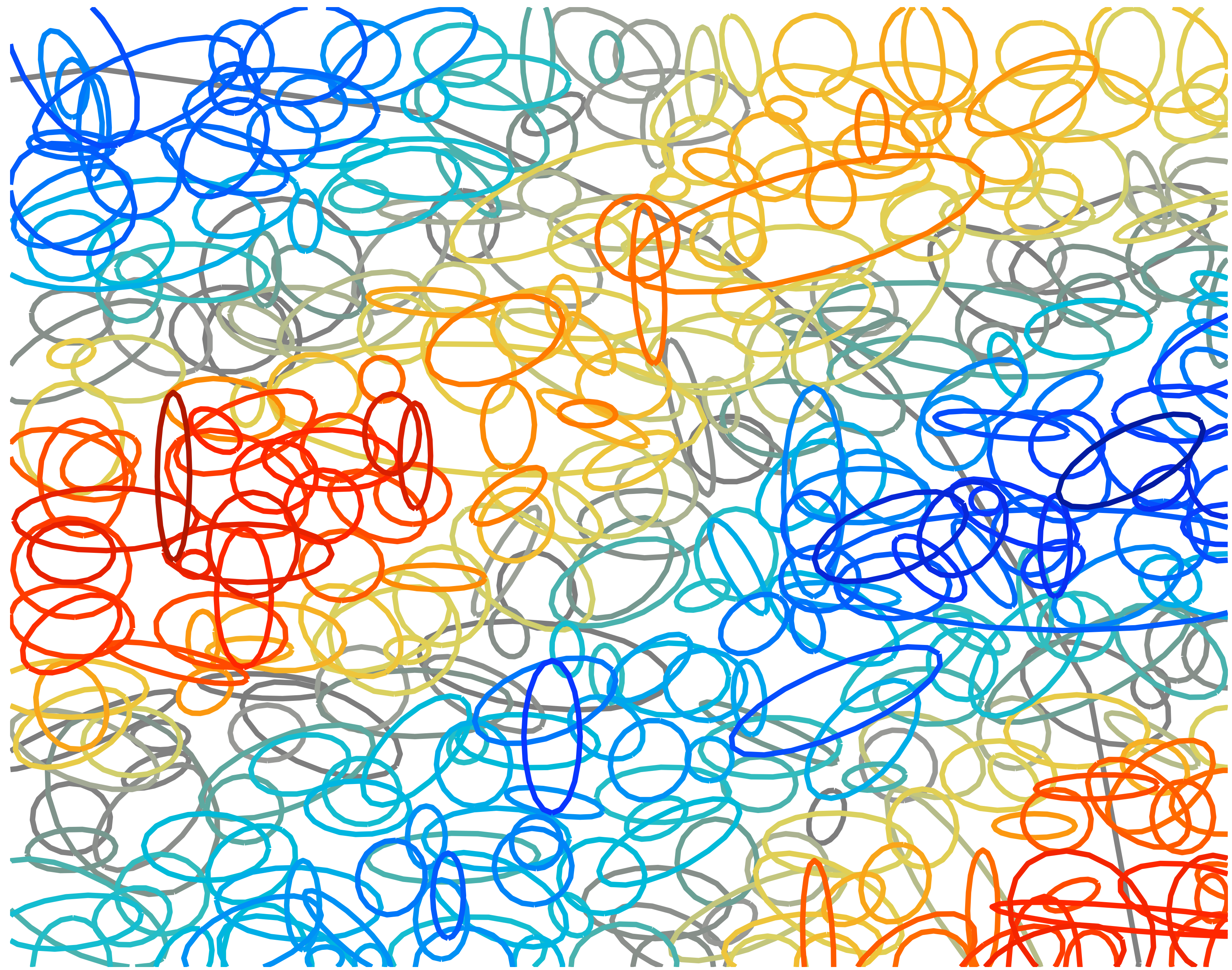


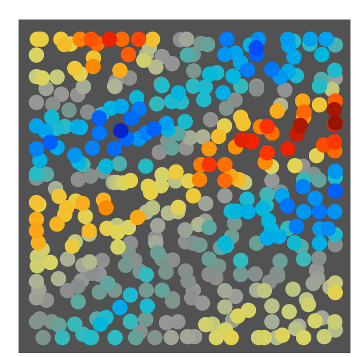




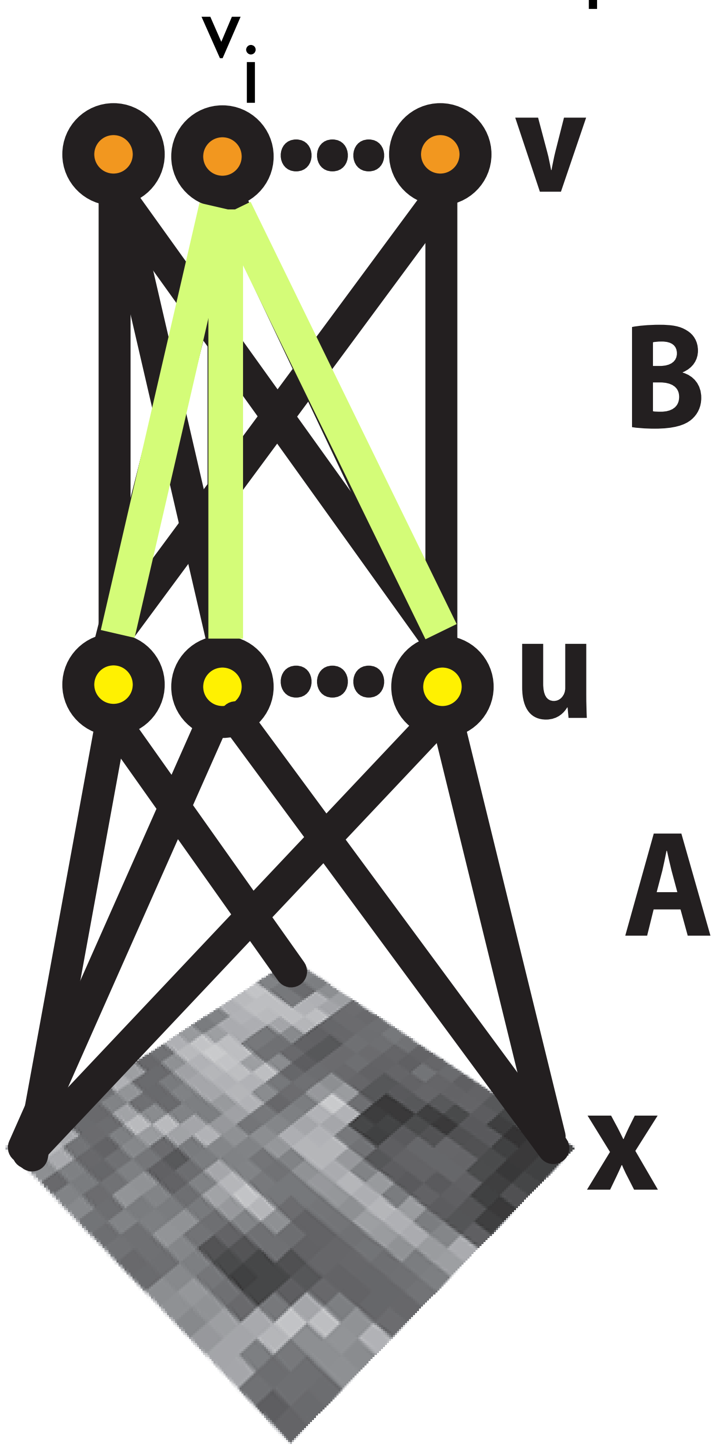
Interpreting higher-order density components

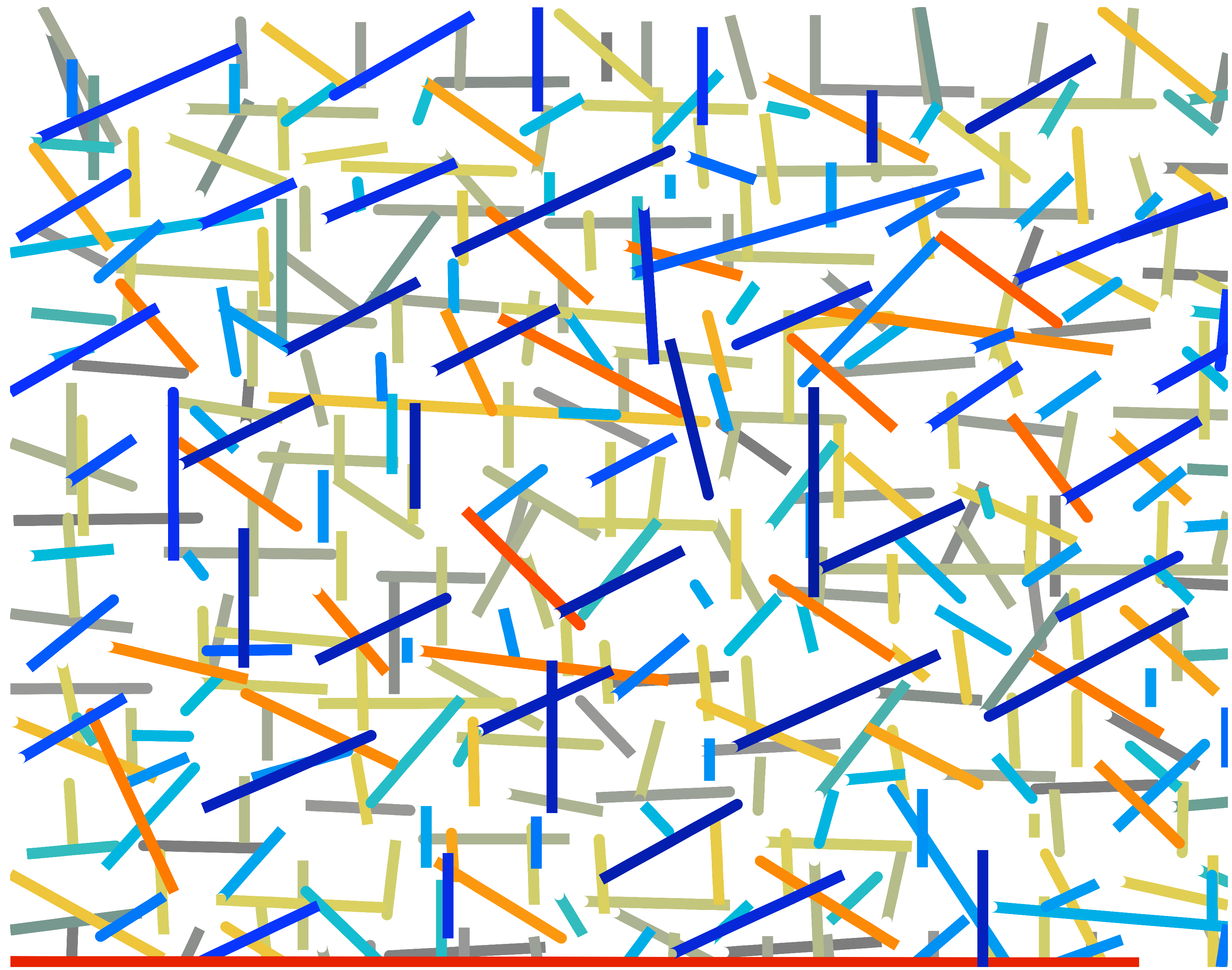


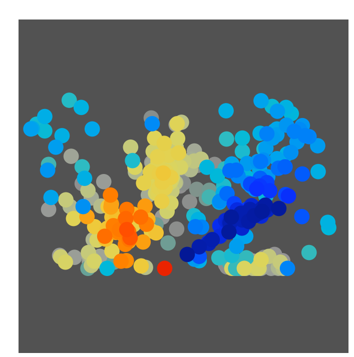




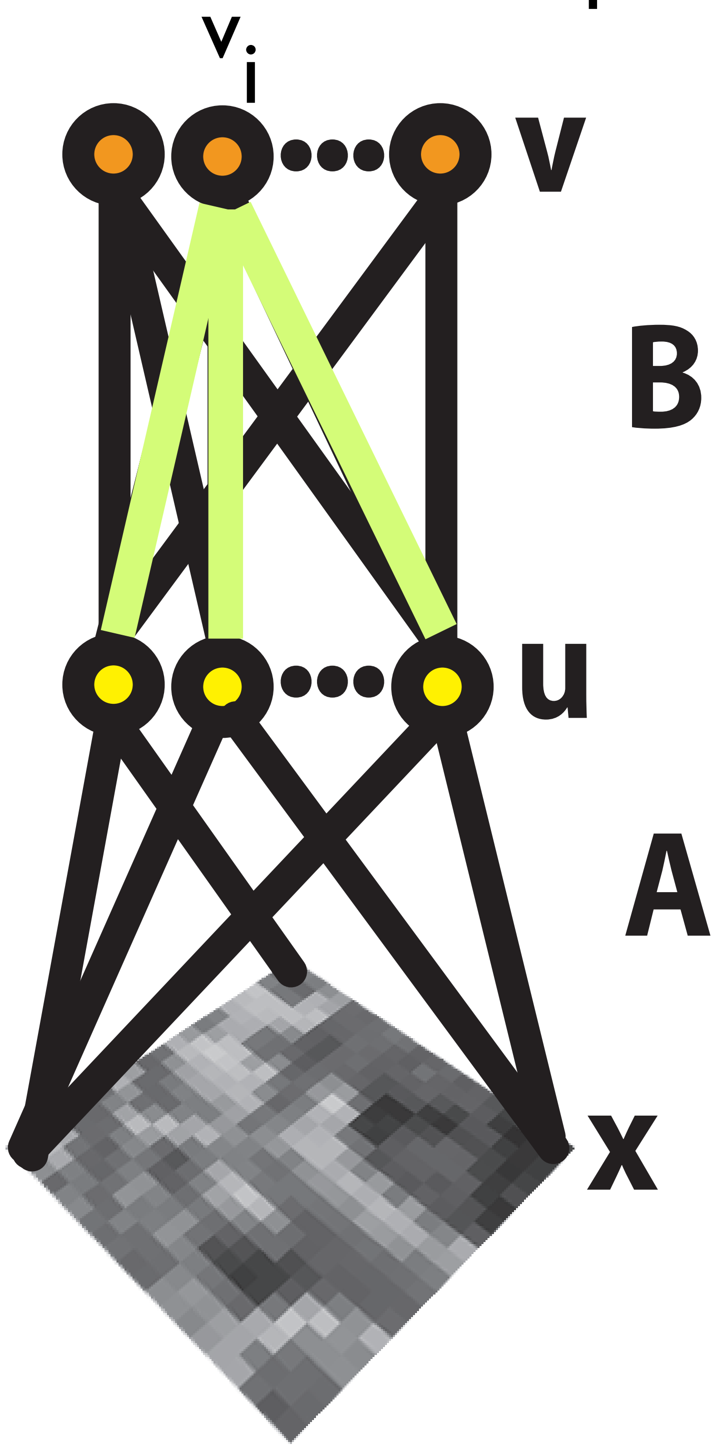
Interpreting higher-order density components

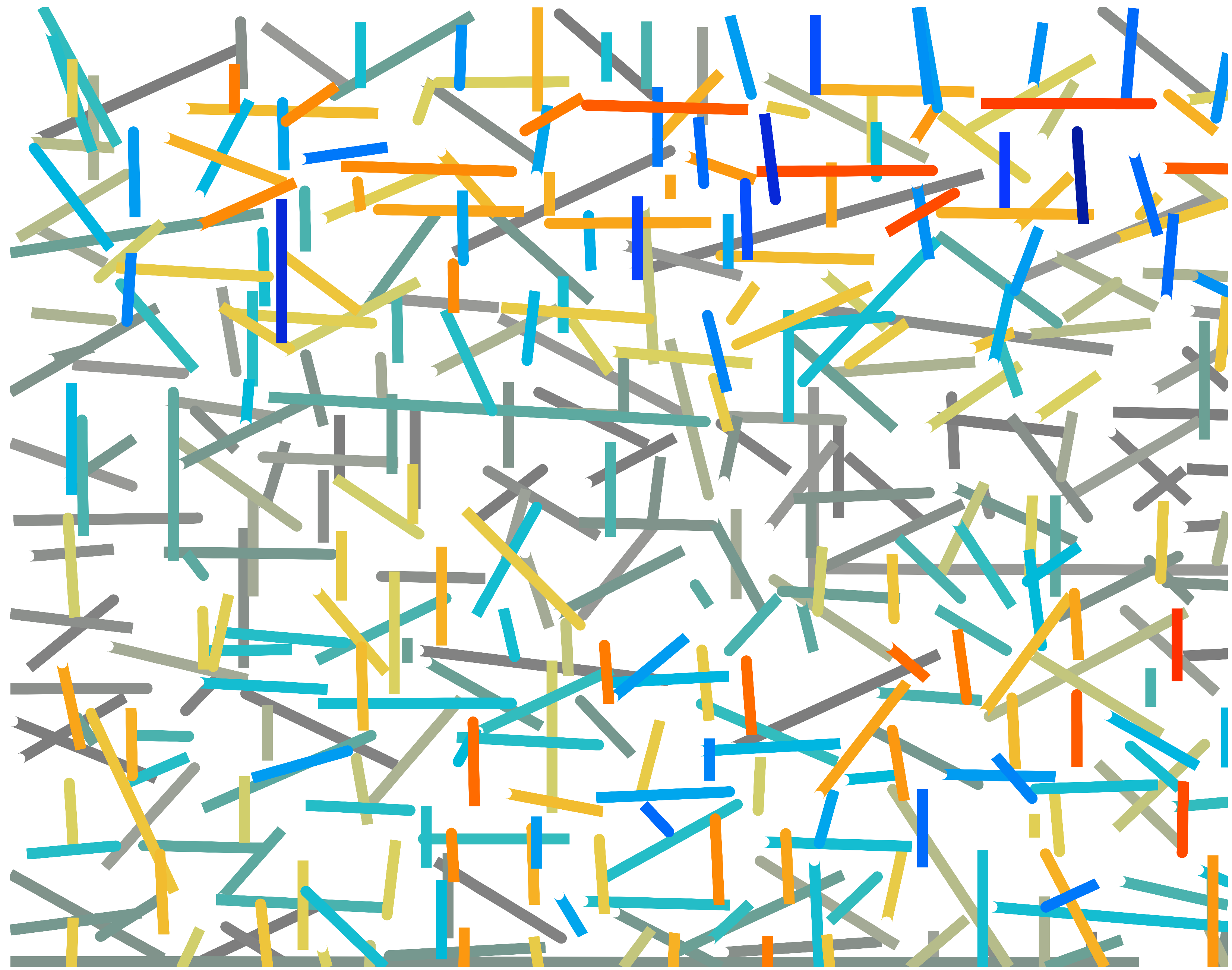


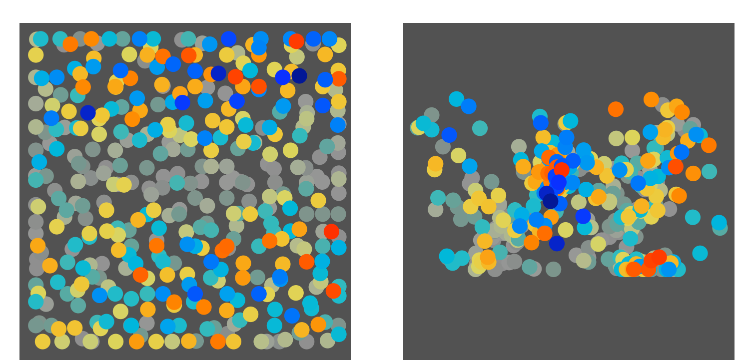




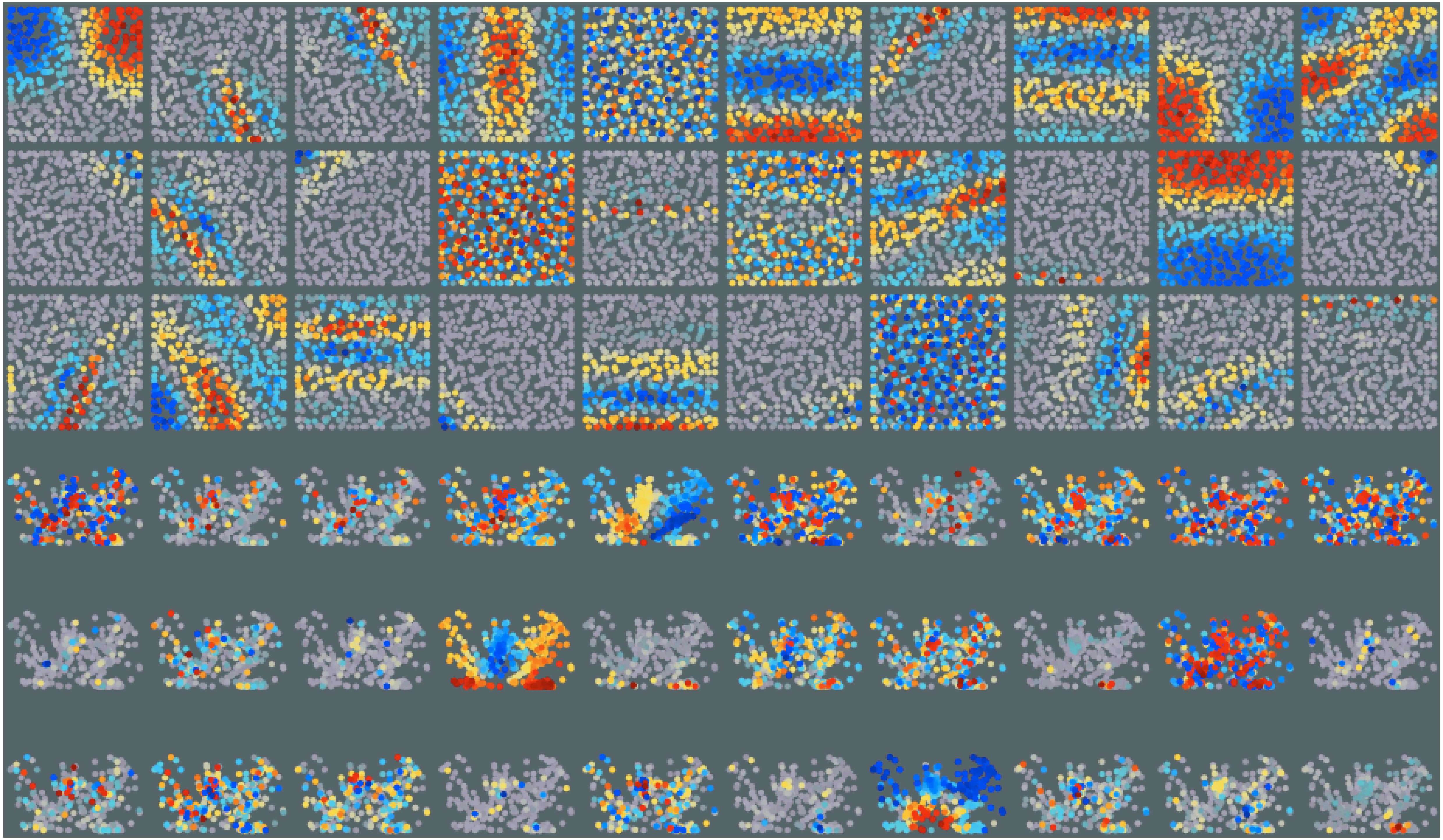
Interpreting higher-order density components





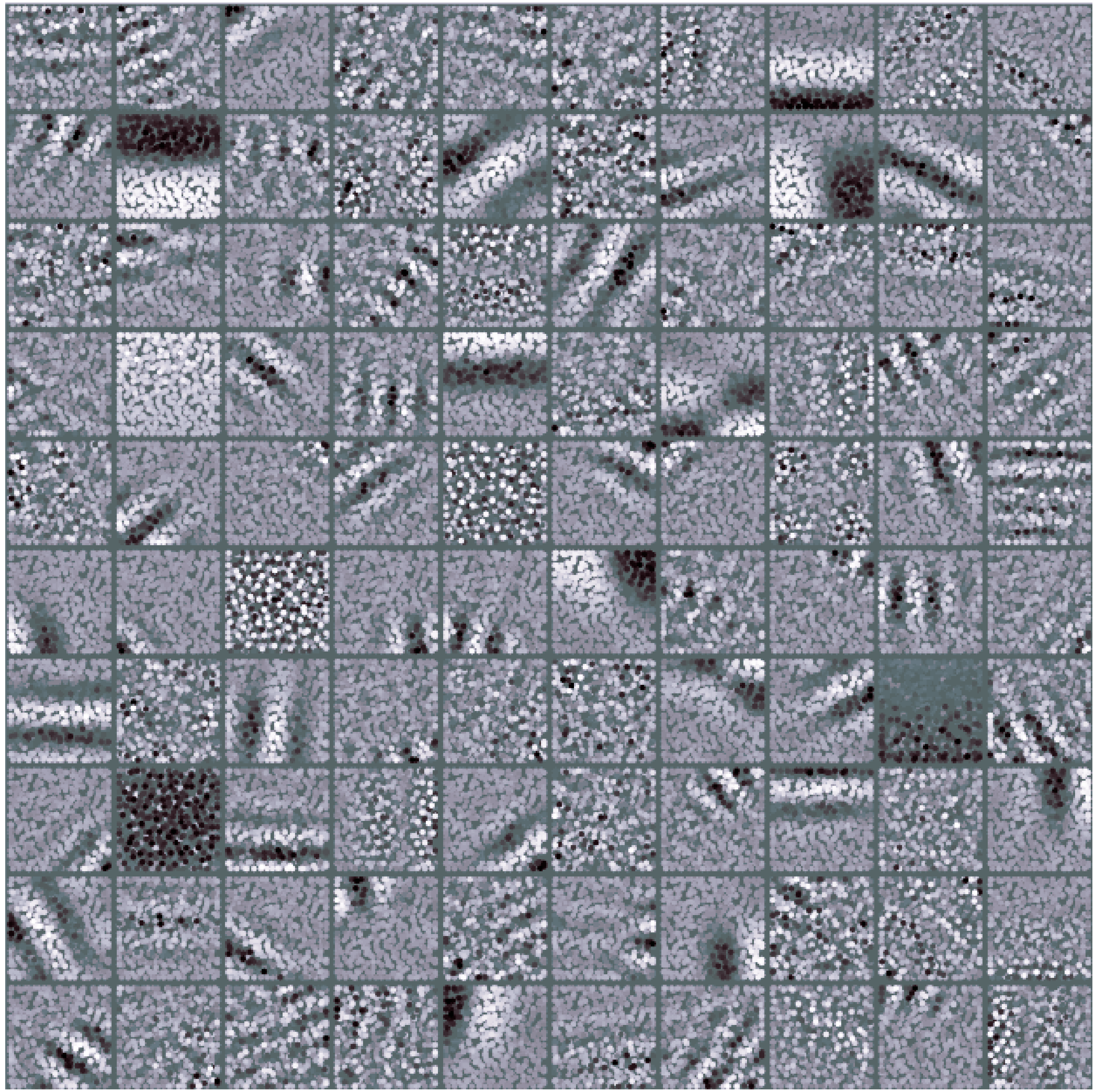


Learned density components of natural images



(30 out of 100 shown)

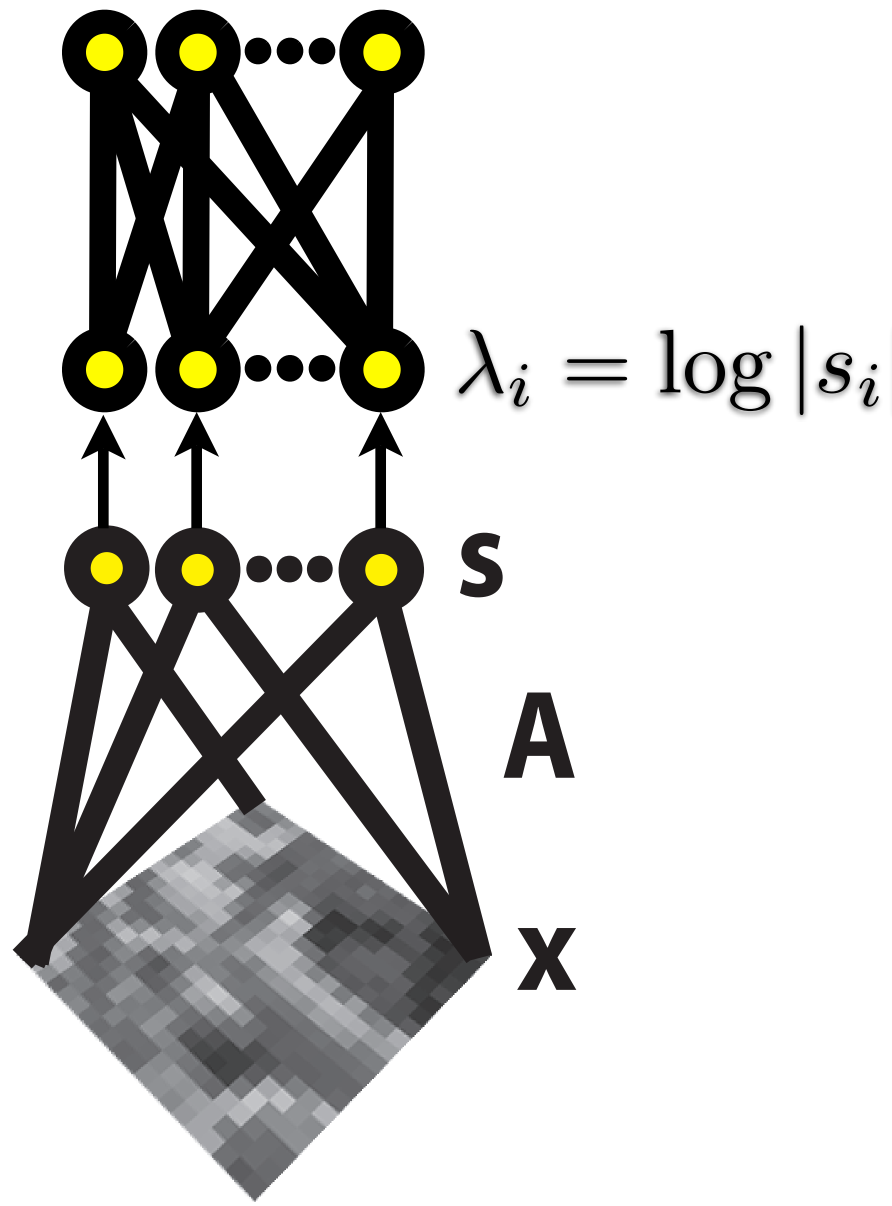
Full set of natural image density components



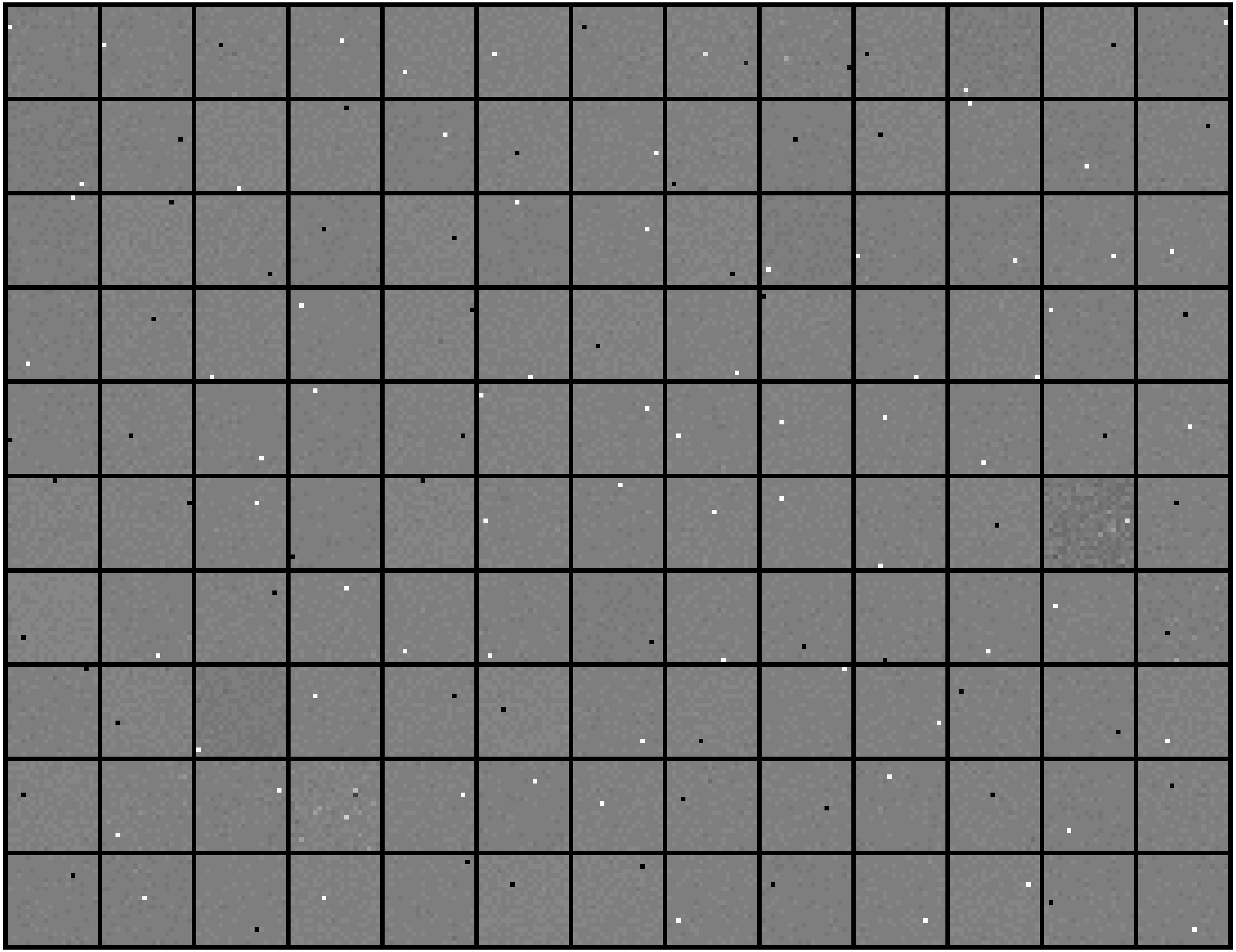
What about a feed-forward non-linearity?

Isn't this the same?

3. Do ICA again on output:
2. Add non-linearity:
1. Take standard ICA model:

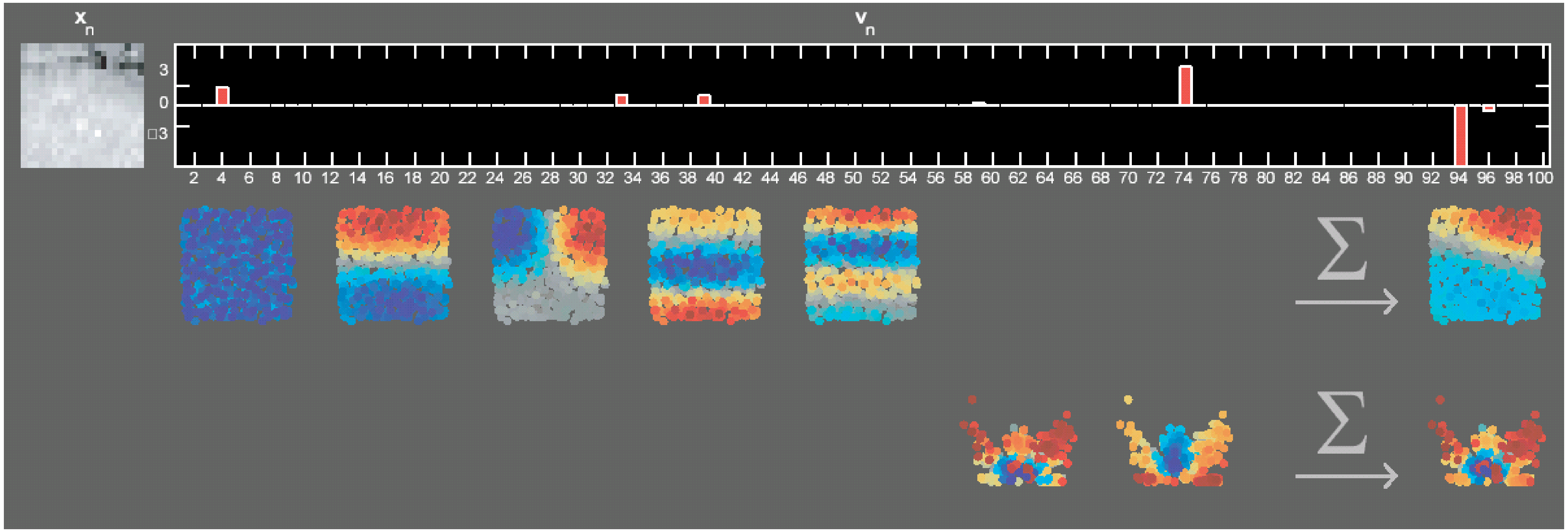


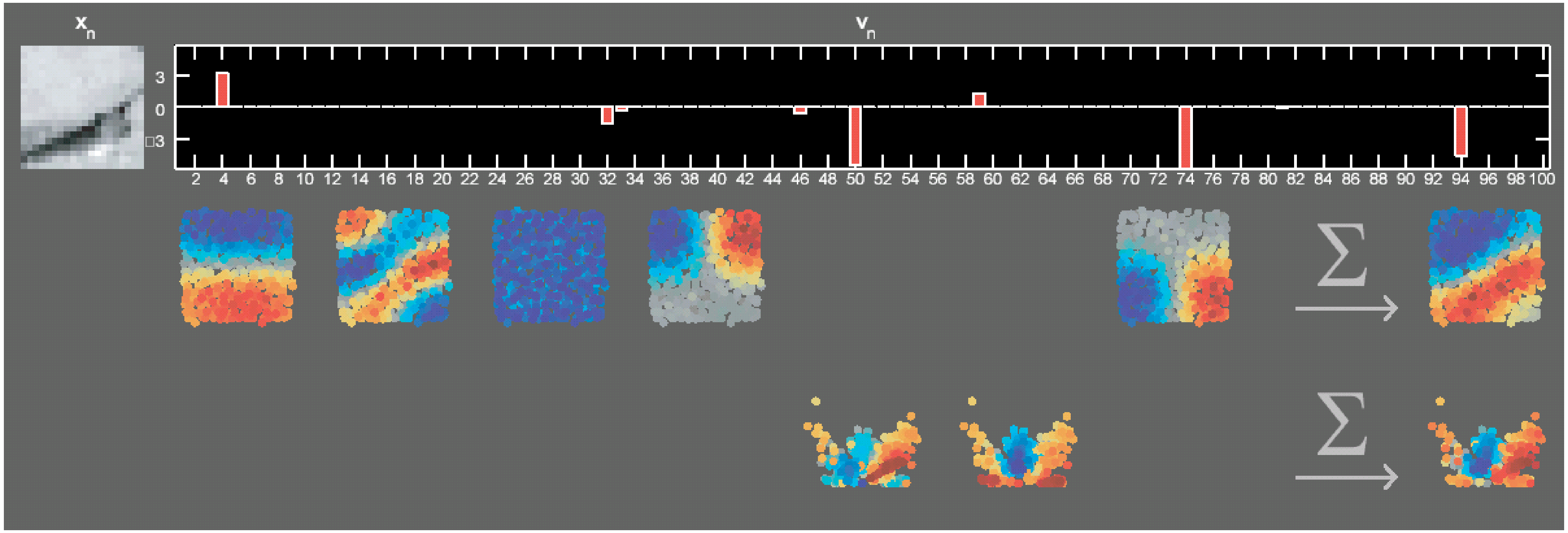
ICA on non-linearity reveals no structure



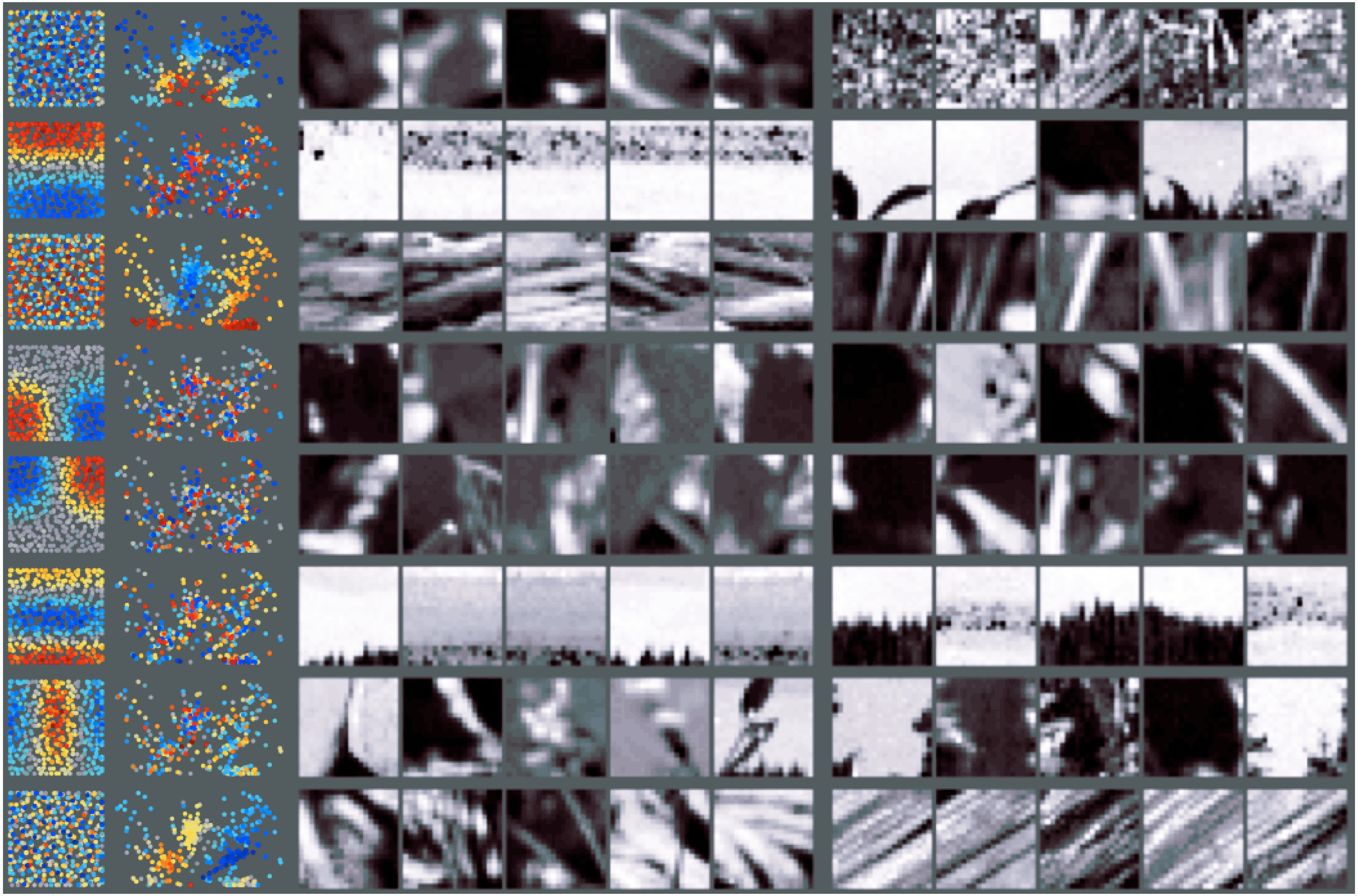
subset of ICA basis functions on log |s|

Inferred v forms a sparse distributed code

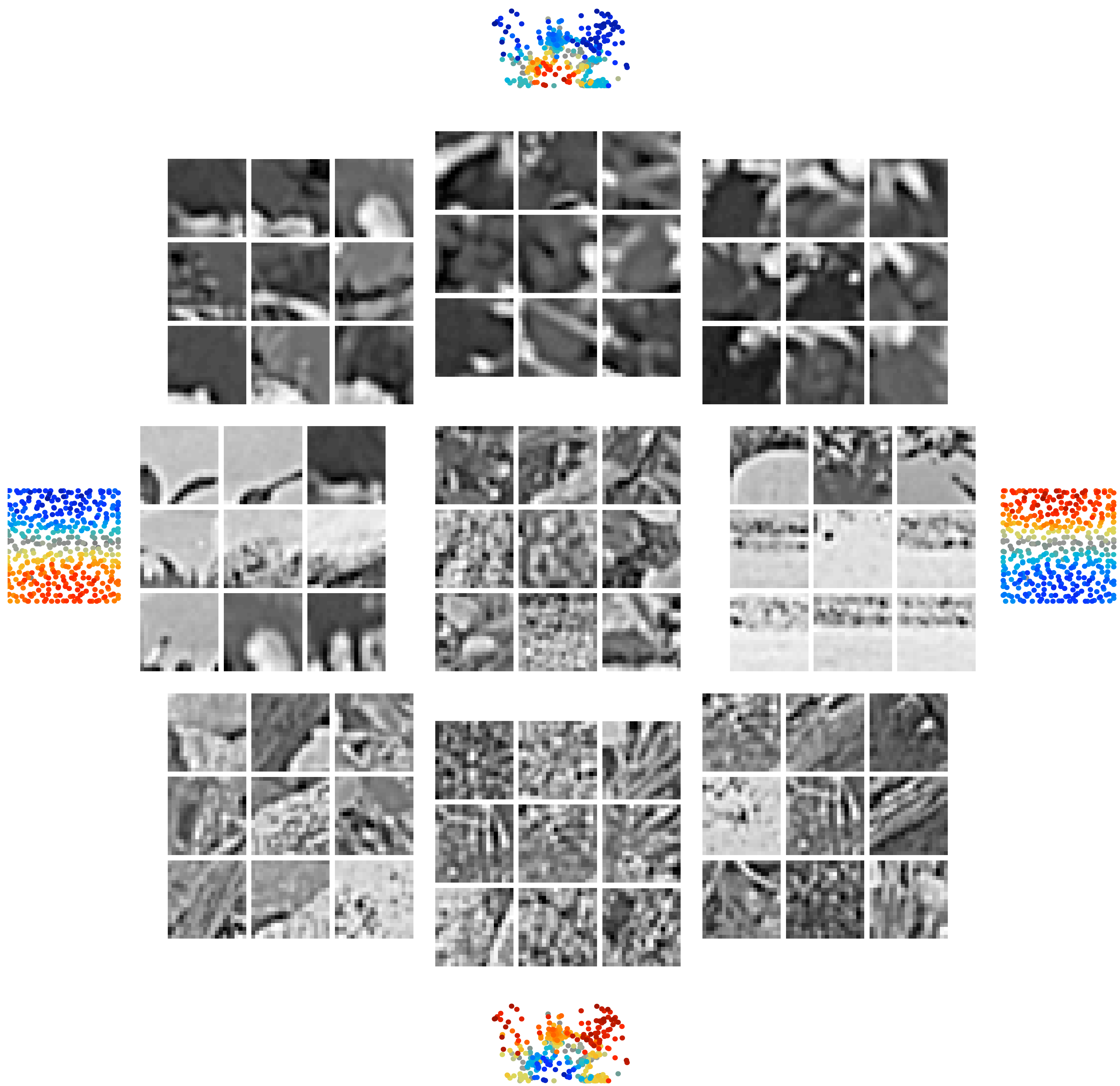




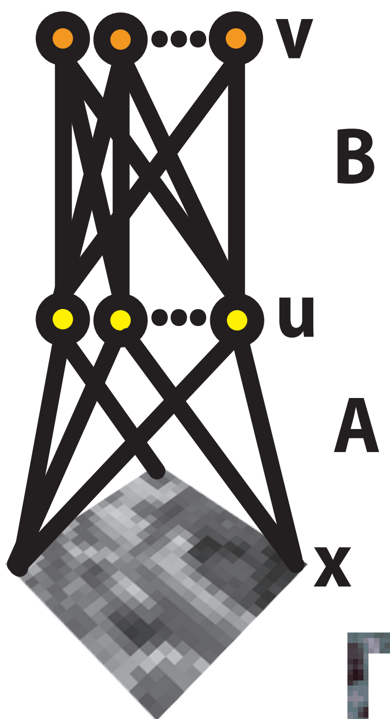
Most typical images for selected density components

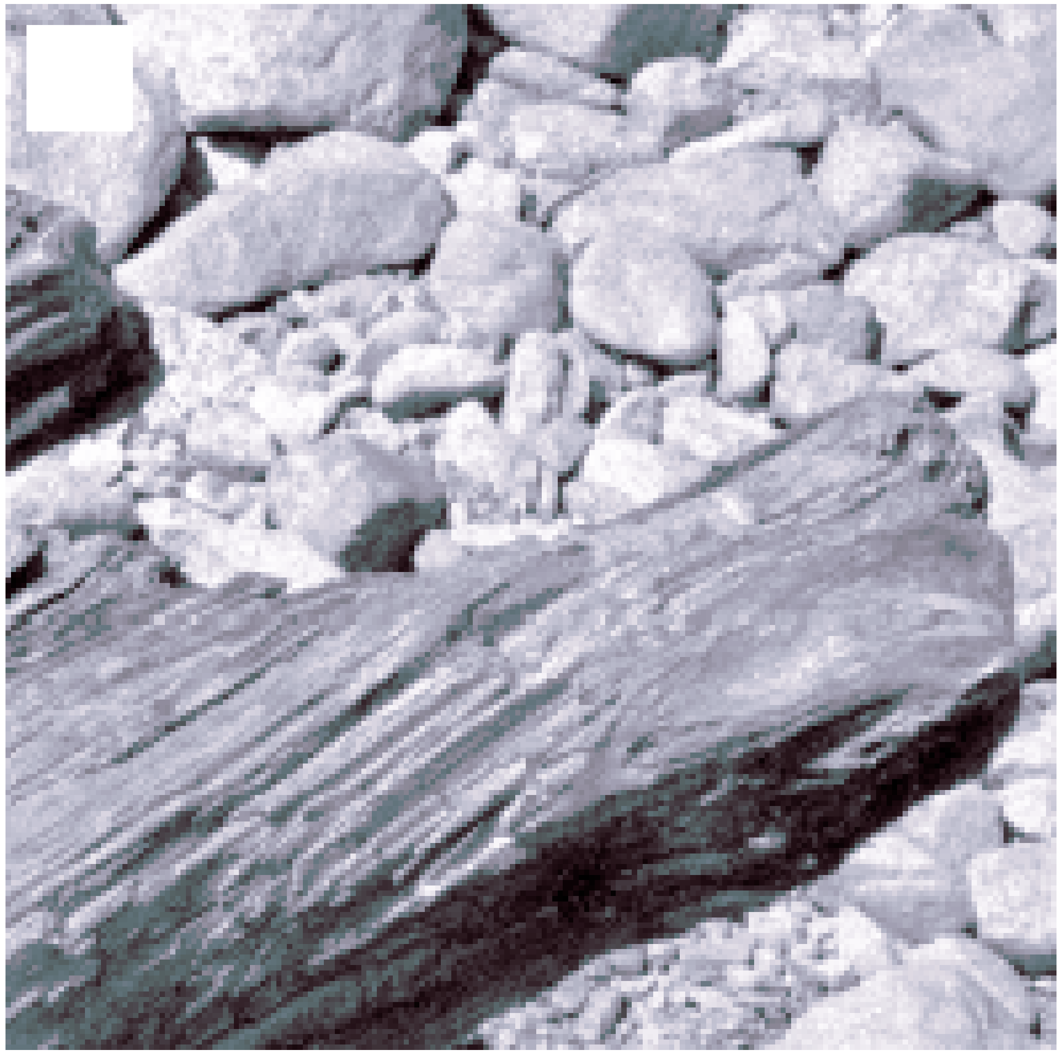


Distributed representation of natural image densities

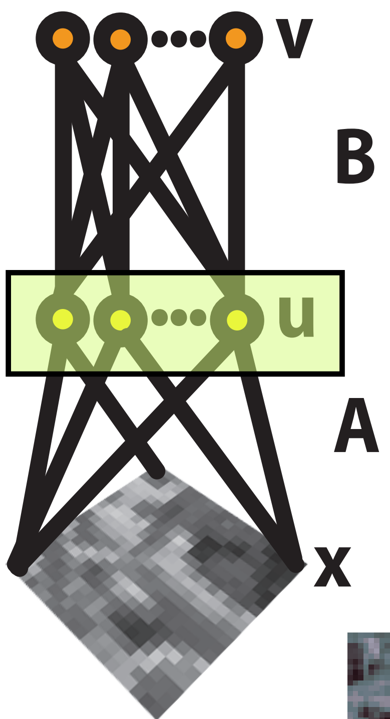


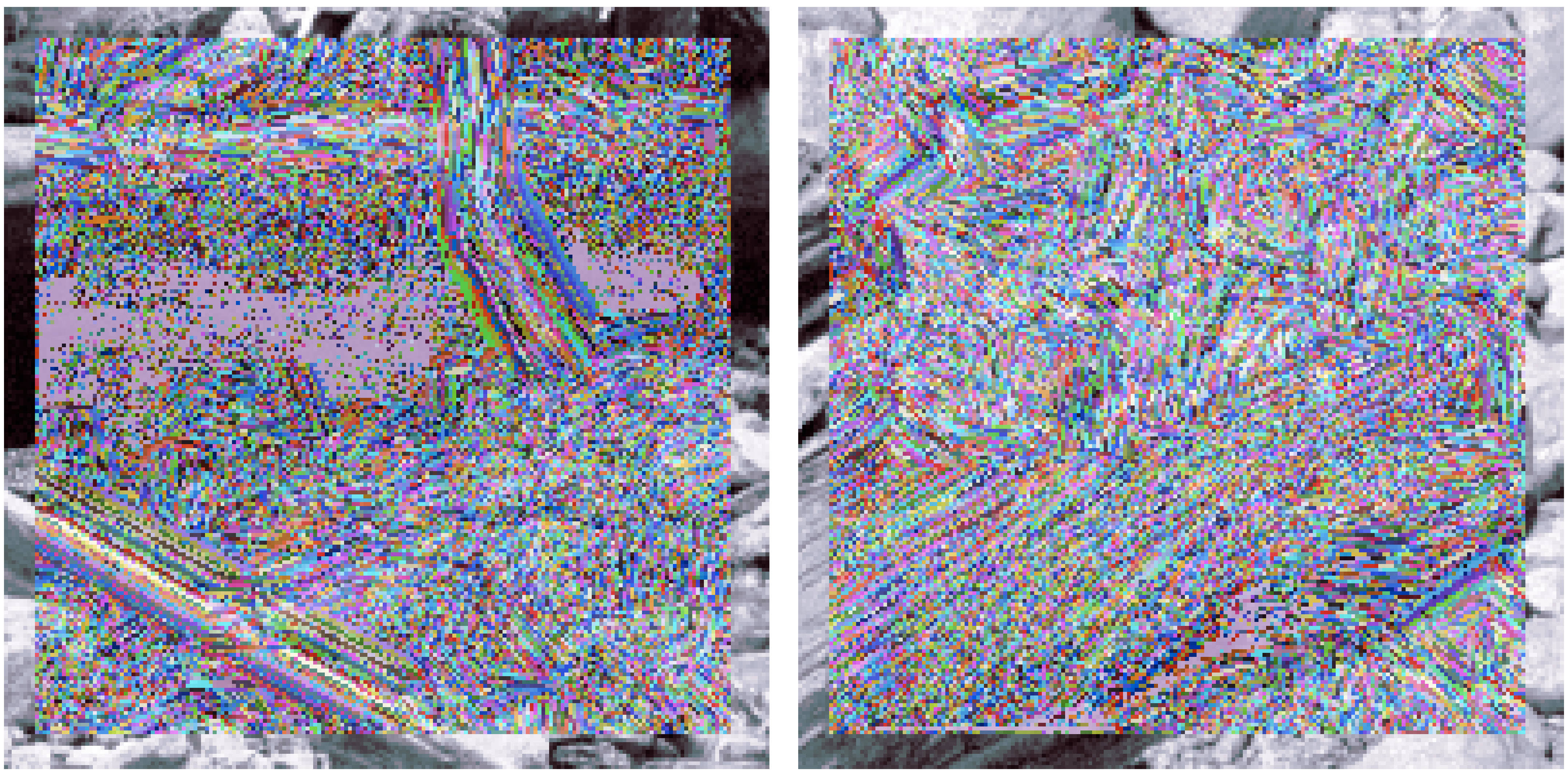
Comparing the degree of abstraction



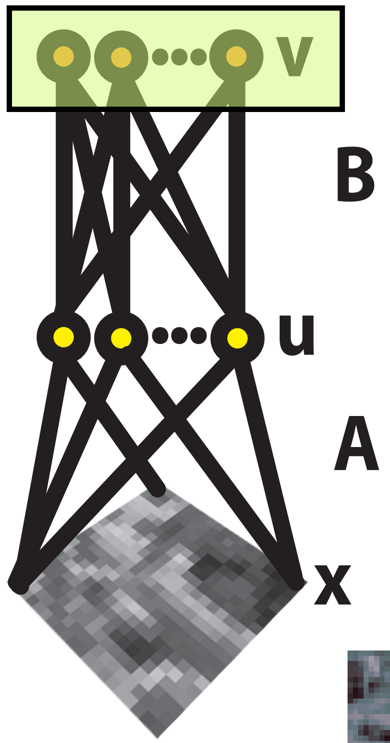


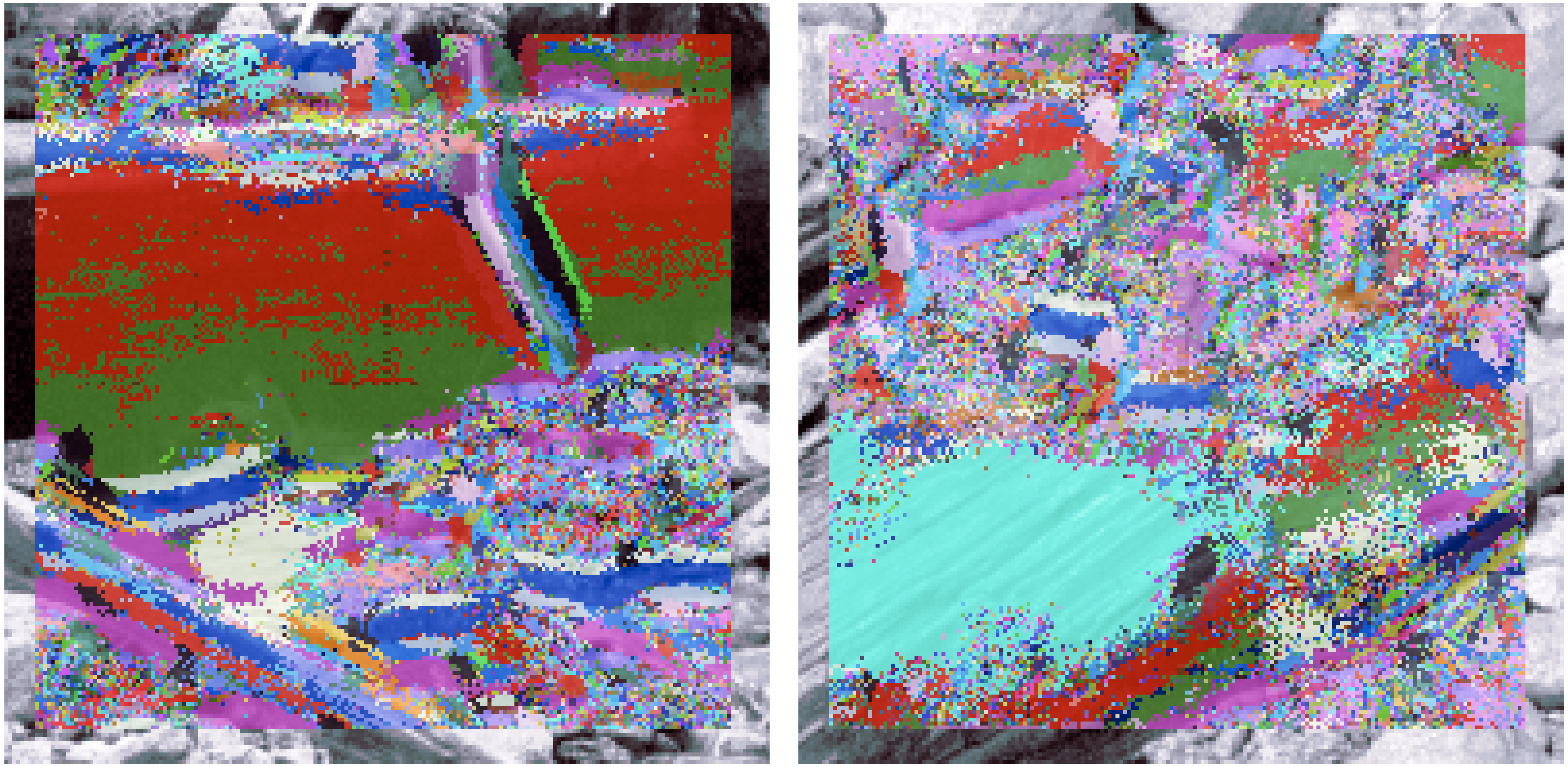
Maximum $|u_i|$ for each pixel



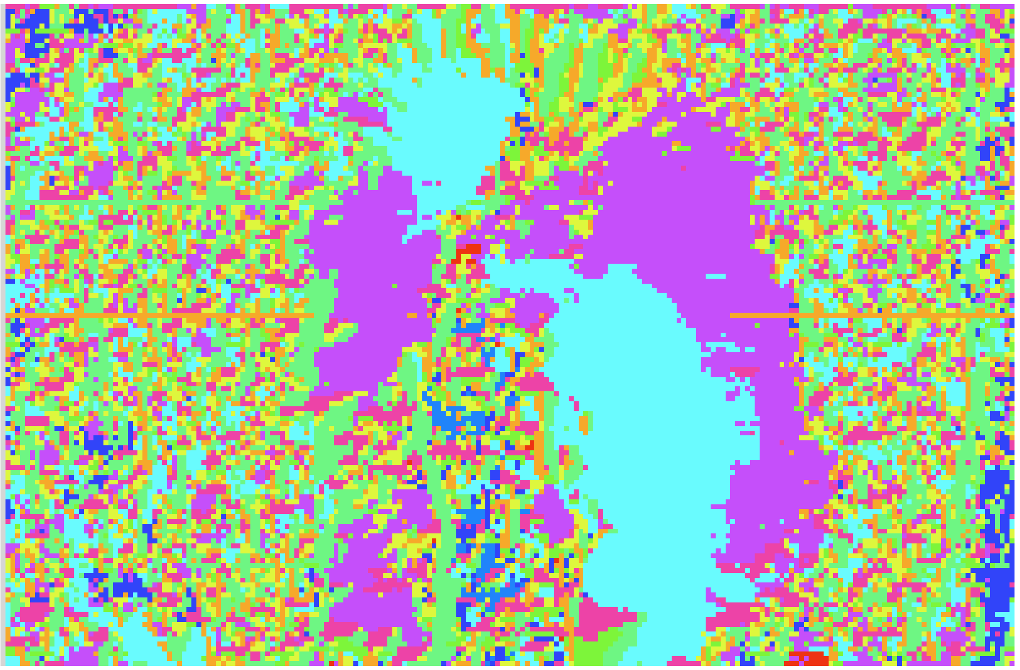


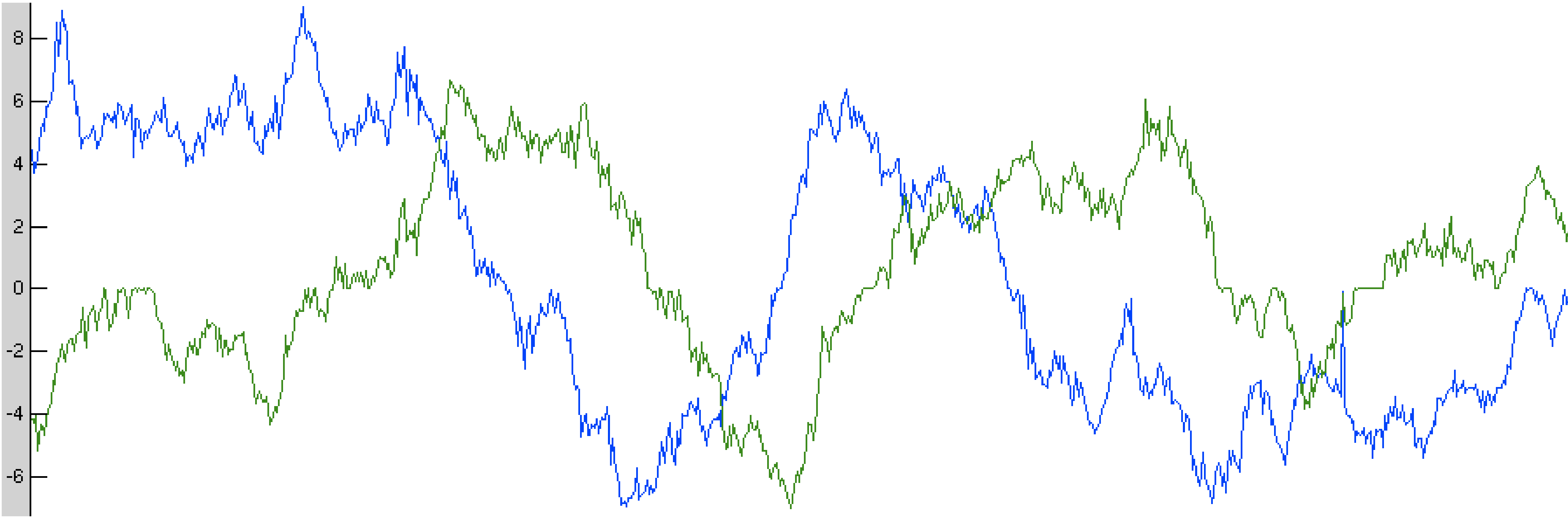
Maximum $|v_i|$ for each pixel



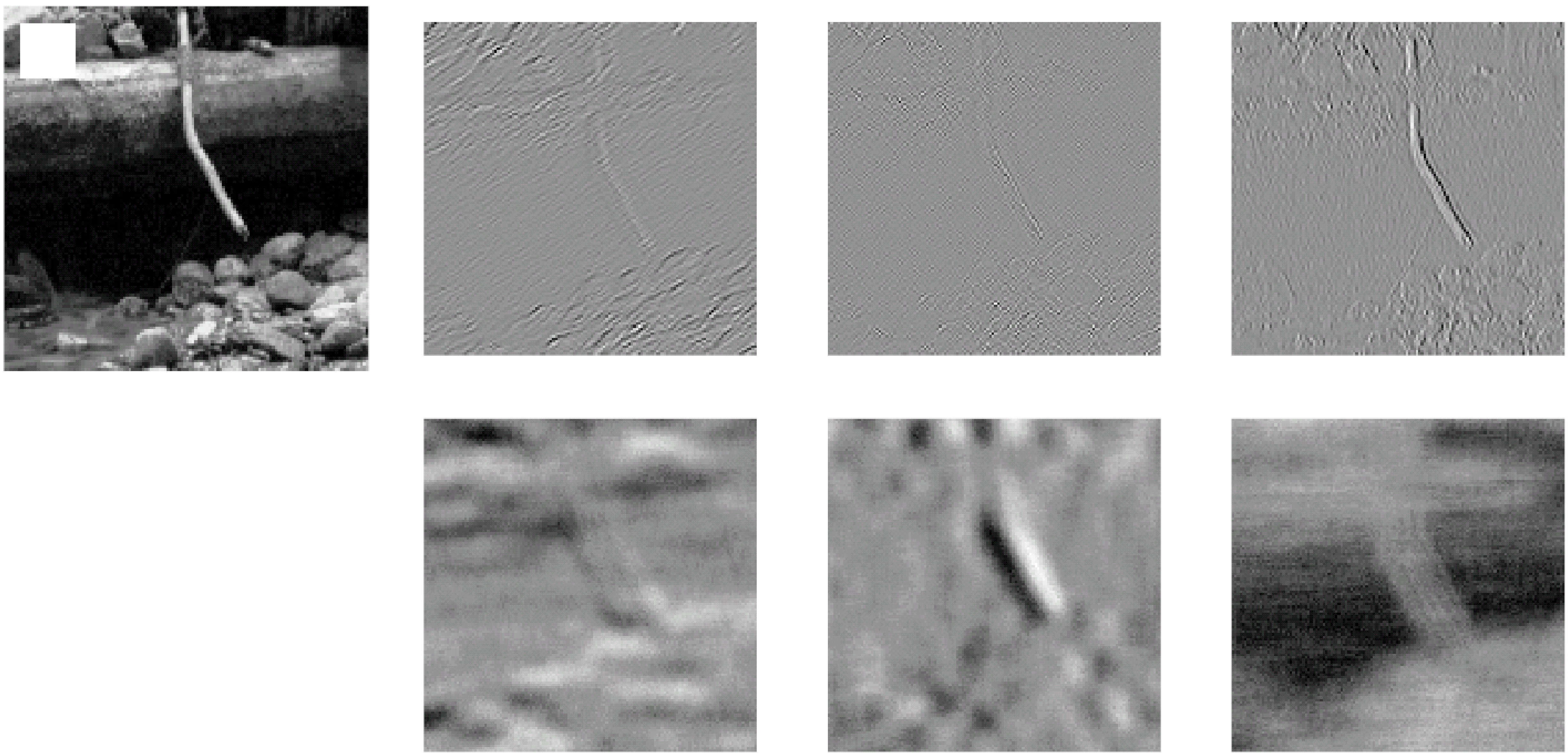


A distributed code for visual surfaces





Activation maps for three different v units



Learning higher-order codes for scanned newpaper

ıy Picture-Perfect Slice of the

ing on Route 4 over a dark and winding mountain pass, the visitor suddenly emerges into a "Lost Horizon" world of hot springs, trout streams and meadows of wildflowers, where cattle and the state's largest elk herd graze side by side.

But the same sense of wide-open Western independence evoked by the vistas has prevented the sale of the land for years. And the deal that is being negotiated for the ranch, which has been owned by one family for almost 40 years, is as much about Western attitudes toward public land as it is about money.

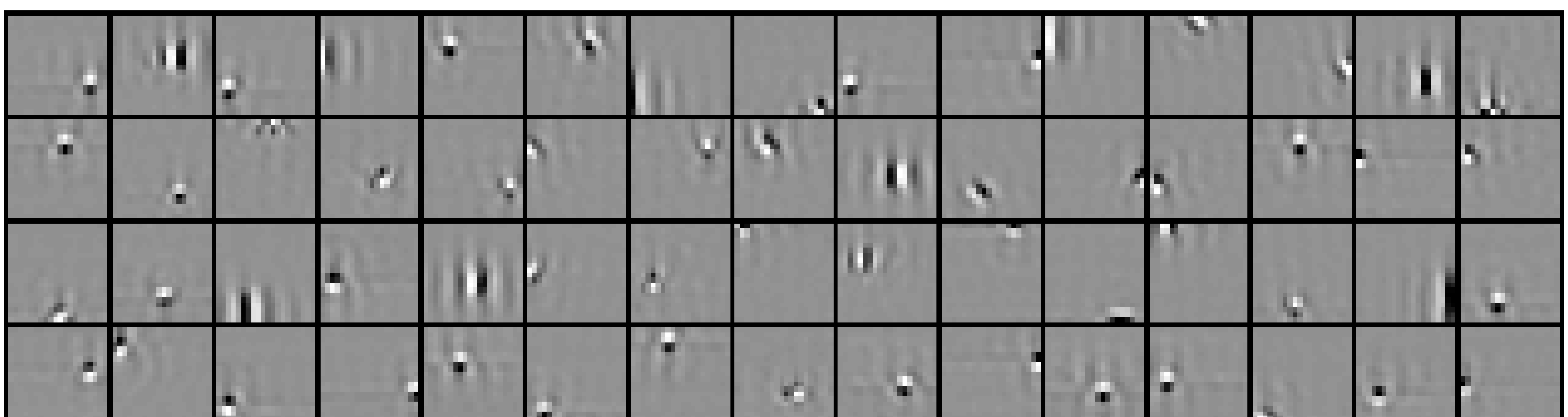
The Administration has long supported the purchase of the ranch, which has been called "the hole in the doughnut" because it is an island surrounded by the Santa Fe National Forest. Last February on a visit to

New Mexico, President
had Air Force One mak
fly over the ranch for a l
its dominant feature —
wide crater of the dorn

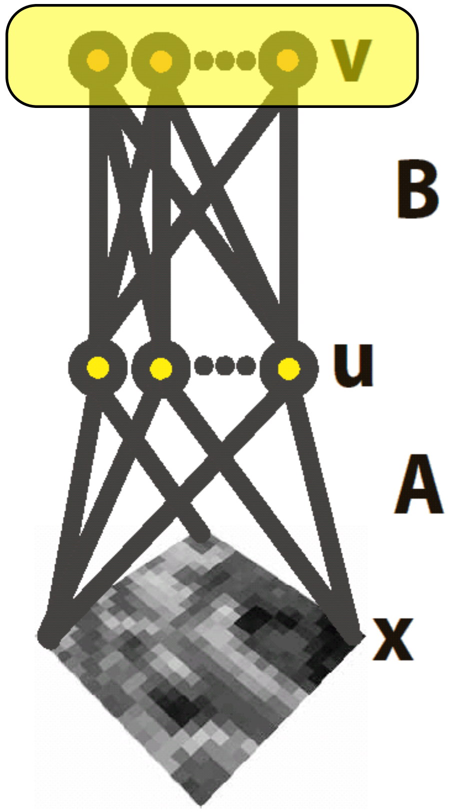
Republicans, noting
one-third of New Mexi
owned by the Federal
have long opposed
chases. But in August
sentiment began shift
cally.

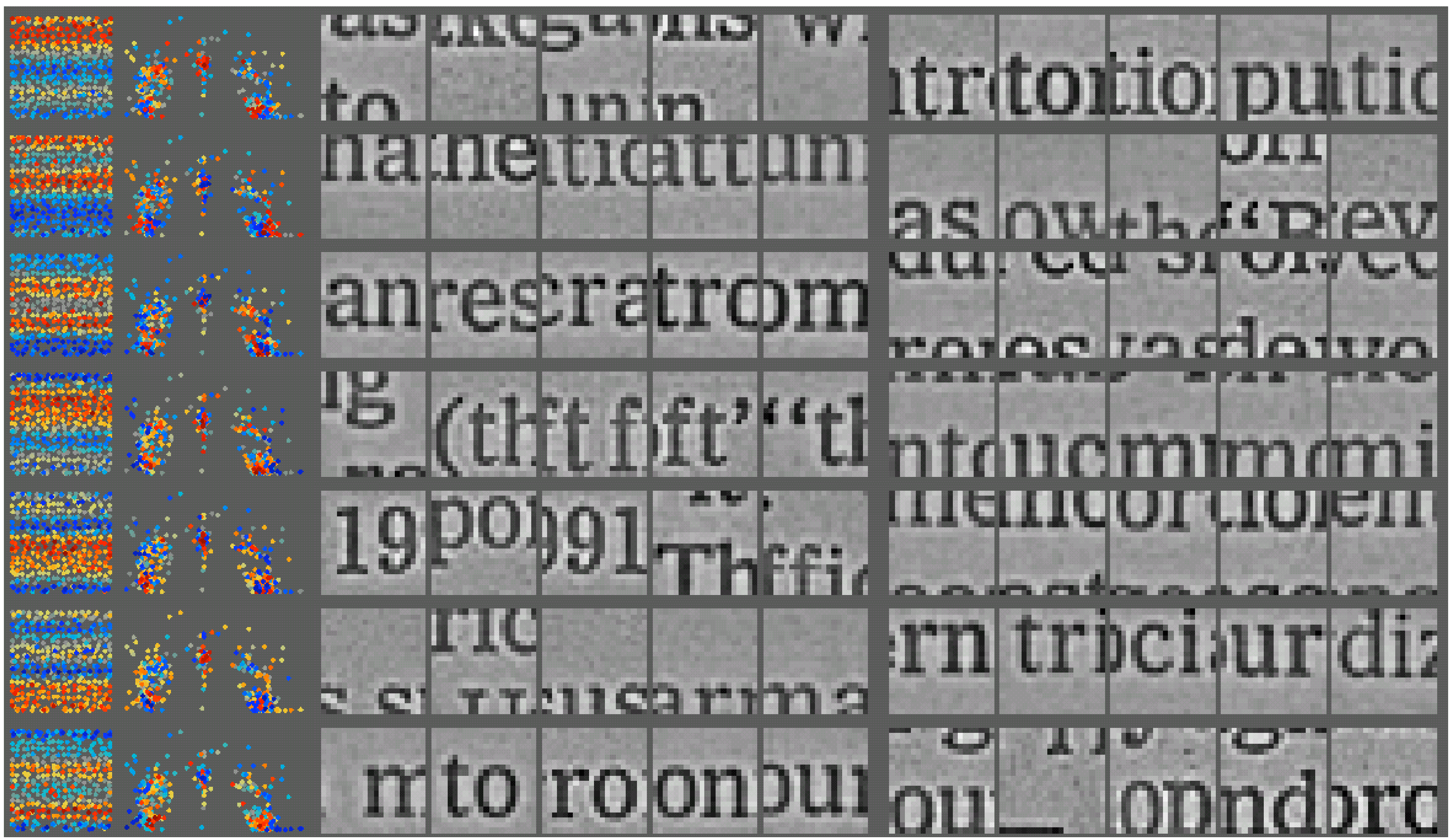
Under legislation dra
tor Pete V. Domenici,
New Mexico, the Baca
separate unit of the Na
system, owned by the
Forest Service, but m
trust, comprised of n
pointed by the Presiden

Continued on Pa

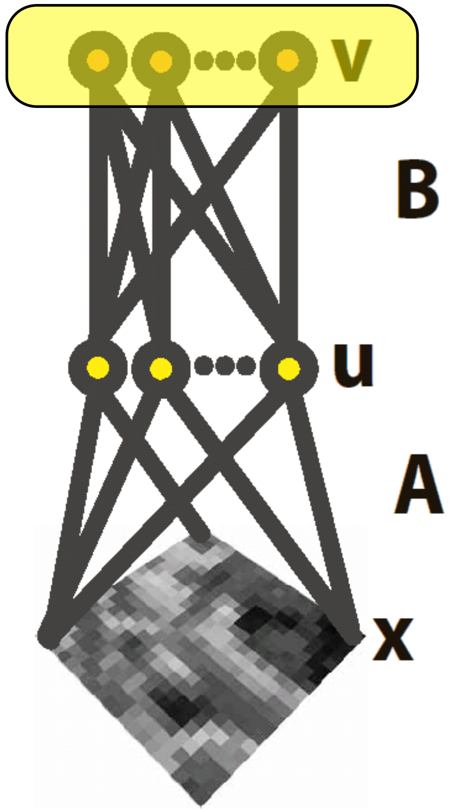


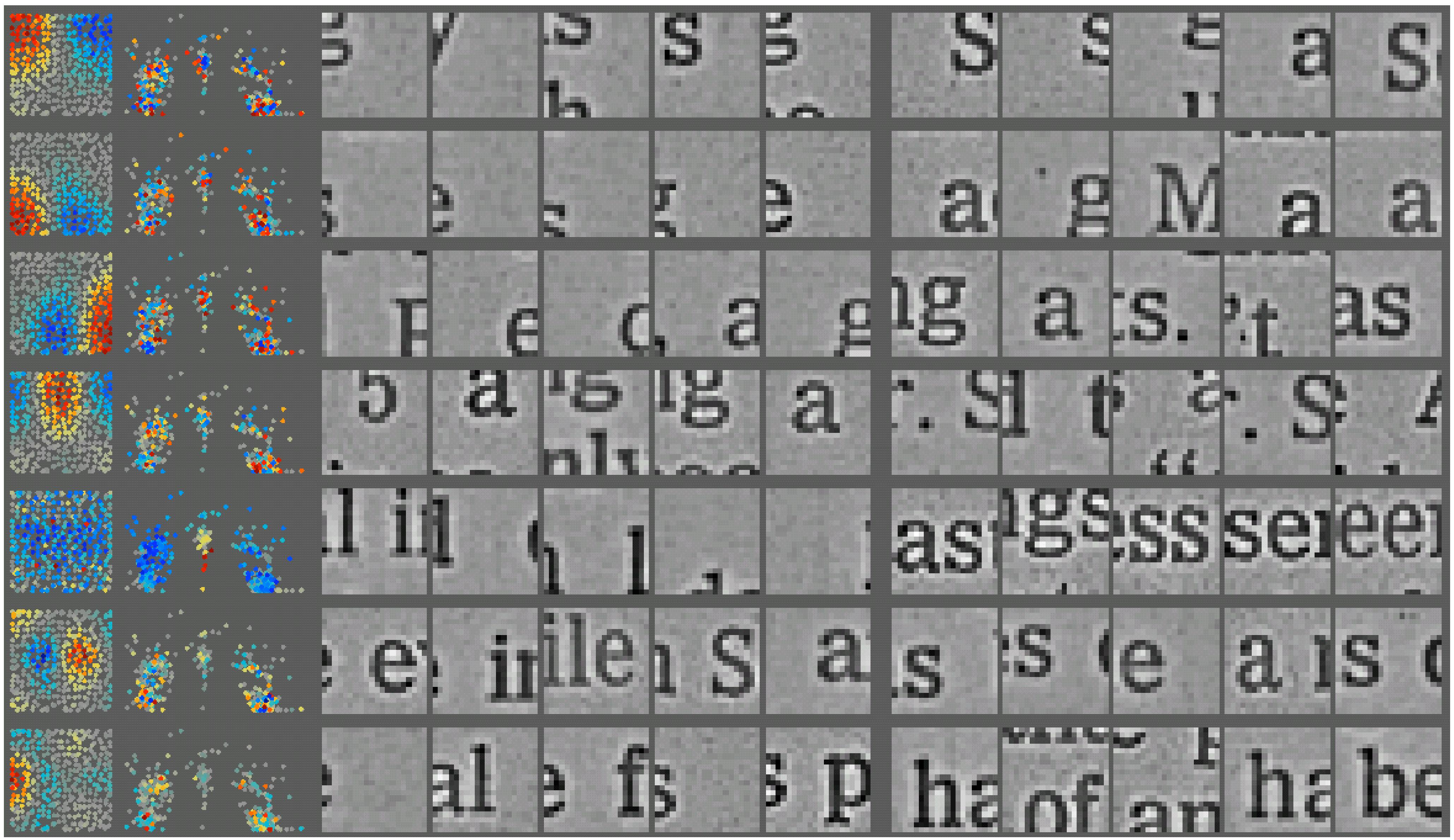
Most typical images for selected density components



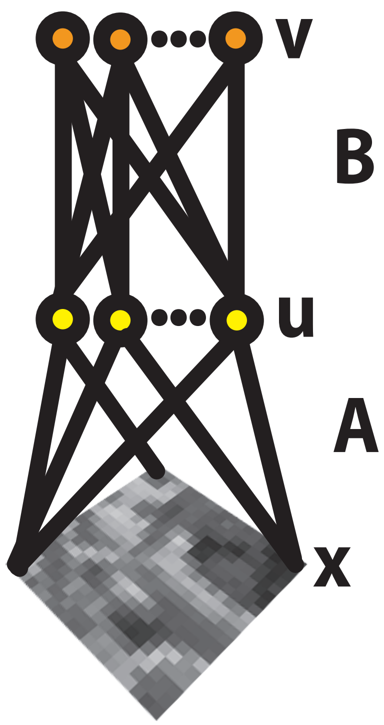


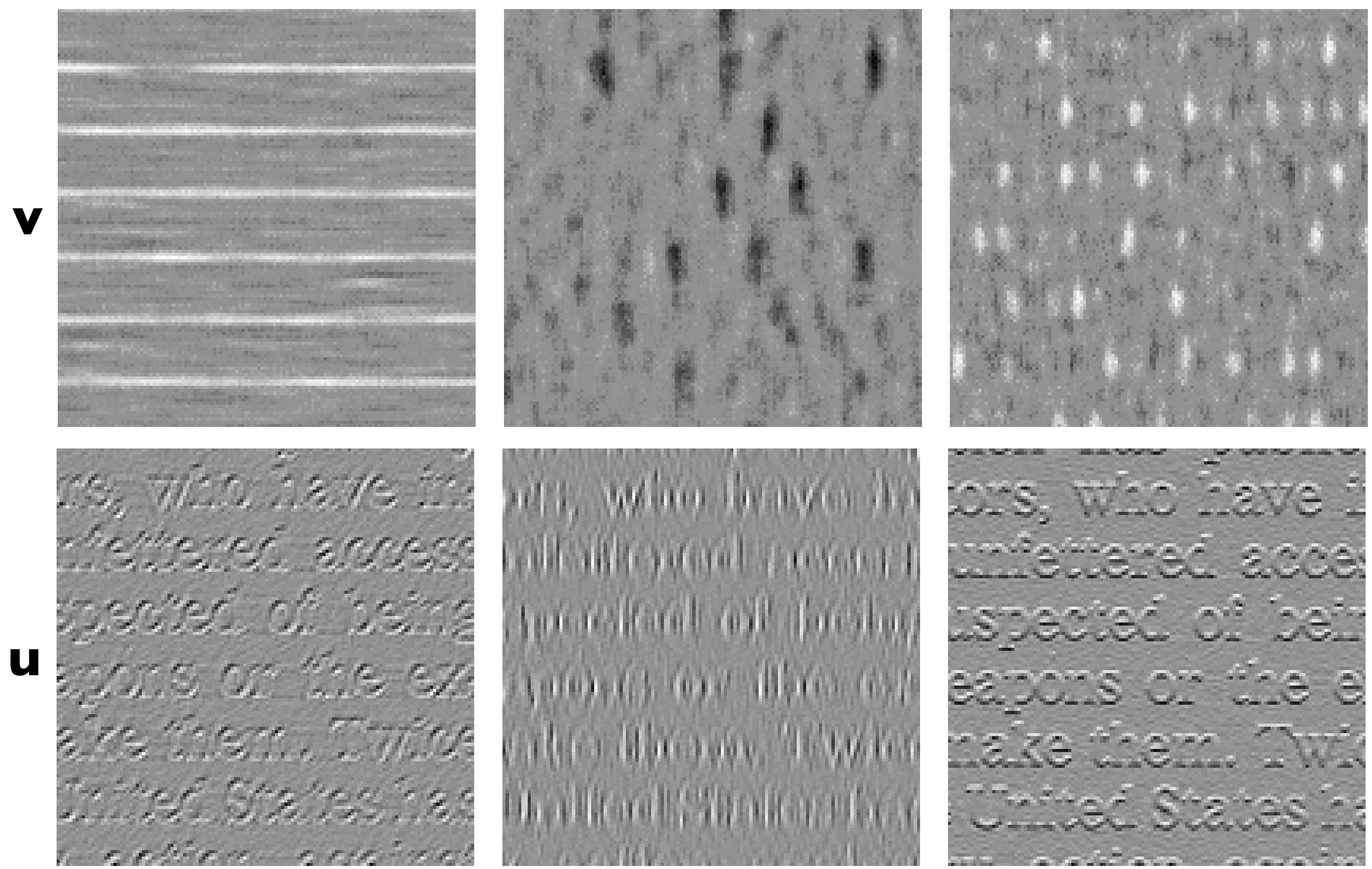
Most typical images for selected density components



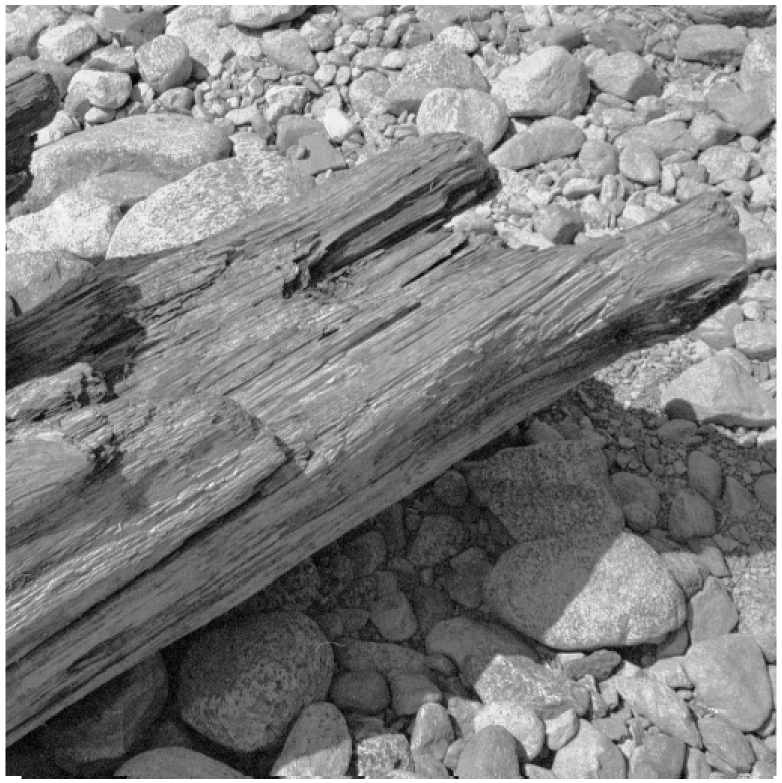


Higher-order codes for scanned newspaper text

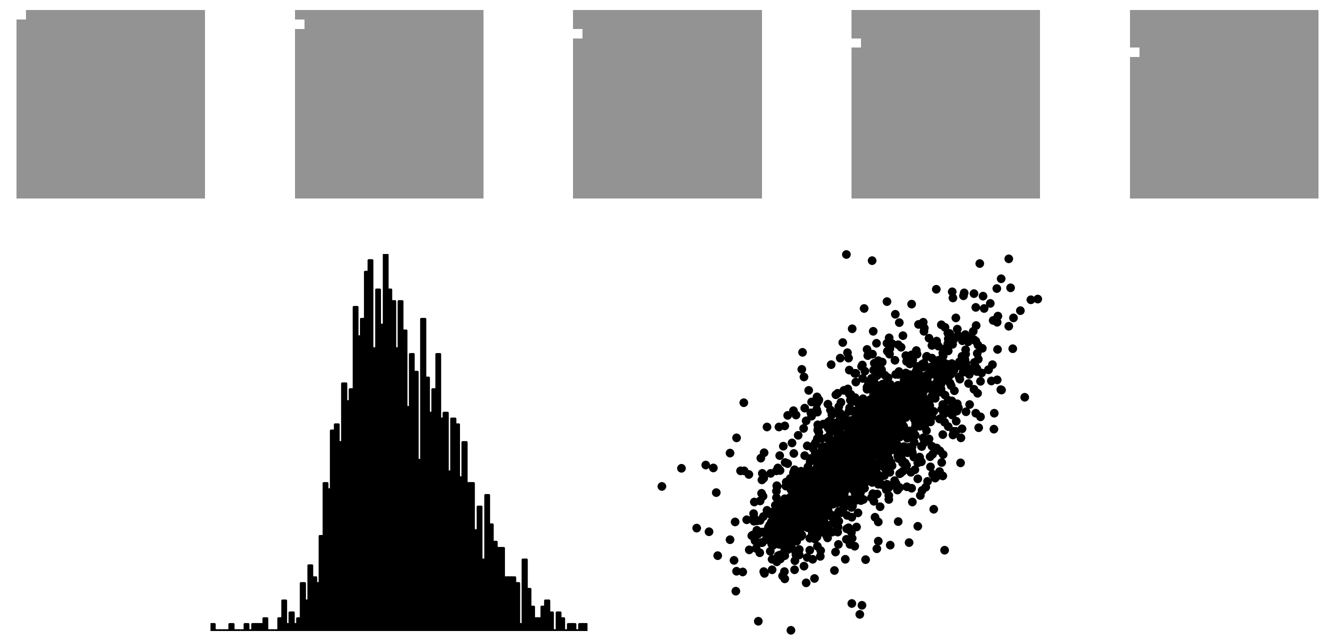




Evaluating a pixel model by synthesis



Representation



Sampling from the model we get...

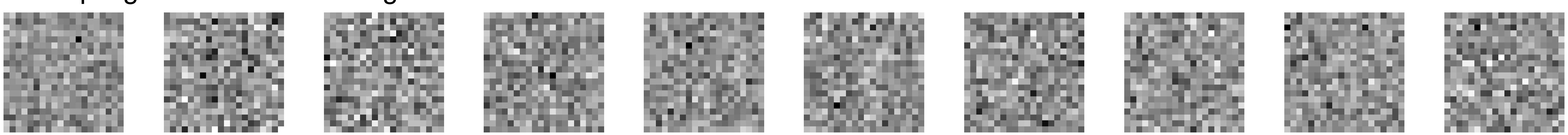
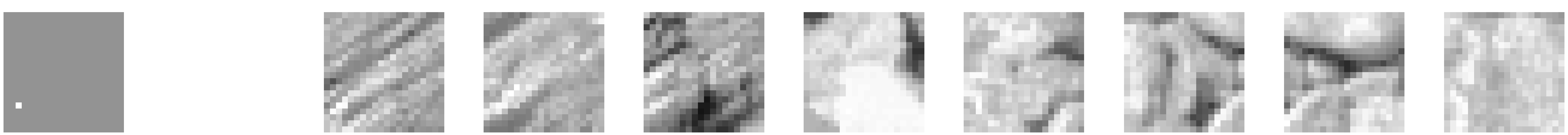
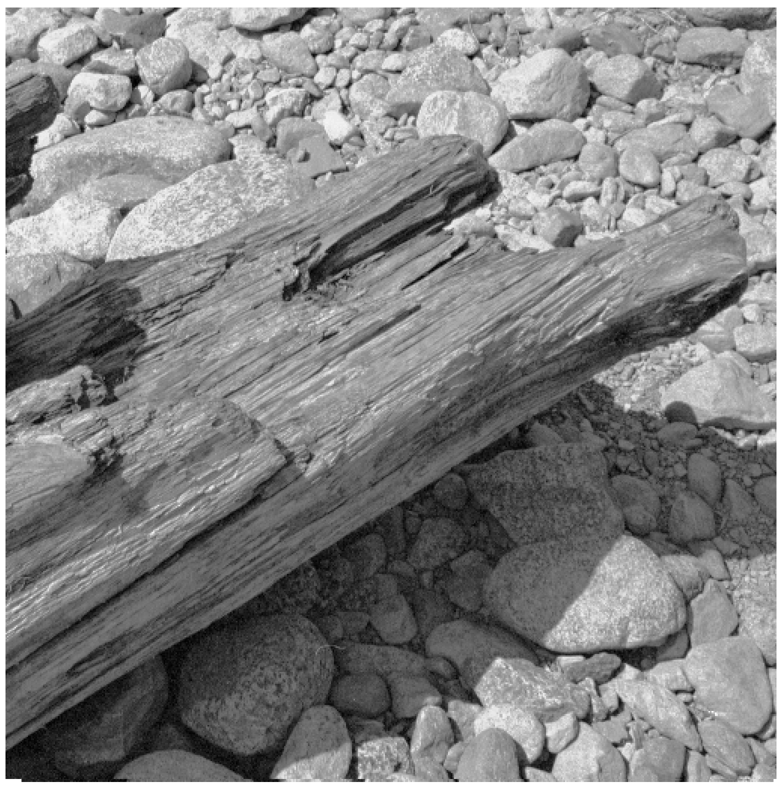


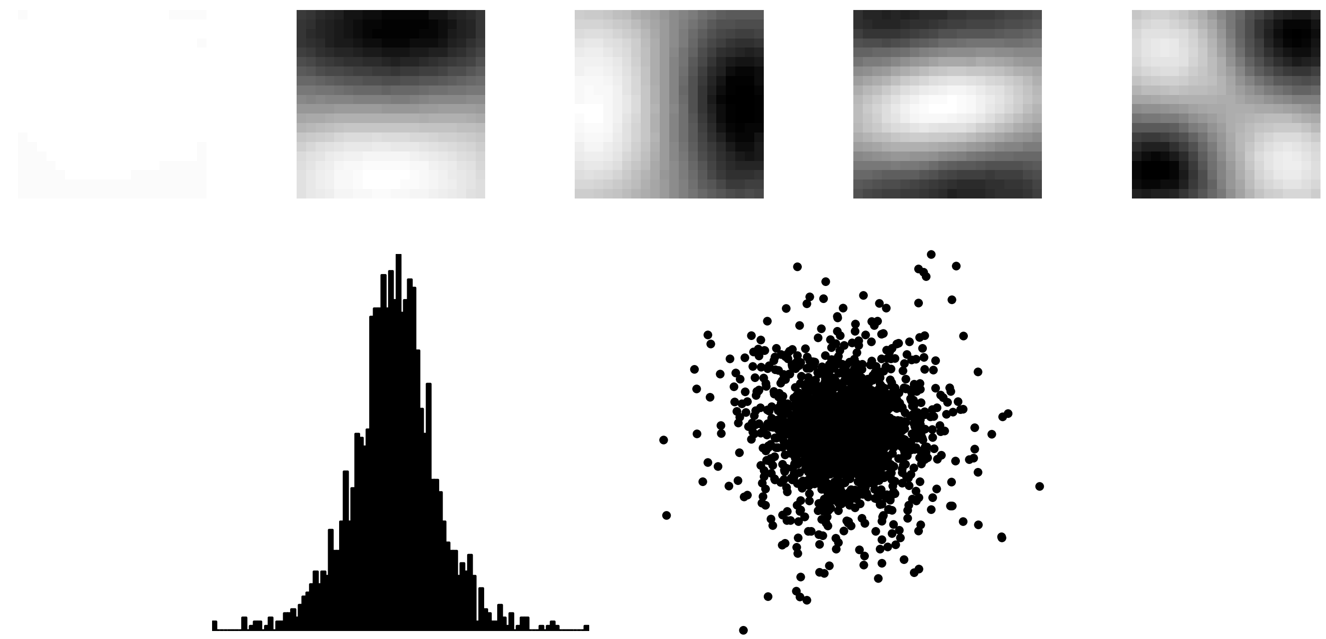
Image data that maximally activates this unit...



PCA model



Representation



Sampling from the model we get...

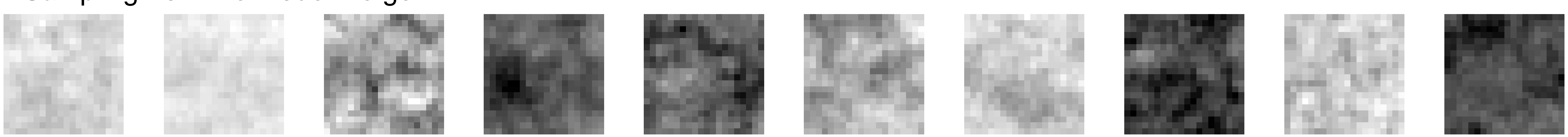
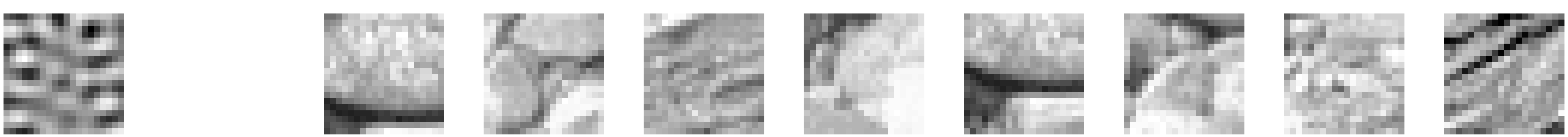
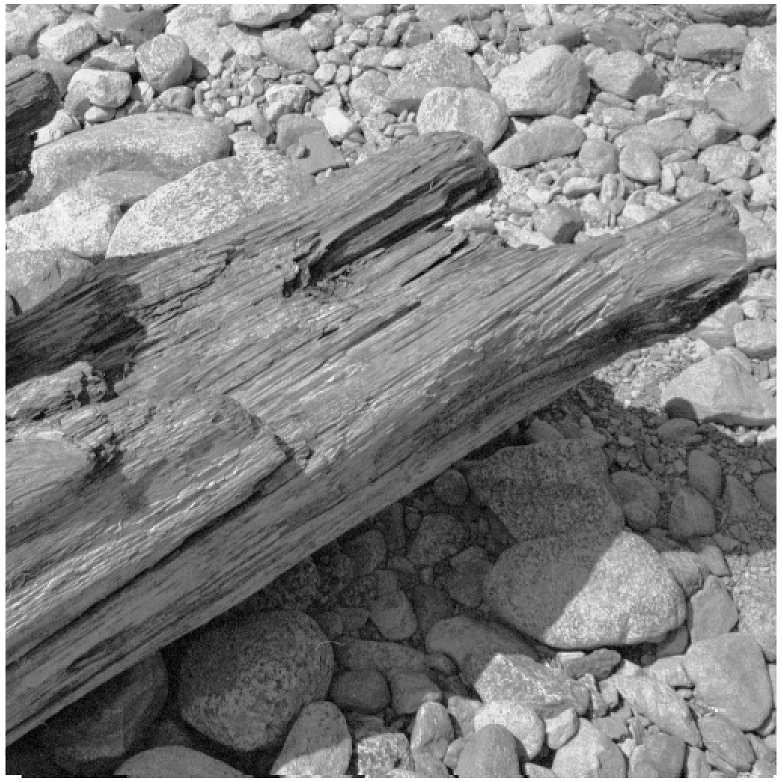


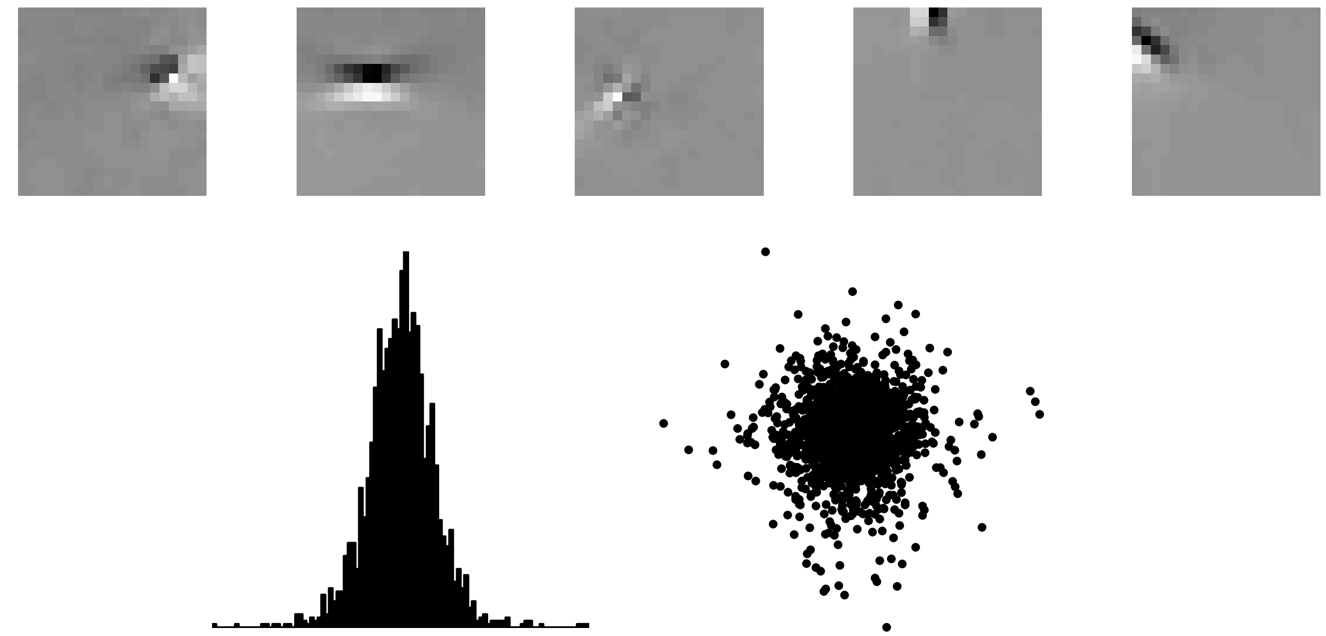
Image data that maximally activates this unit...



ICA model



Representation



Sampling from the model we get...

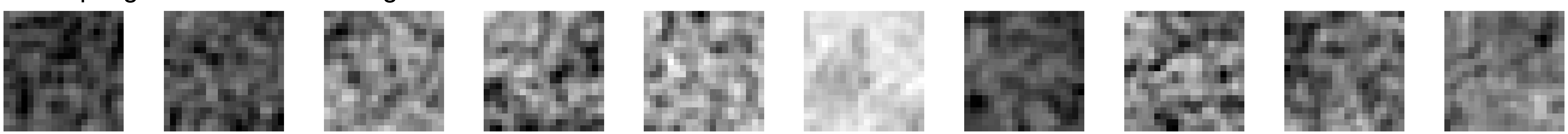
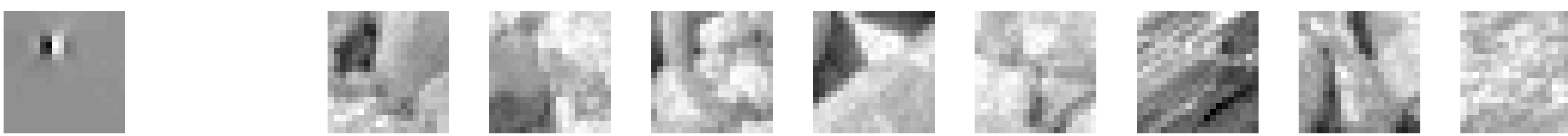
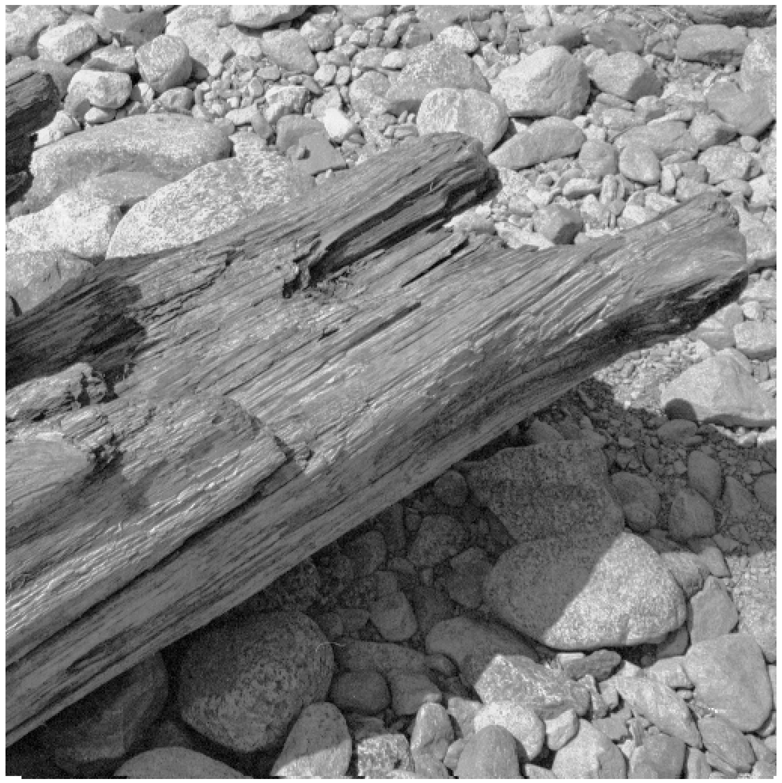


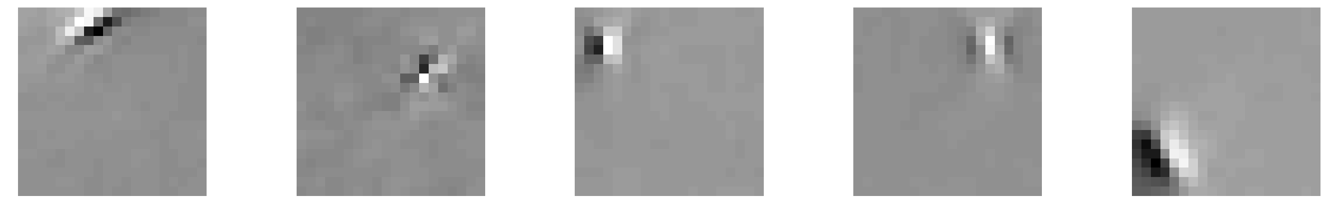
Image data that maximally activates this unit...

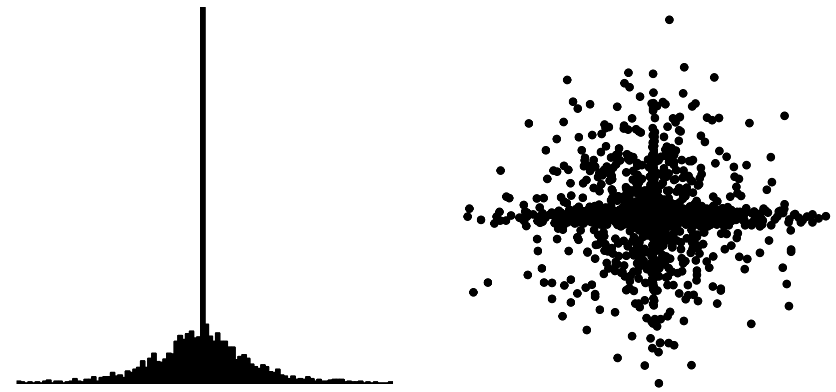


Hierarchical, non-linear model



Representation





Sampling from the model we get...

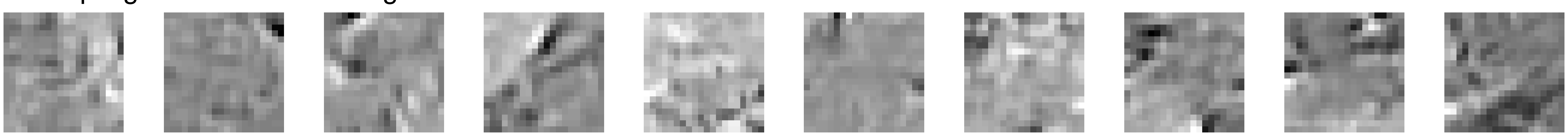
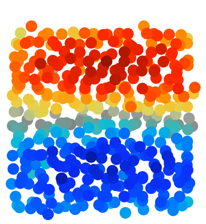
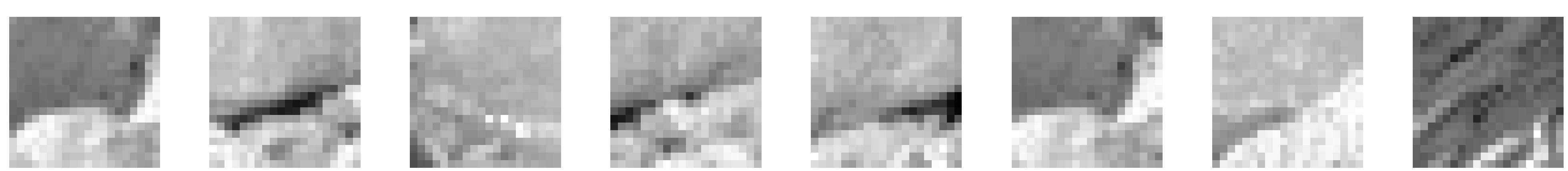


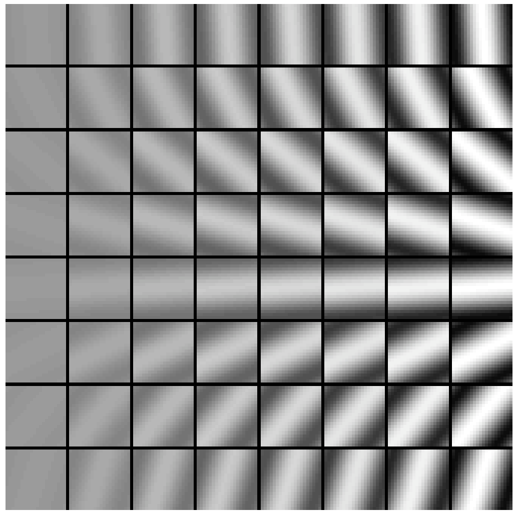
Image data that maximally activates this unit...



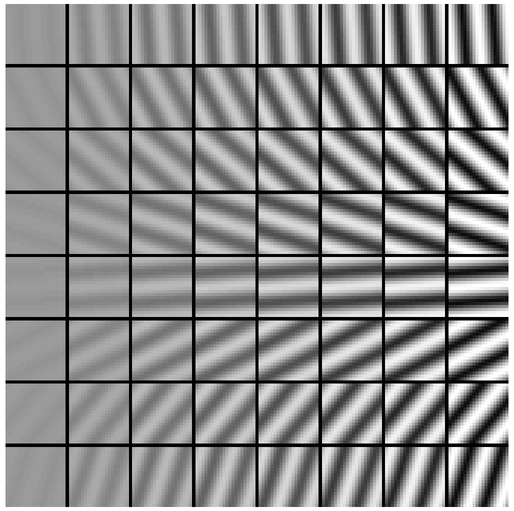


Conventional physiology: orientation vs contrast by freq

freq=1

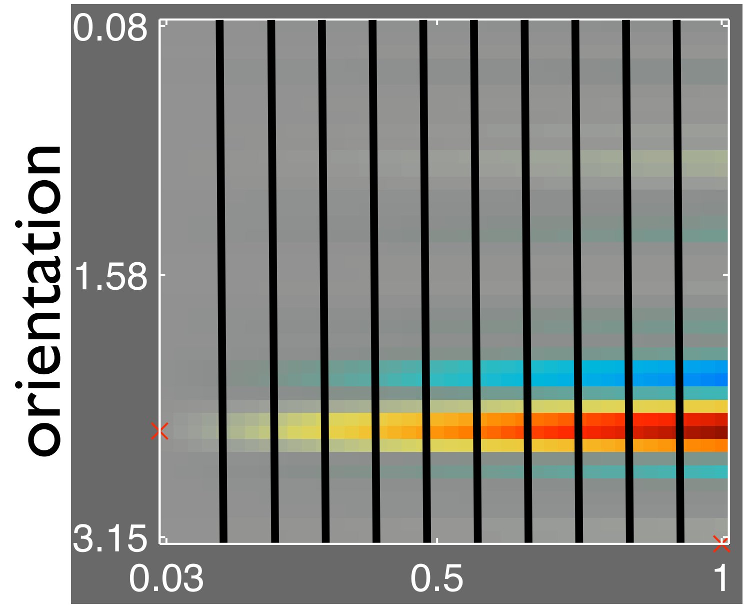


freq=2



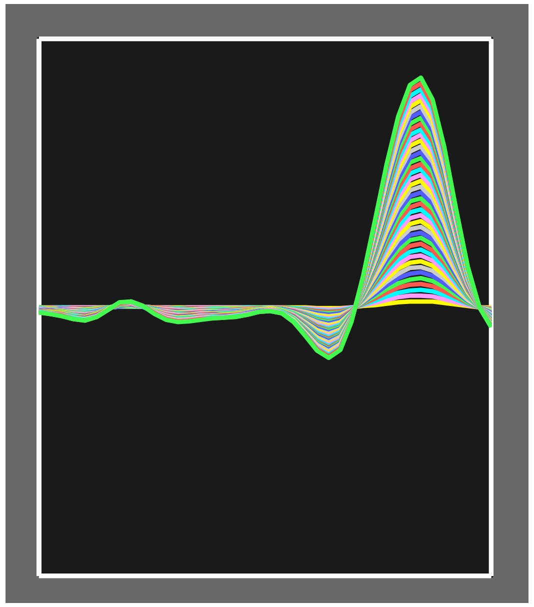
•
•
•

2D u map at best freq



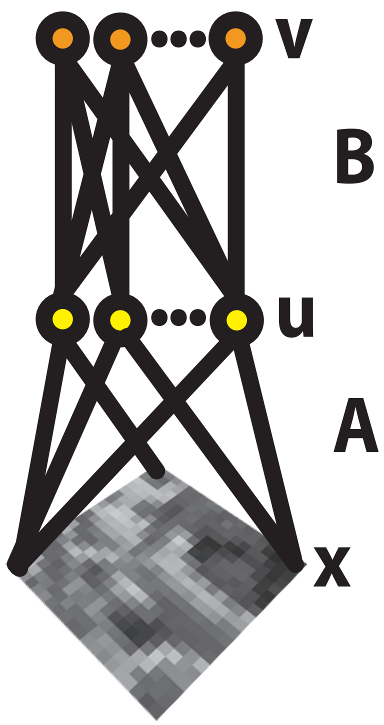
contrast

line u map

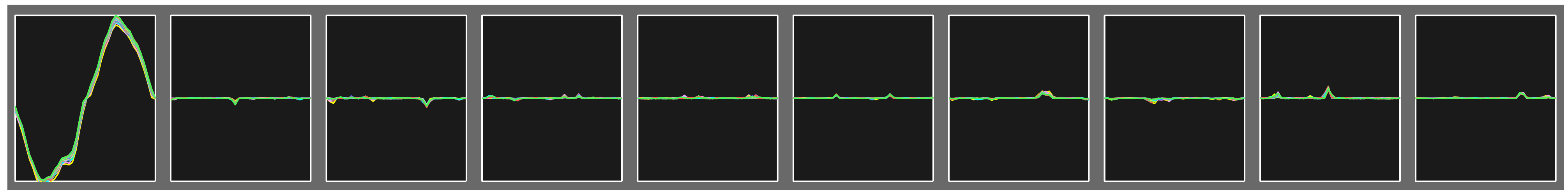


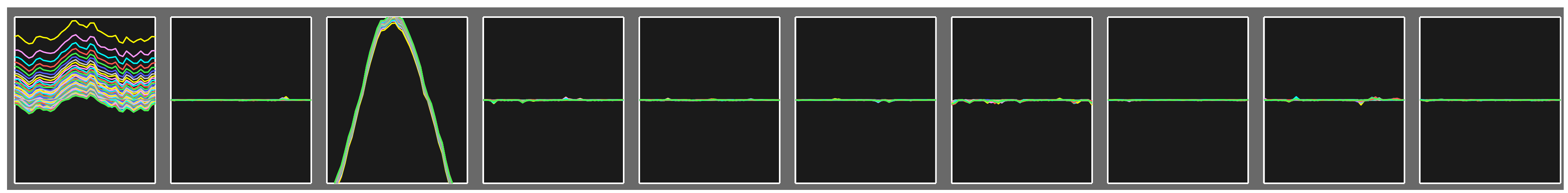
orientation

Higher-order representation is contrast independent

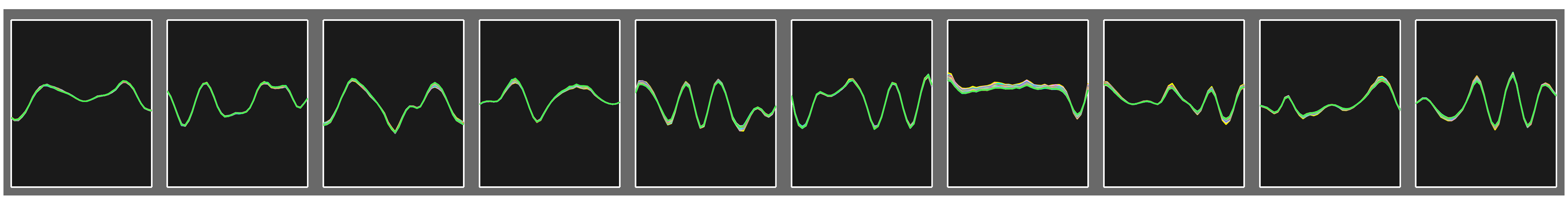


orientation (x-axis) by contrast (lines) at freq=3

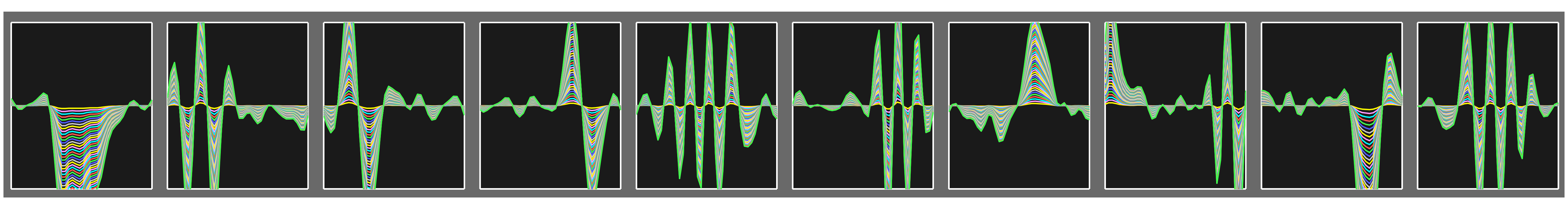




v activation is much more sparse and only a few units show contrast dependence

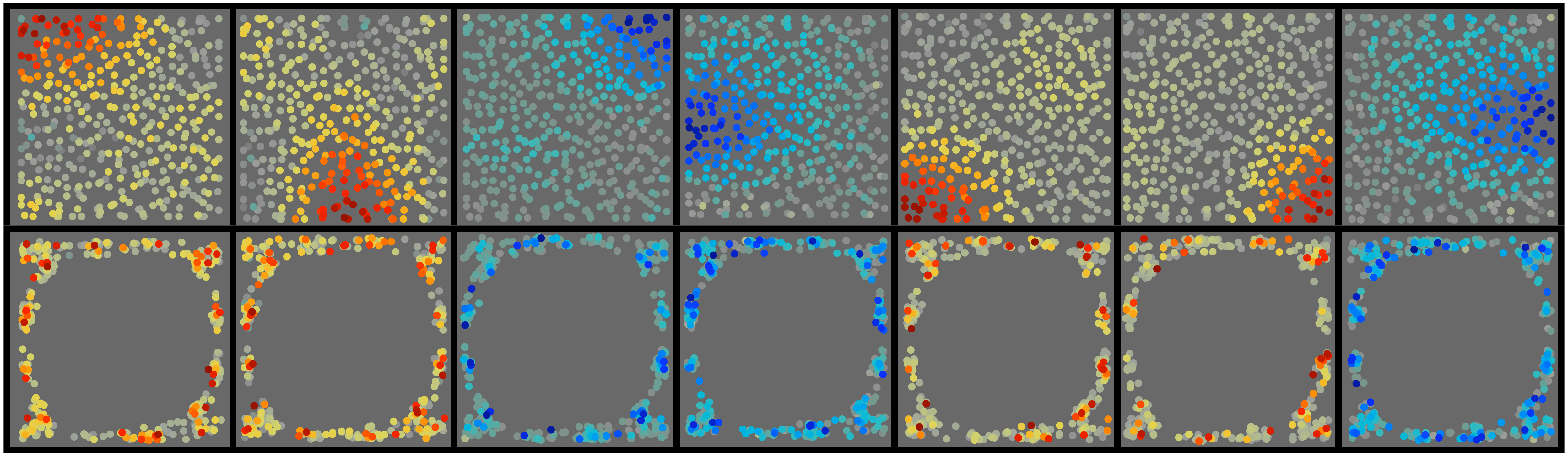


normalized u activation shows minimal dependence on contrast

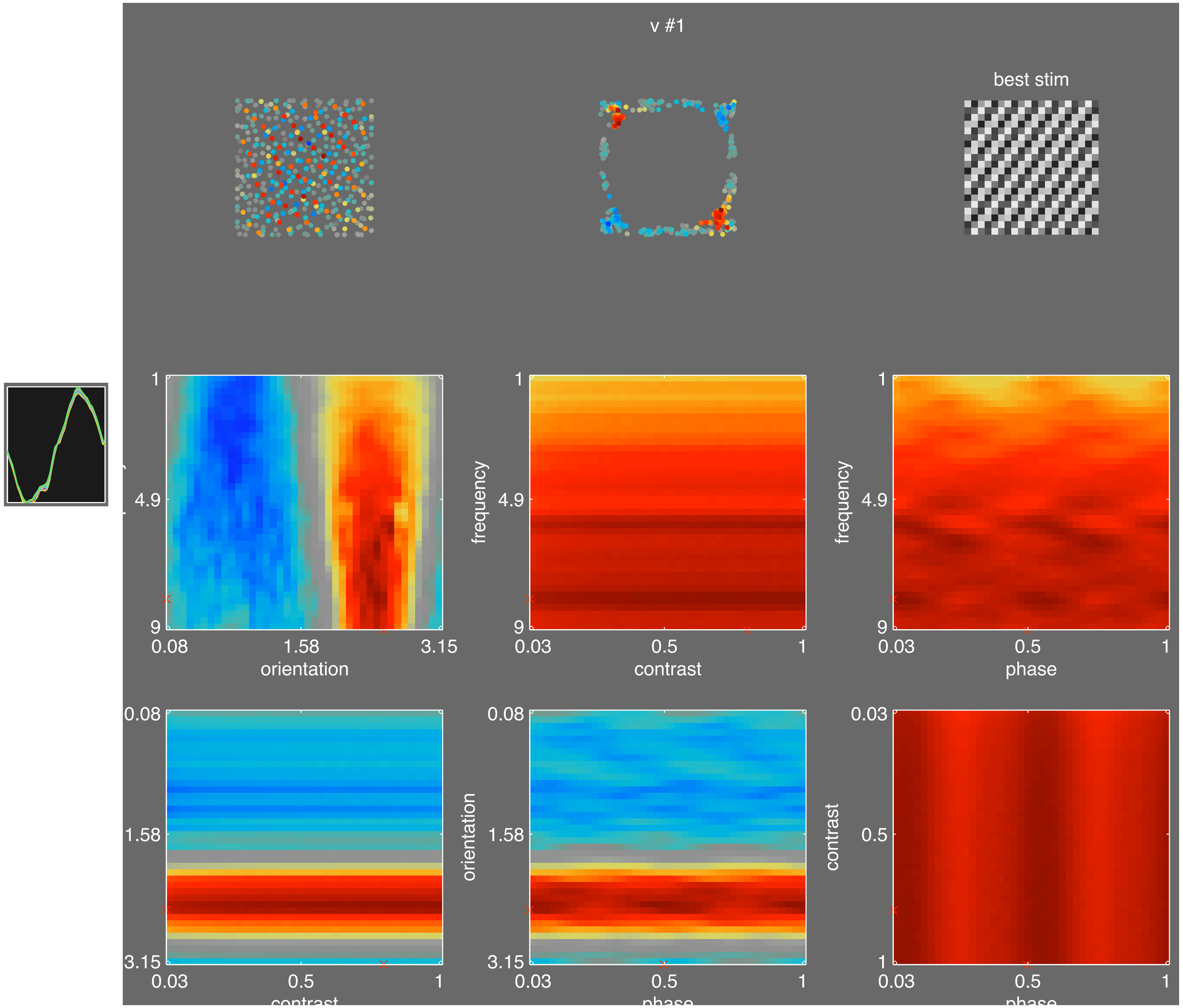


u activation shows strong dependence on contrast

The full set of contrast dependent v's

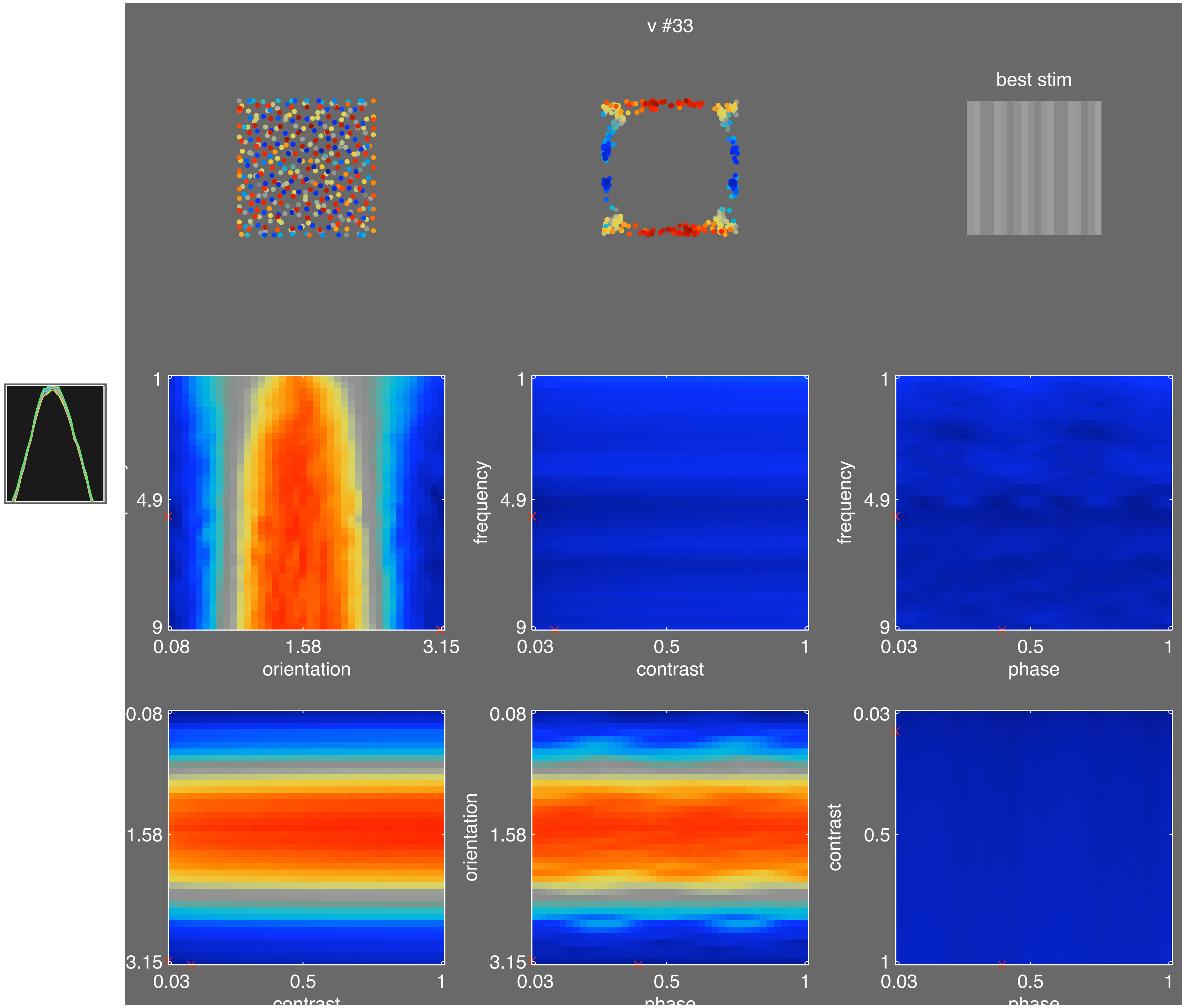


orientation dependence / contrast invariance

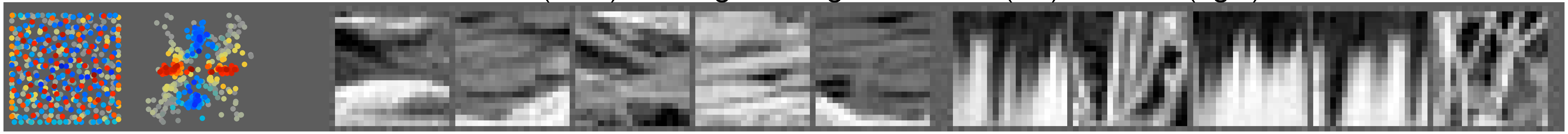


a similar function from Karklin and Lewicki (2003) showing the images that most (left) and least (right) activated v

another example of the same type

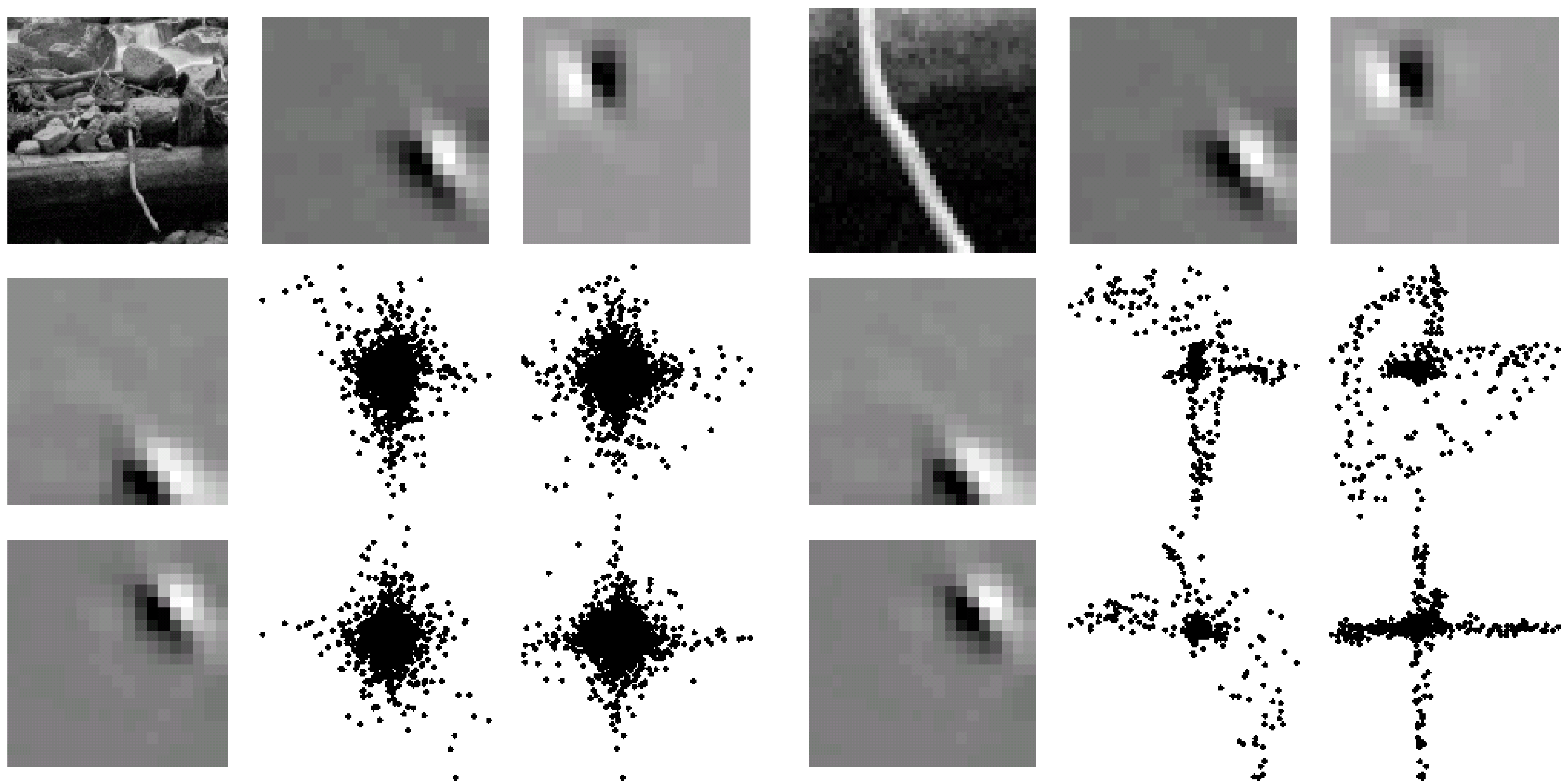


a similar function from Karklin and Lewicki (2003) showing the images that most (left) and least (right) activated v



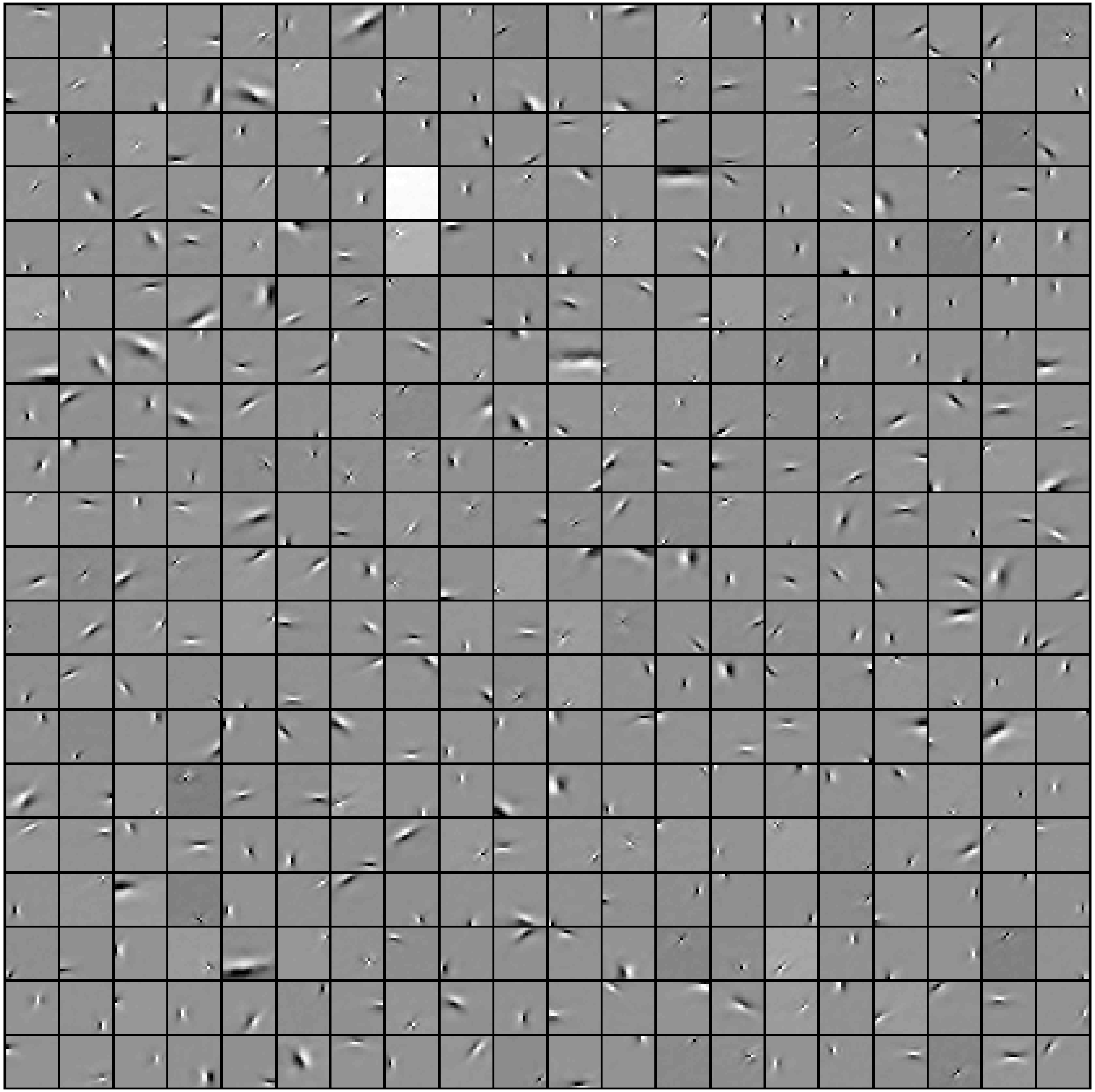
Learning more image structure and higher-order dependencies

Joint activation patterns of u's over different image patches

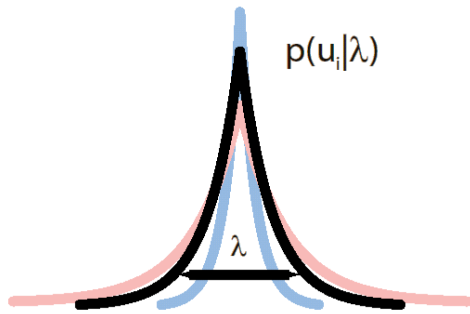


How do we extend the model to learn these types of regularities?

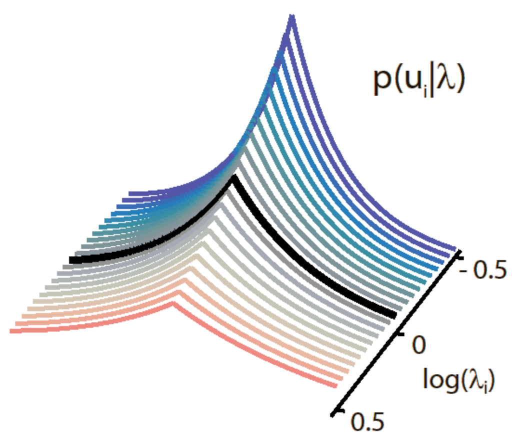
Edge transform demo

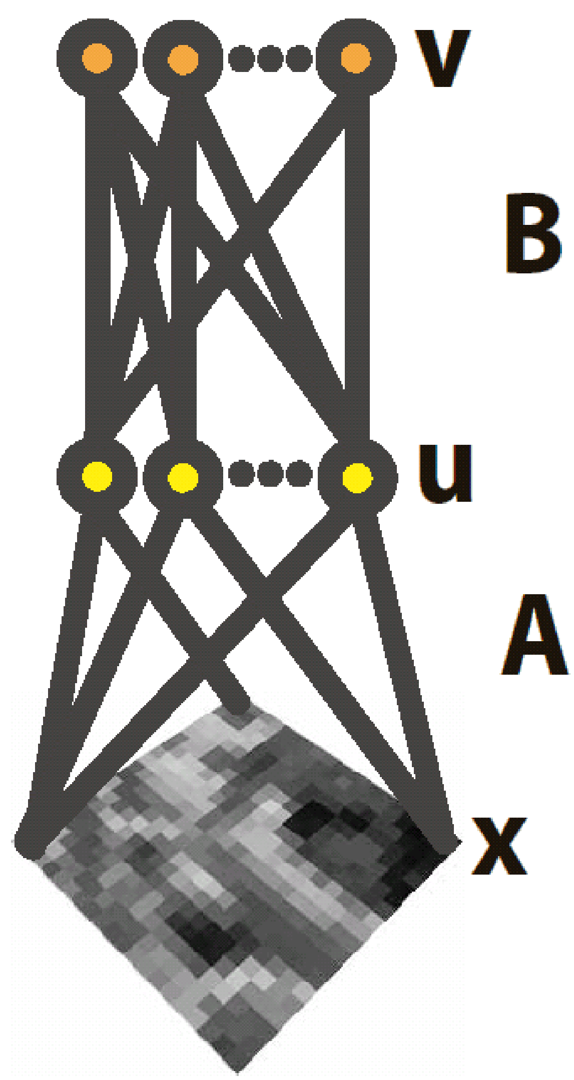


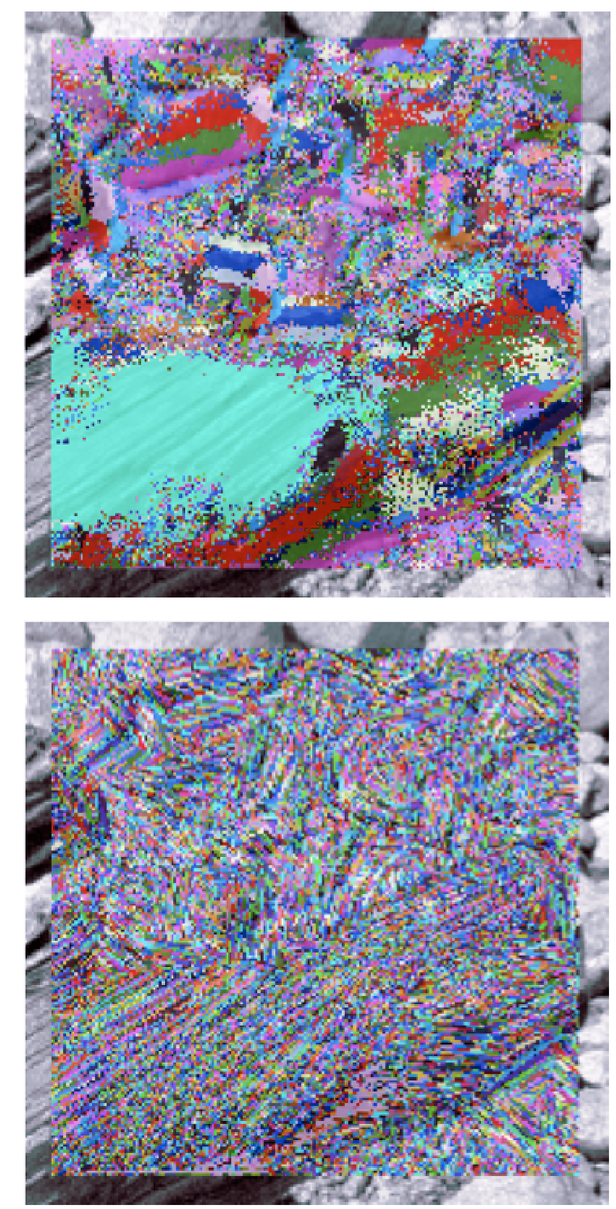
Summary

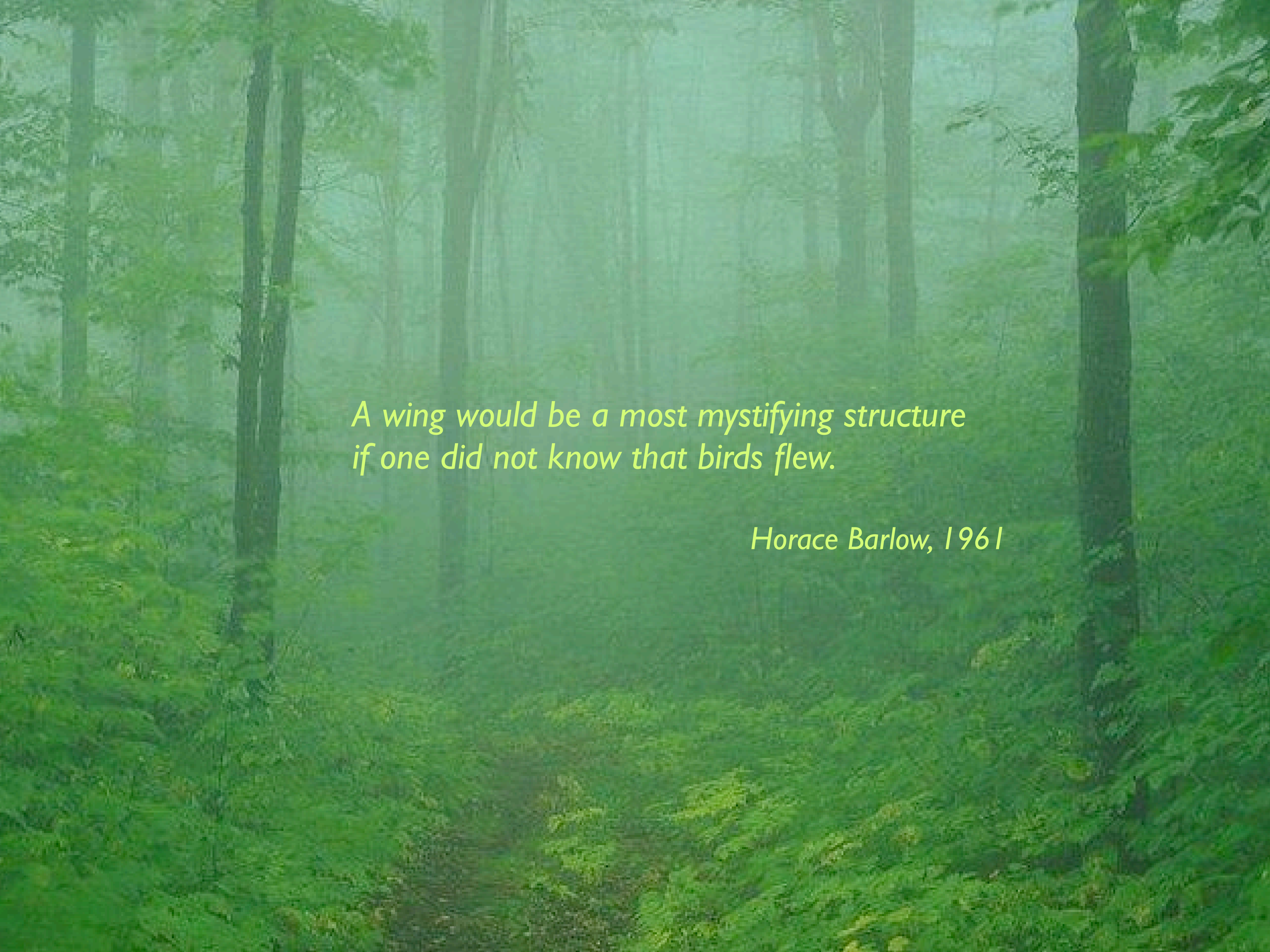


$$\log p(\mathbf{u}|\mathbf{B},\mathbf{v}) \propto -\sum \left|\frac{u_i}{e^{[\mathbf{Bv}]_i}}\right|^{q_i}$$








A wing would be a most mystifying structure
if one did not know that birds flew.
Horace Barlow, 1961