Analog Memory and High-Dimensional Computation

Bruno Olshausen

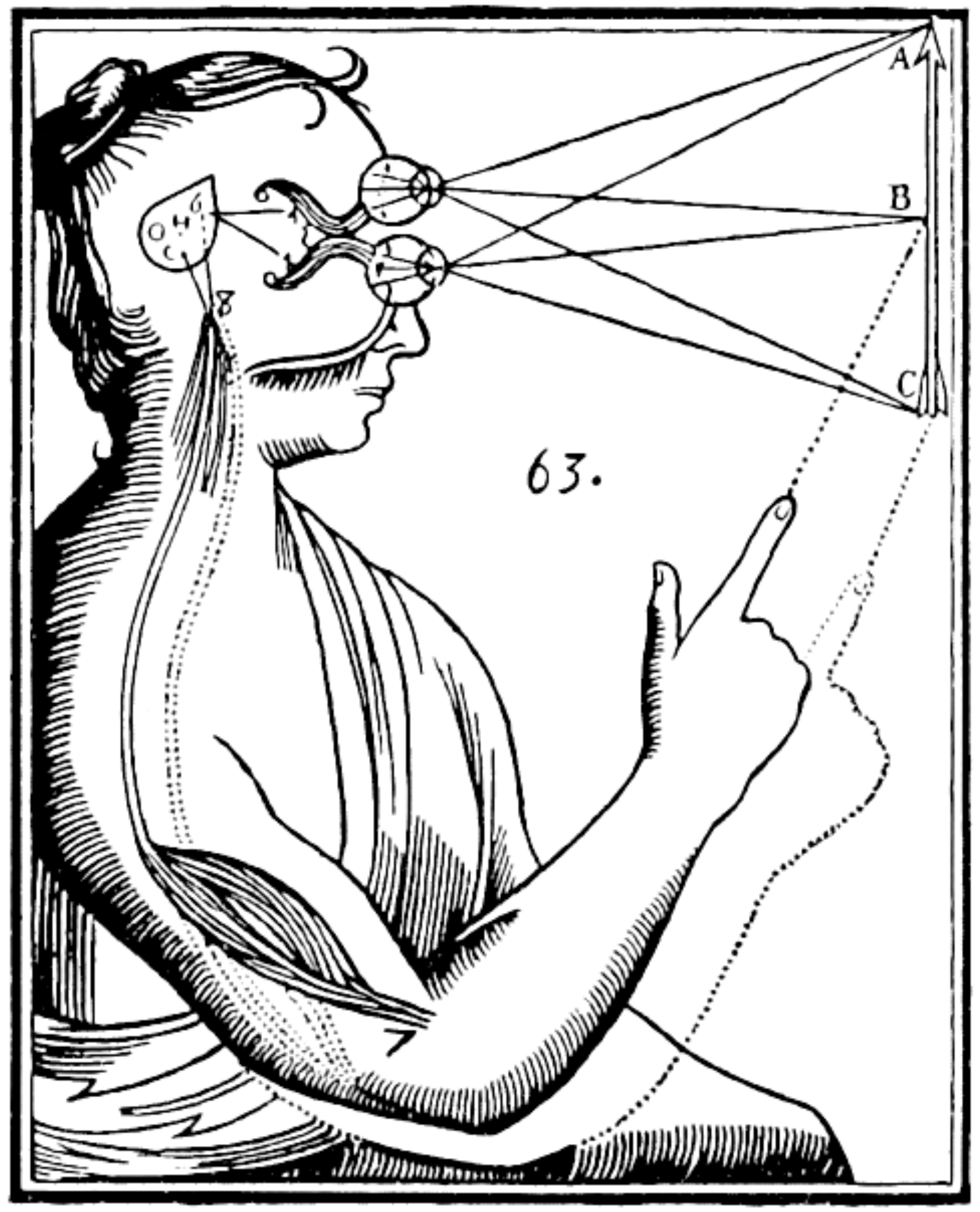
Ryan Zarcone, Jesse Engel

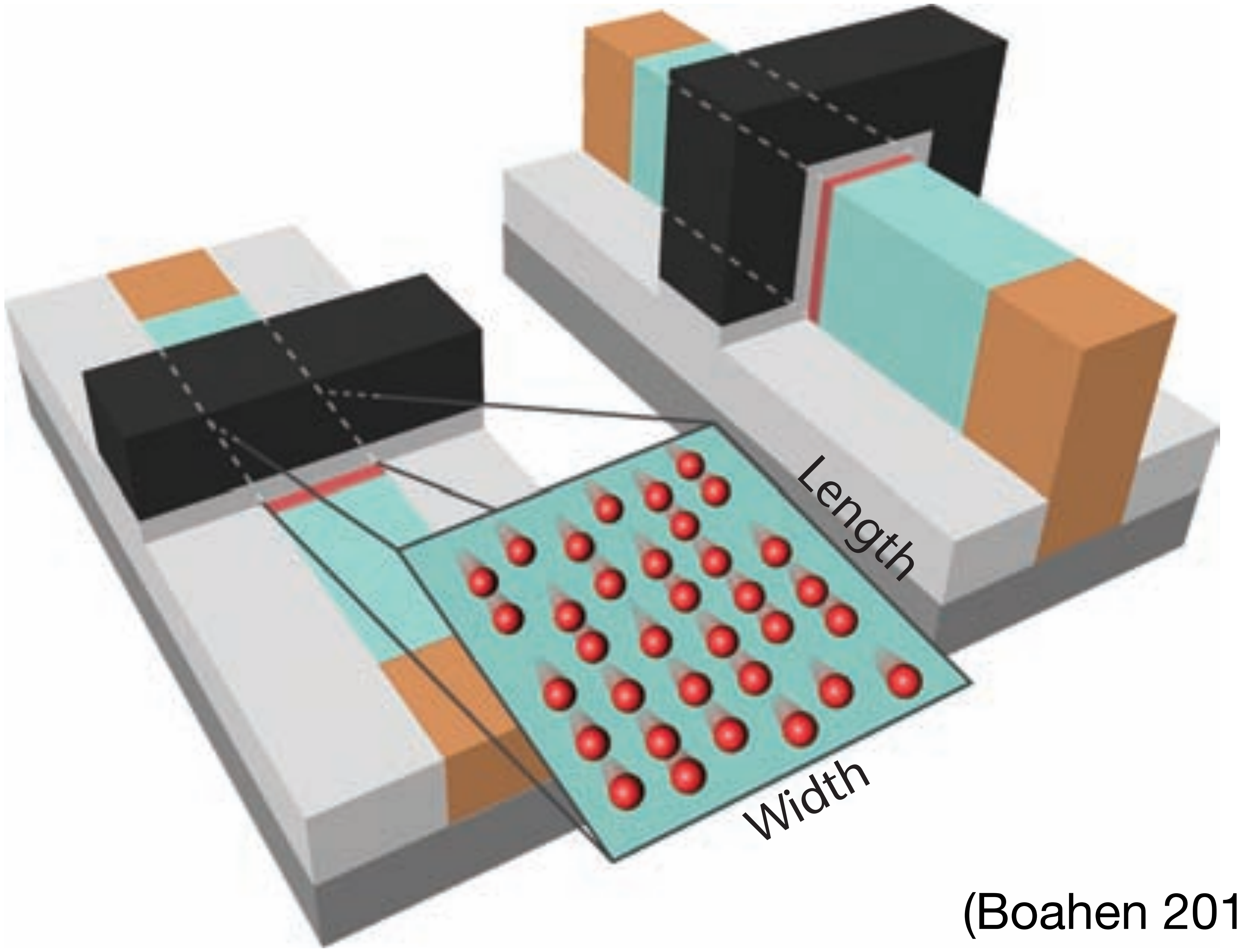
Spencer Kent, E. Paxon Frady
Fritz Sommer

Xin Zheng, Joon Sohn, Weier Wan
Philip Wong
Dept. of Electrical Engineering, Stanford University

REDWOOD CENTER
for Theoretical Neuroscience

Brains vs. machines

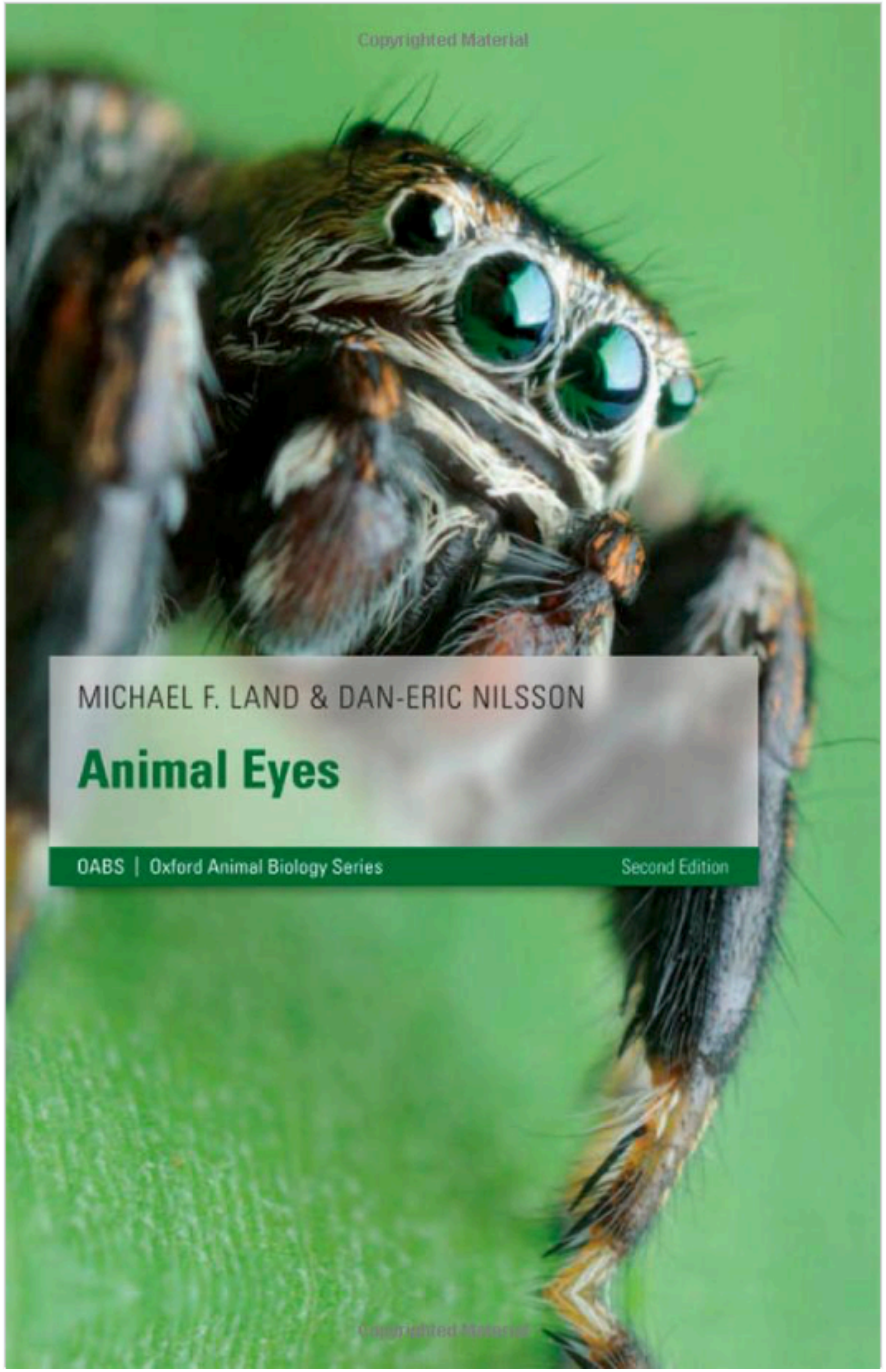


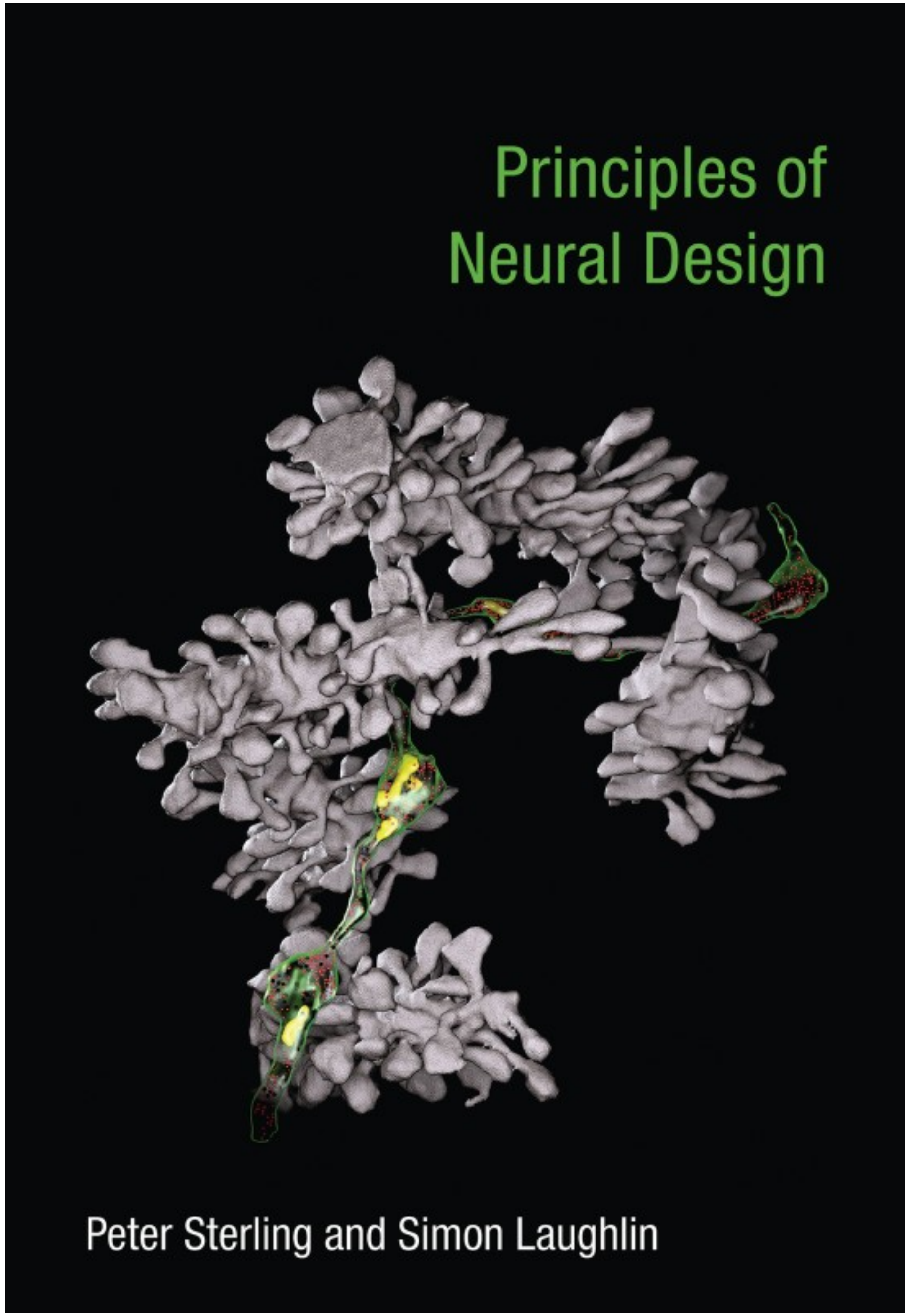


(Boahen 2017)

Brain-like functions are more probabilistic in nature and use different data representations.

How to compute with nanoscale, low-power, *stochastic* circuit components?


Copyrighted Material
MICHAEL F. LAND & DAN-ERIC NILSSON
Animal Eyes
OABS | Oxford Animal Biology Series
Second Edition


Principles of Neural Design
Peter Sterling and Simon Laughlin

Two questions

1. How to realize the potential of low-power, compact phase-change memory (PCM) and Resistive RAM (RRAM) crossbar arrays for *analog data storage*?

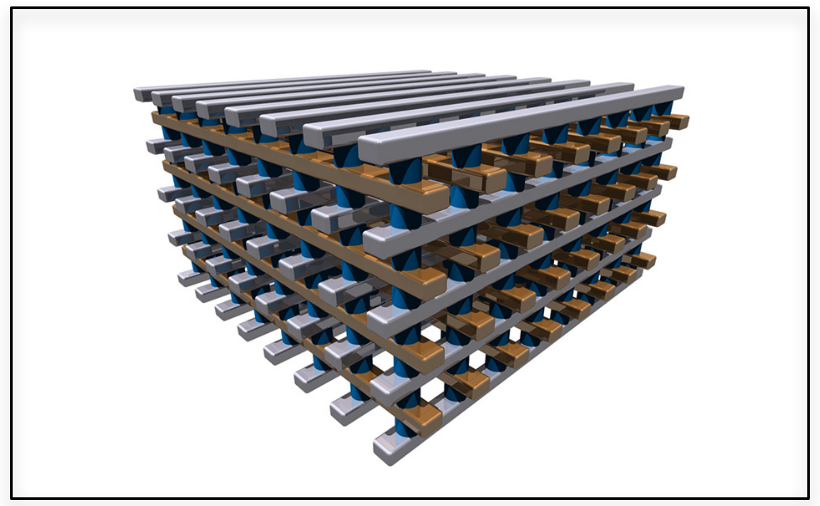
2. How to compute *holistically* with large populations of neurons - i.e., with high-dimensional data representations?

Adaptive Error-Correcting Codes for Analog Data Storage in PCM/RRAM

source

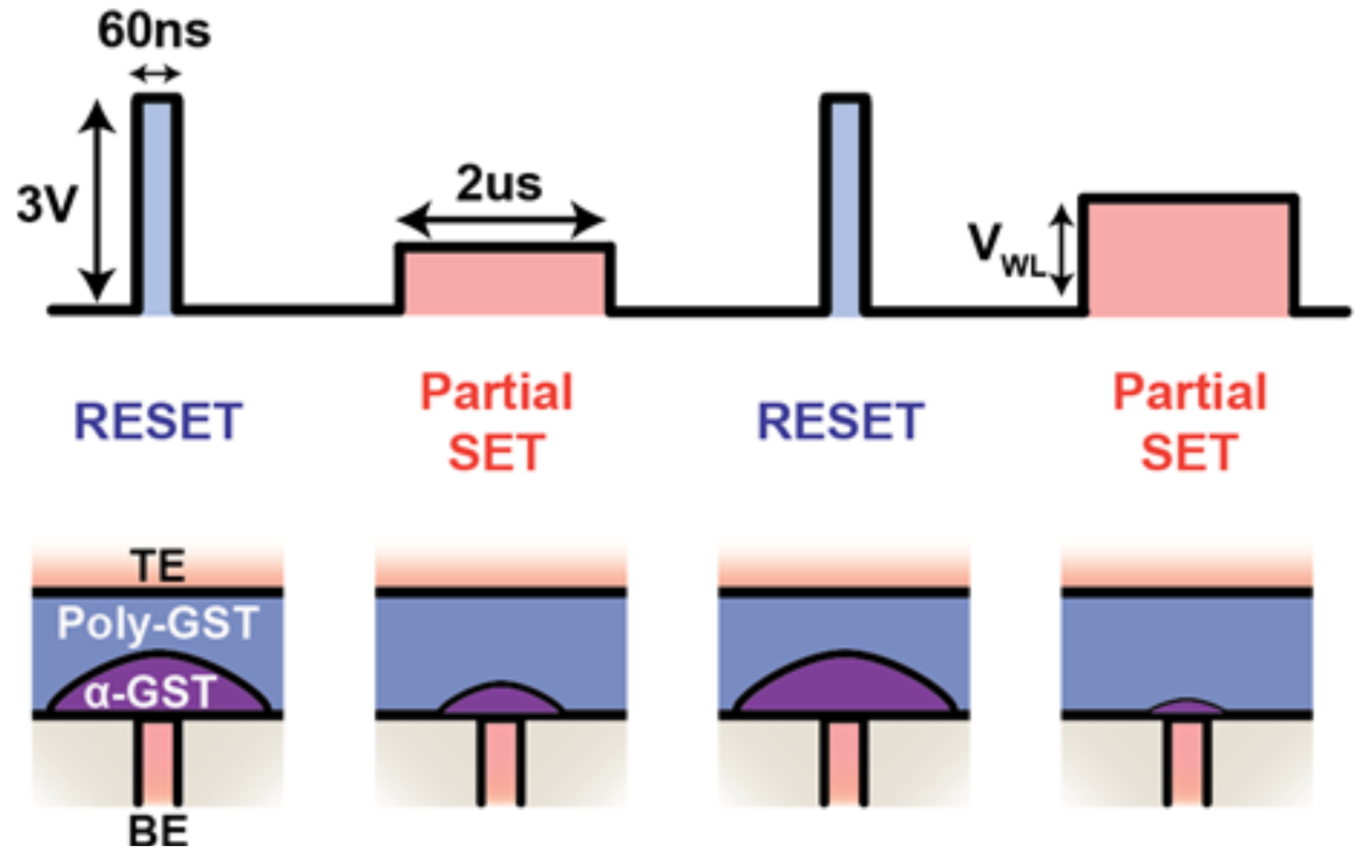
encode

channel

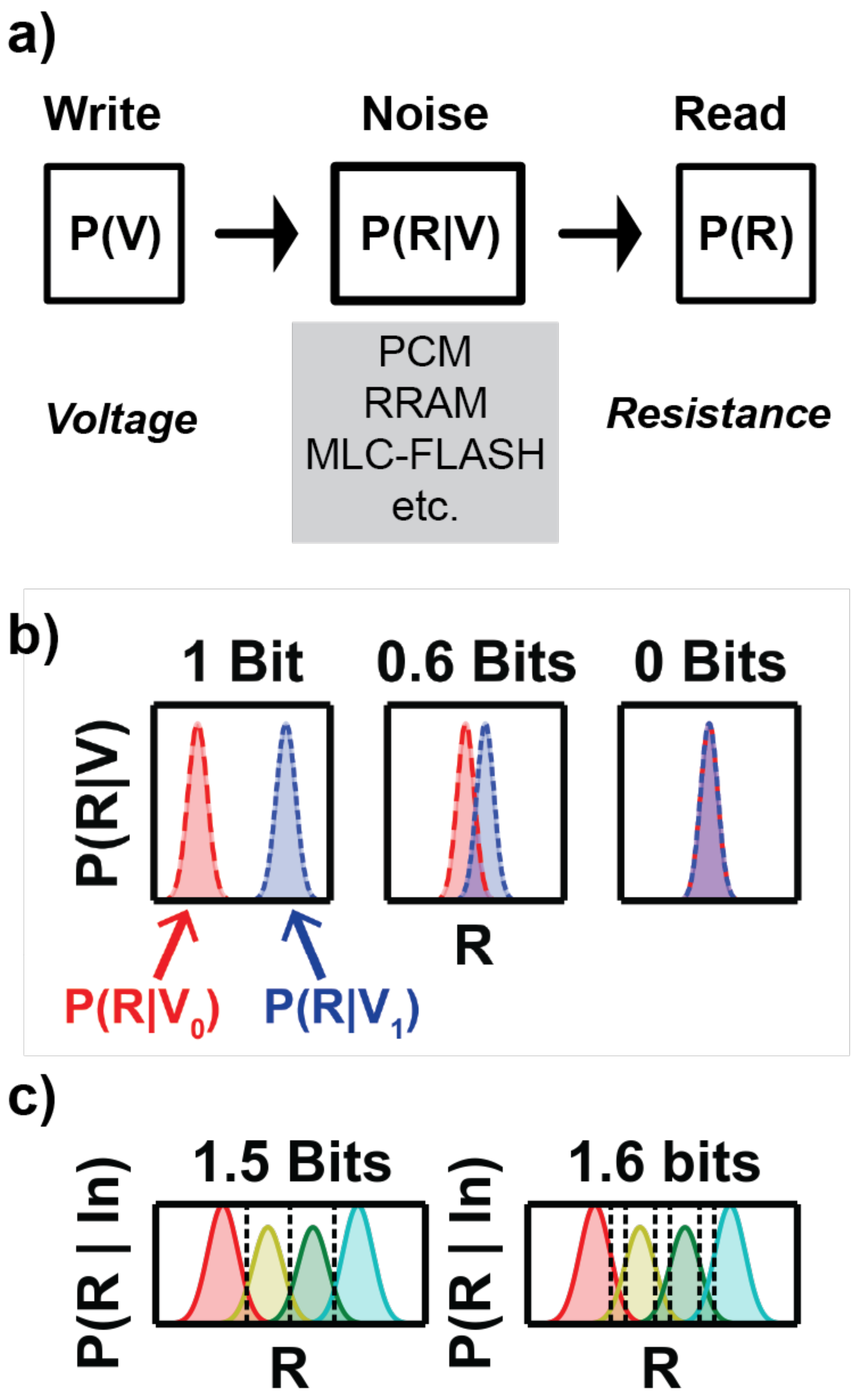


decode

receiver



Analog memory as a noisy channel

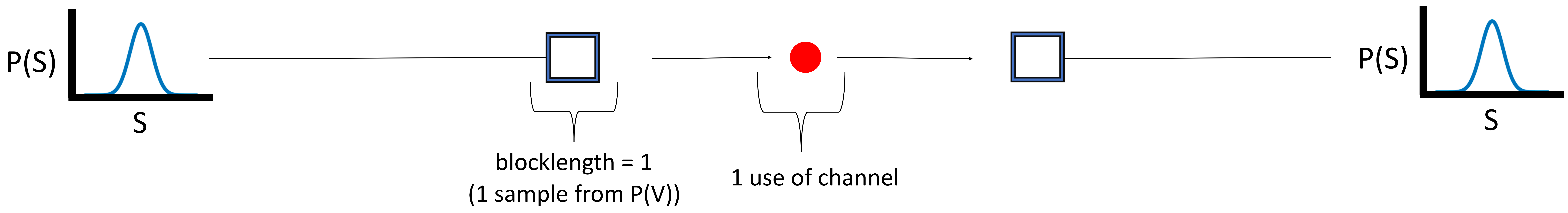


P(R|V) determines capacity

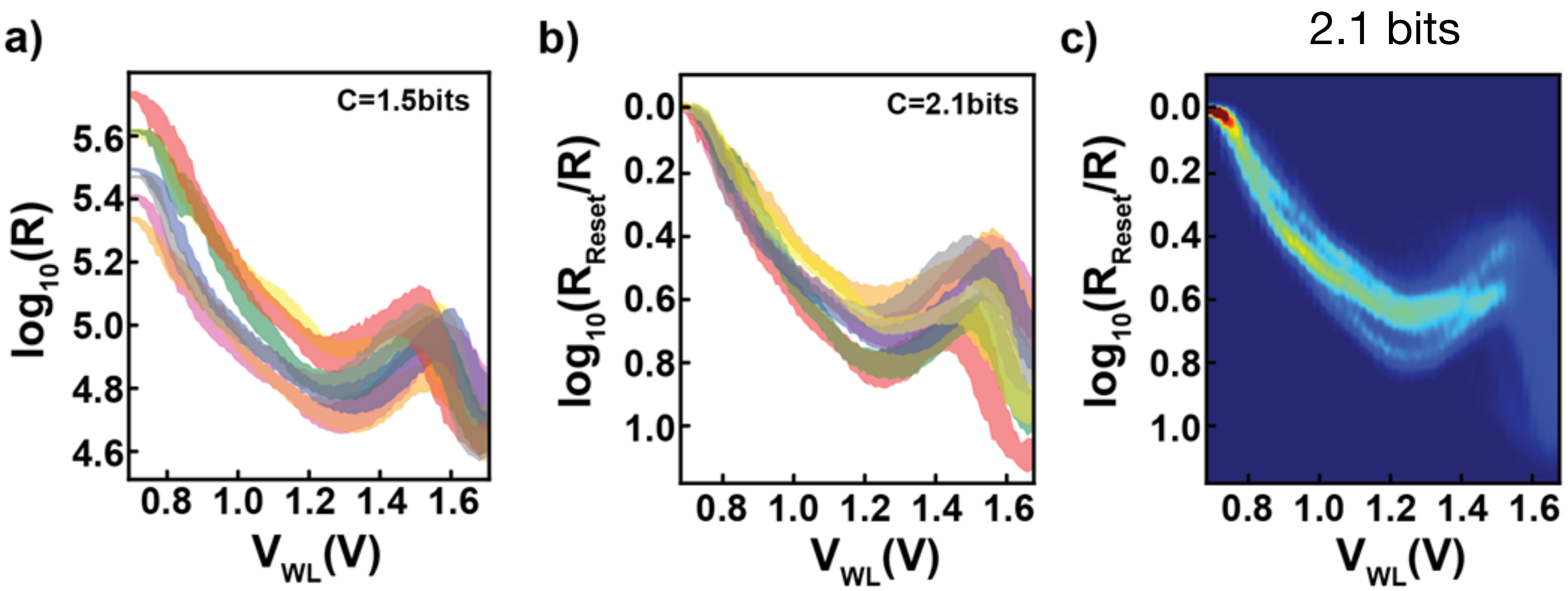
$$C = \max_{P(V)} \sum_{V,R} P(V)\, P(R|V)\, \log_2 \frac{P(R|V)}{P(R)}$$

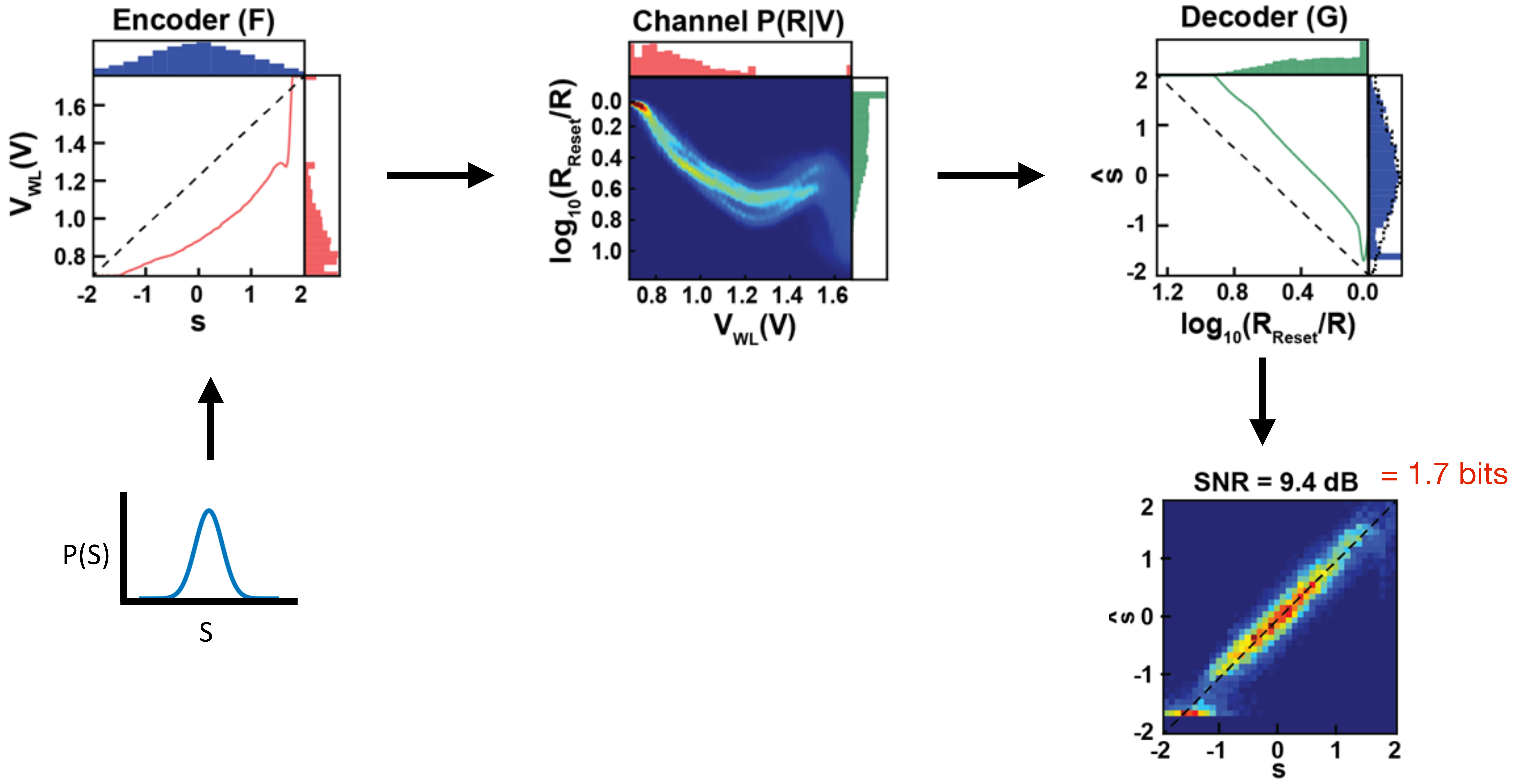
Separate Source-Channel Coding

Joint Source-Channel Coding



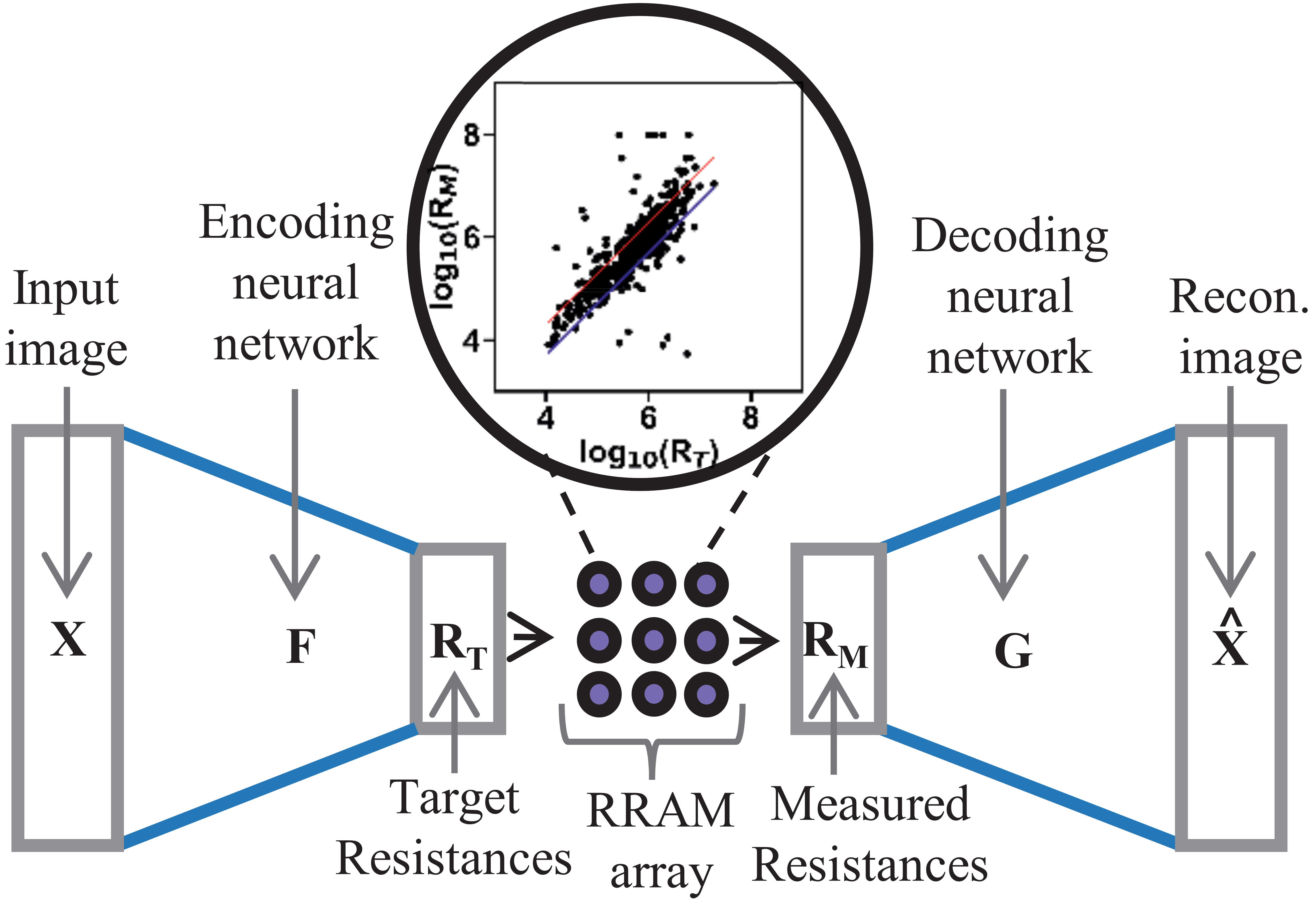
P(R|V) for seven devices on a PCM array





Encoder (F)
Channel P(R|V)
Decoder (G)
V_{WL}(V)
s
$\log_{10}(R_{Reset}/R)$
V_{WL}(V)
$\hat{s}$
$\log_{10}(R_{Reset}/R)$
P(S)
S
SNR = 9.4 dB
= 1.7 bits
$\hat{s}$
s

Autoencoder framework for multidimensional signals (images)



$$\min_{F,G} \left\langle [\mathbf{x} - \hat{\mathbf{x}}]^2 \right\rangle$$

(Zheng, Zarcone, Paiton, Sohn, Wan, Olshausen & Wong, IEDM 2018)

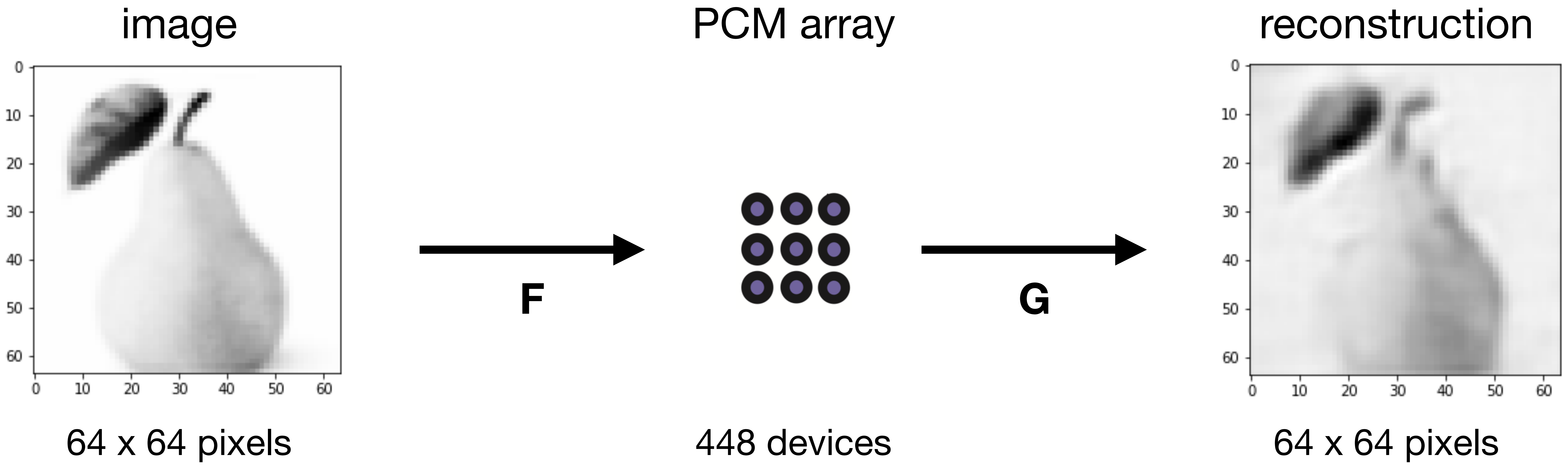
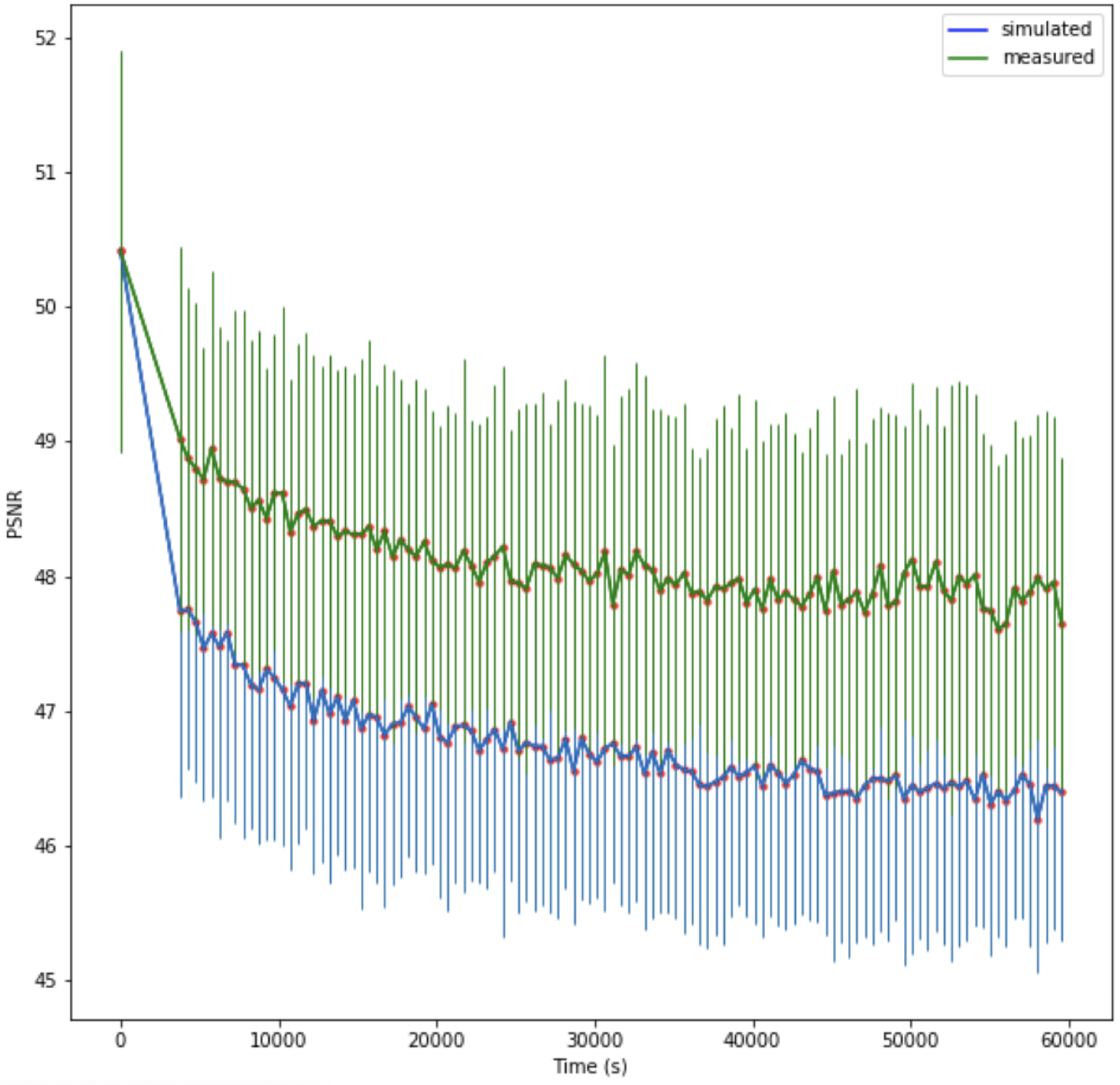


image
PCM array
reconstruction
F
G
64 x 64 pixels
448 devices
64 x 64 pixels

Effect of device drift on image reconstruction

17 dB →



Computing with high-dimensional representations

Single neuron recording ⇒ Single neuron thinking

1940 PROCEEDINGS OF THE IRE November

What the Frog's Eye Tells the Frog's Brain*

J. Y. LETTVIN†, H. R. MATURANA‡, W. S. McCULLOCH||, SENIOR MEMBER, IRE, AND W. H. PITTS||

Perception, 1972, volume 1, pages 371–394

67

Single units and sensation: A neuron doctrine for perceptual psychology?

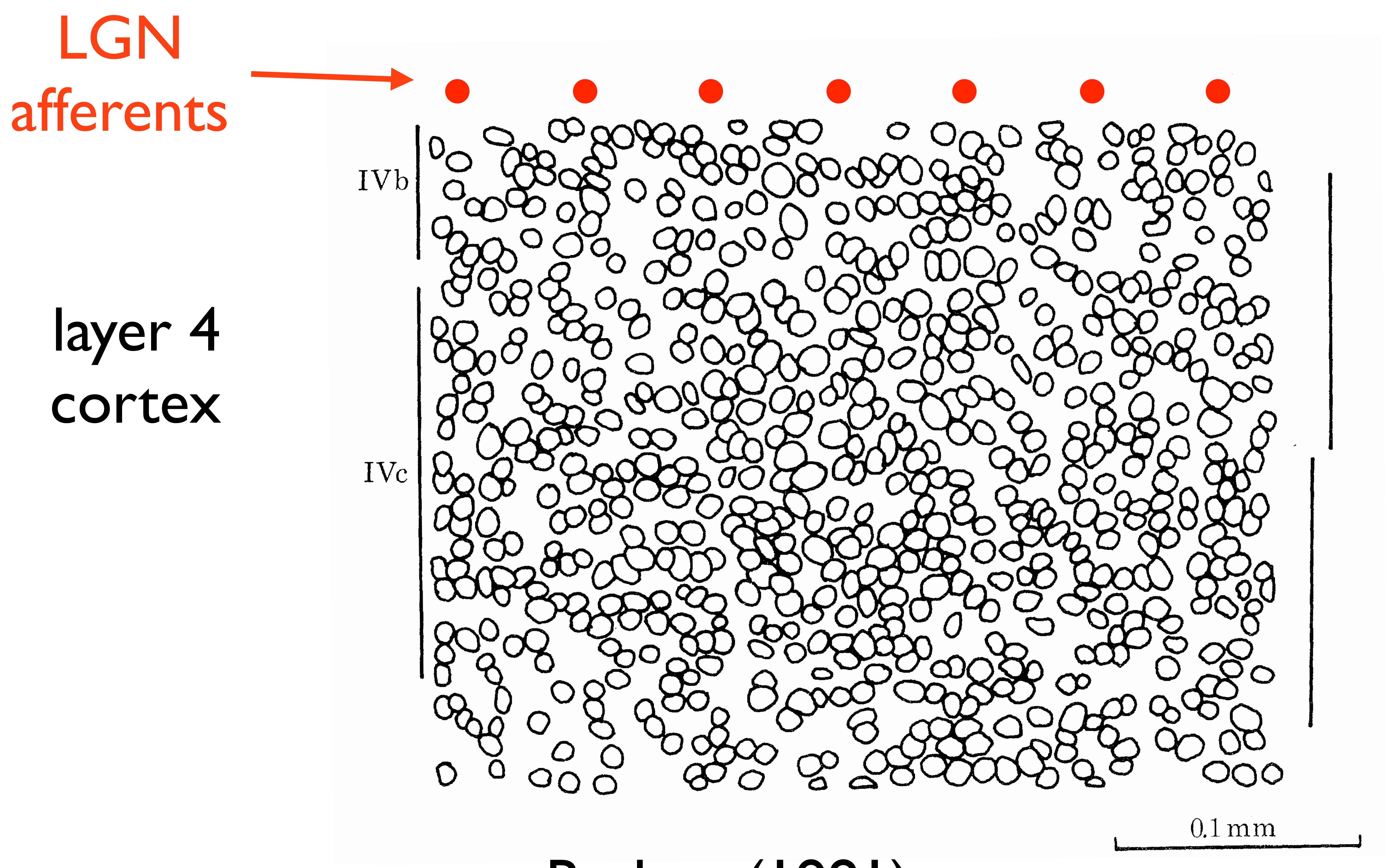
H B Barlow
Department of Physiology–Anatomy, University of California, Berkeley, California 94720
Received 6 December 1972

Abstract. The problem discussed is the relationship between the firing of single neurons in sensory pathways and subjectively experienced sensations. The conclusions are formulated as the following five dogmas:

1. To understand nervous function one needs to look at interactions at a cellular level, rather than either a more macroscopic or microscopic level, because behaviour depends upon the organized pattern of these intercellular interactions.
2. The sensory system is organized to achieve as complete a representation of the sensory stimulus as possible with the minimum number of active neurons.
3. Trigger features of sensory neurons are matched to redundant patterns of stimulation by experience as well as by developmental processes.
4. Perception corresponds to the activity of a small selection from the very numerous high-level neurons, each of which corresponds to a pattern of external events of the order of complexity of the events symbolized by a word.
5. High impulse frequency in such neurons corresponds to high certainty that the trigger feature is present.

The development of the concepts leading up to these speculative dogmas, their experimental basis, and some of their limitations are discussed.

The brain's circuits are high-dimensional



Barlow (1981)

Computing with high-dimensional vectors

Pentti Kanerva

Concepts, variables, attributes are represented as high-dimensional vectors (e.g., 10,000 bits)

Three fundamental operations:

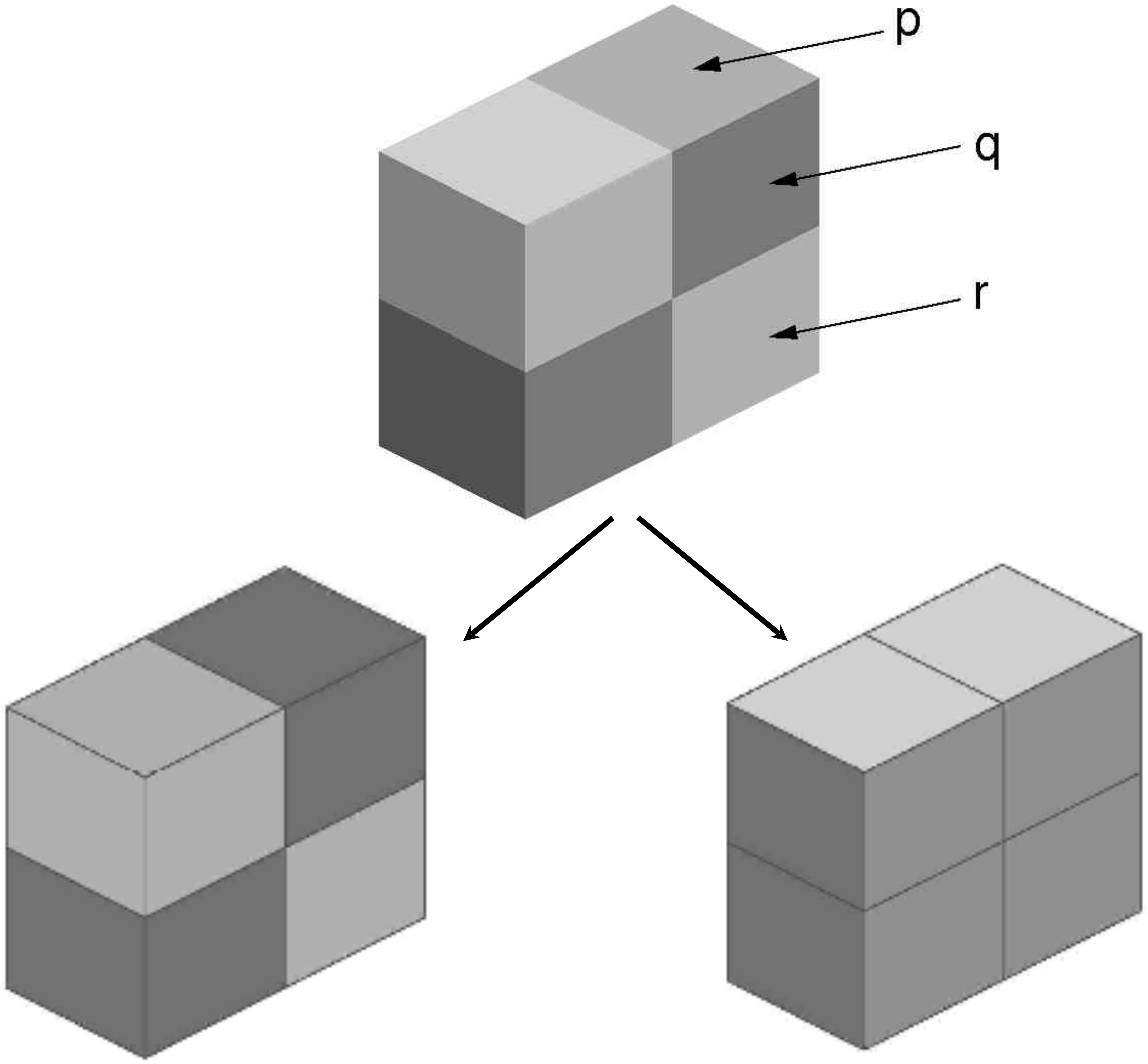
- multiplication (binding)
- addition (combining)
- permutation (sequencing)

Approximates a *field*

Kanerva P (2009) Hyperdimensional Computing: An Introduction to Computing in Distributed Representation with High-Dimensional Random Vectors. *Cognitive Computing, 1*: 139-159.

Plate, T. A. (1995). Holographic reduced representations. *IEEE Transactions on Neural networks*, *6*(3), 623-641.

Factorization of shape and reflectance



reflectance

shading

(Adelson, 2000)

We approach this problem within the framework of High-Dimensional (HD) Computing:

- Visual scene attributes such as *position*, *shape* or *color* are represented as HD vectors.
- An image is encoded into a HD vector so that it expresses a *product* of these attributes.
- The problem of scene analysis amounts to *factorizing* an HD scene vector into its attributes.
- A scene containing multiple objects may be expressed as a *superposition* of products.

Factorization in HD

Let $\mathbf{b} = \mathbf{x} \otimes \mathbf{y} \otimes \mathbf{z}$

$$\mathbf{x} \in \mathbb{X} := \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n\}$$
$$\mathbf{y} \in \mathbb{Y} := \{\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_n\}$$
$$\mathbf{z} \in \mathbb{Z} := \{\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_n\}$$

Problem: You are given $\mathbf{b}$, what are $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$?

Solution: Resonate

$$\hat{\mathbf{x}}_{t+1} = g\big(\mathbf{X}\mathbf{X}^\top(\mathbf{b} \otimes \hat{\mathbf{y}}_t^{-1} \otimes \hat{\mathbf{z}}_t^{-1})\big)$$
$$\hat{\mathbf{y}}_{t+1} = g\big(\mathbf{Y}\mathbf{Y}^\top(\mathbf{b} \otimes \hat{\mathbf{x}}_t^{-1} \otimes \hat{\mathbf{z}}_t^{-1})\big)$$
$$\hat{\mathbf{z}}_{t+1} = g\big(\mathbf{Z}\mathbf{Z}^\top(\mathbf{b} \otimes \hat{\mathbf{x}}_t^{-1} \otimes \hat{\mathbf{y}}_t^{-1})\big)$$

$$\mathbf{X} = \begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_n \\ | & | & & | \end{bmatrix}$$
$$\mathbf{Y} = \begin{bmatrix} | & | & & | \\ \mathbf{y}_1 & \mathbf{y}_2 & \ldots & \mathbf{y}_n \\ | & | & & | \end{bmatrix}$$
$$\mathbf{Z} = \begin{bmatrix} | & | & & | \\ \mathbf{z}_1 & \mathbf{z}_2 & \ldots & \mathbf{z}_n \\ | & | & & | \end{bmatrix}$$

$$g(x) = \mathrm{sgn}(x)$$

Consider the following energy function

$$E = -\mathbf{b} \cdot \overbrace{(\mathbf{x} \otimes \mathbf{y} \otimes \mathbf{z})}^{(\alpha_1\beta_1\gamma_1\,\mathbf{x}_1 \otimes \mathbf{y}_1 \otimes \mathbf{z}_1 \,+\, \ldots \,+\, \alpha_i\beta_j\gamma_k\,\mathbf{x}_i \otimes \mathbf{y}_j \otimes \mathbf{z}_k \,+\, \ldots \,+\, \alpha_n\beta_n\gamma_n\,\mathbf{x}_n \otimes \mathbf{y}_n \otimes \mathbf{z}_n)}$$

$$\mathbf{x} = \sum_{i=1}^{n} \alpha_i\,\mathbf{x}_i, \quad \mathbf{y} = \sum_{i=1}^{n} \beta_i\,\mathbf{y}_i, \quad \mathbf{z} = \sum_{i=1}^{n} \gamma_i\,\mathbf{z}_i$$

Consider the following energy function

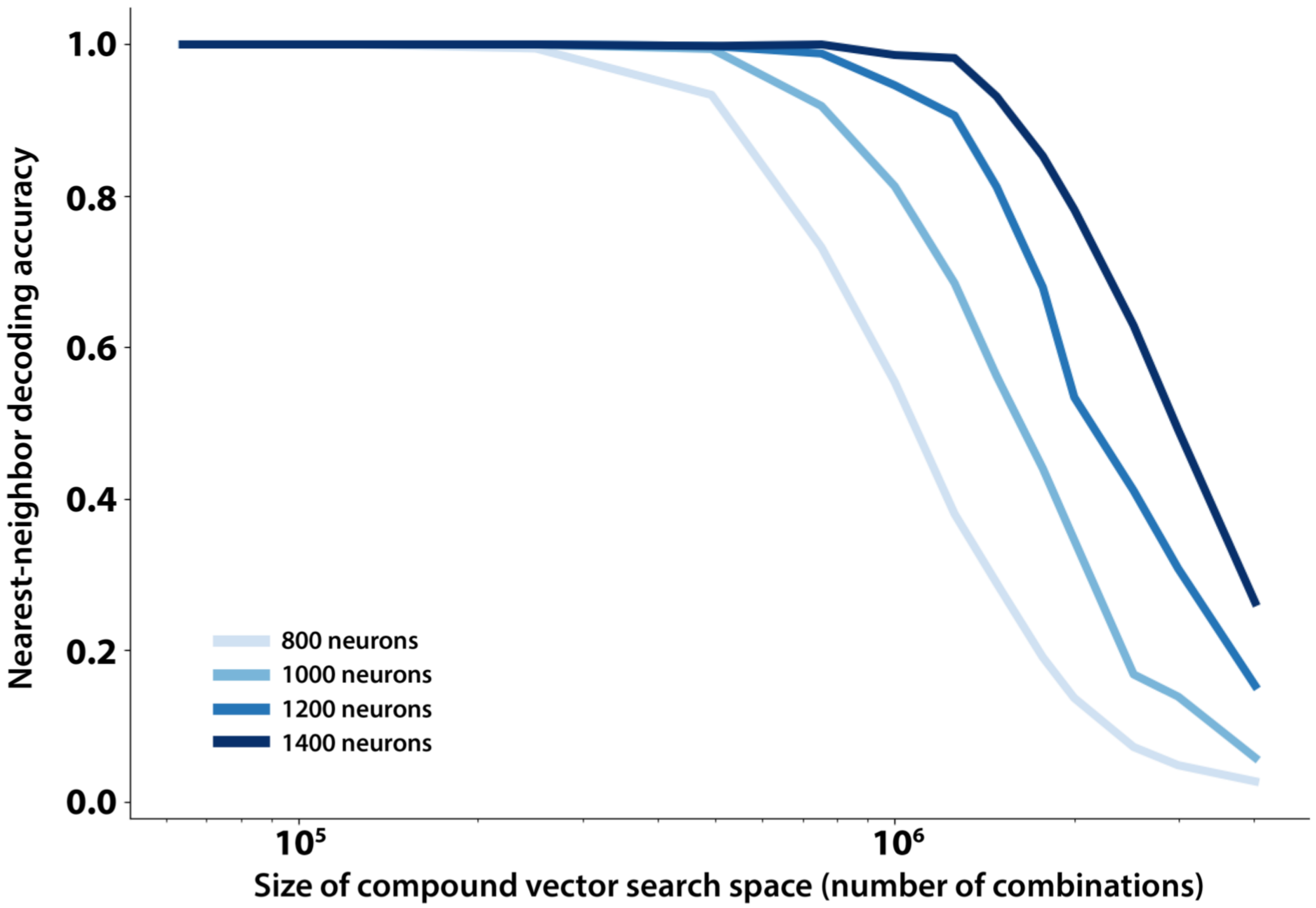
1,000,000 combinations! (n=100)

$$(\alpha_1\beta_1\gamma_1\,\mathbf{x}_1 \otimes \mathbf{y}_1 \otimes \mathbf{z}_1 \; + \; \ldots \; + \; \alpha_i\beta_j\gamma_k\,\mathbf{x}_i \otimes \mathbf{y}_j \otimes \mathbf{z}_k \; + \; \ldots \; + \; \alpha_n\beta_n\gamma_n\,\mathbf{x}_n \otimes \mathbf{y}_n \otimes \mathbf{z}_n)$$

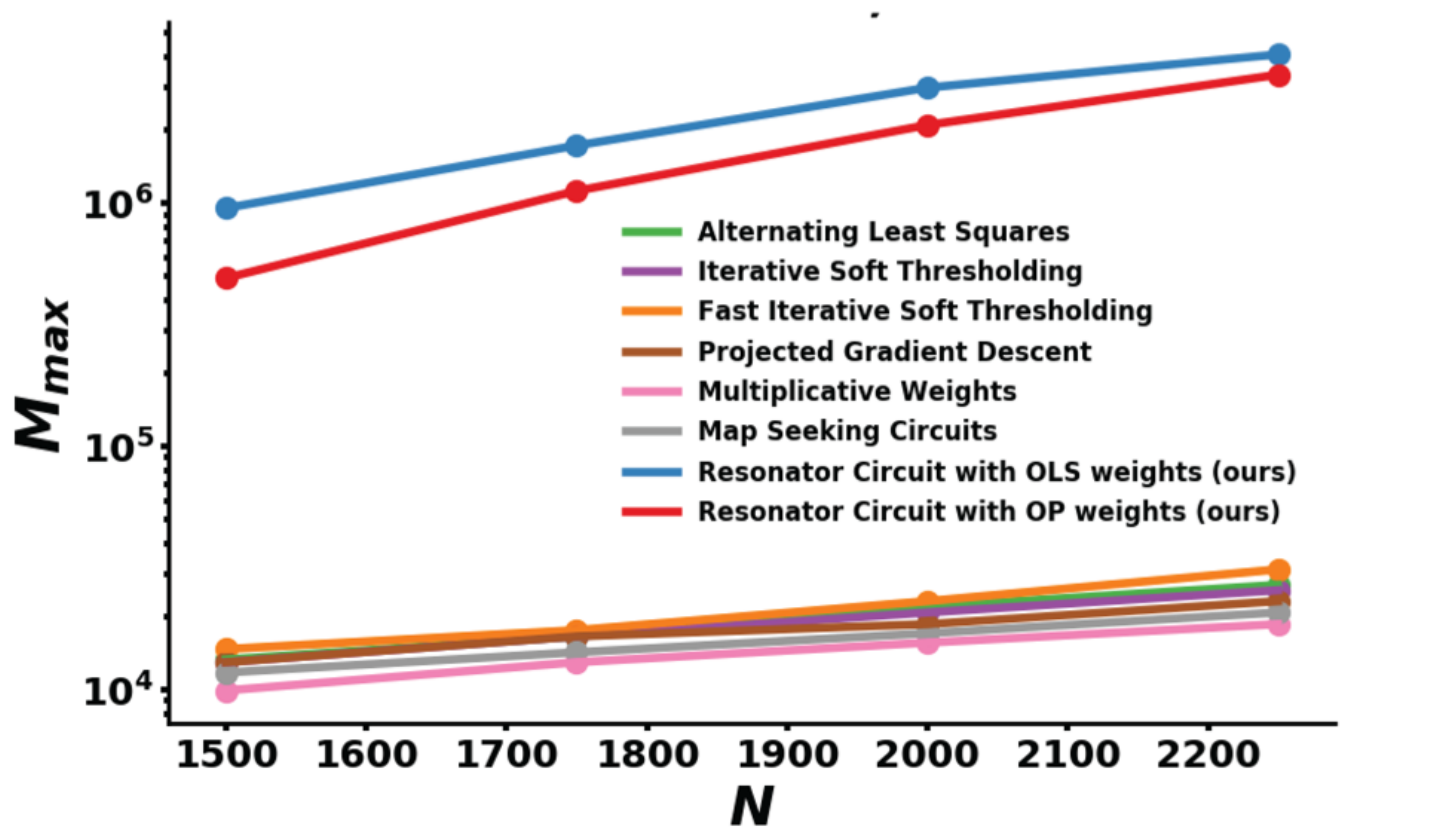
$$E = -\mathbf{b} \cdot (\mathbf{x} \otimes \mathbf{y} \otimes \mathbf{z})$$

$$\mathbf{x} = \sum_{i=1}^{n} \alpha_i\,\mathbf{x}_i, \quad \mathbf{y} = \sum_{i=1}^{n} \beta_i\,\mathbf{y}_i, \quad \mathbf{z} = \sum_{i=1}^{n} \gamma_i\,\mathbf{z}_i$$

Search capacity increases with number of dimensions



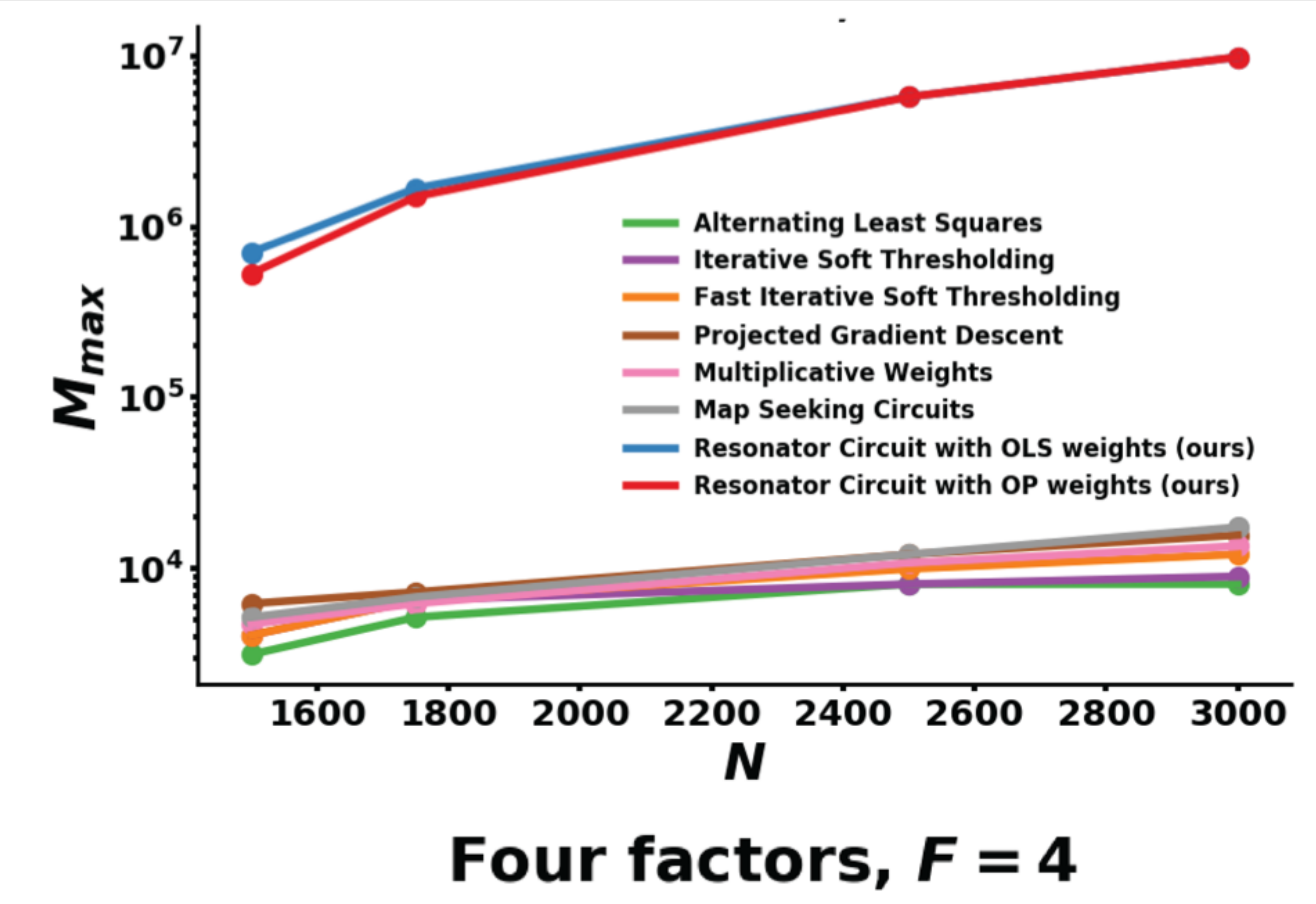
Operational capacity far exceeds gradient-based and other standard optimization methods
(Spencer Kent)



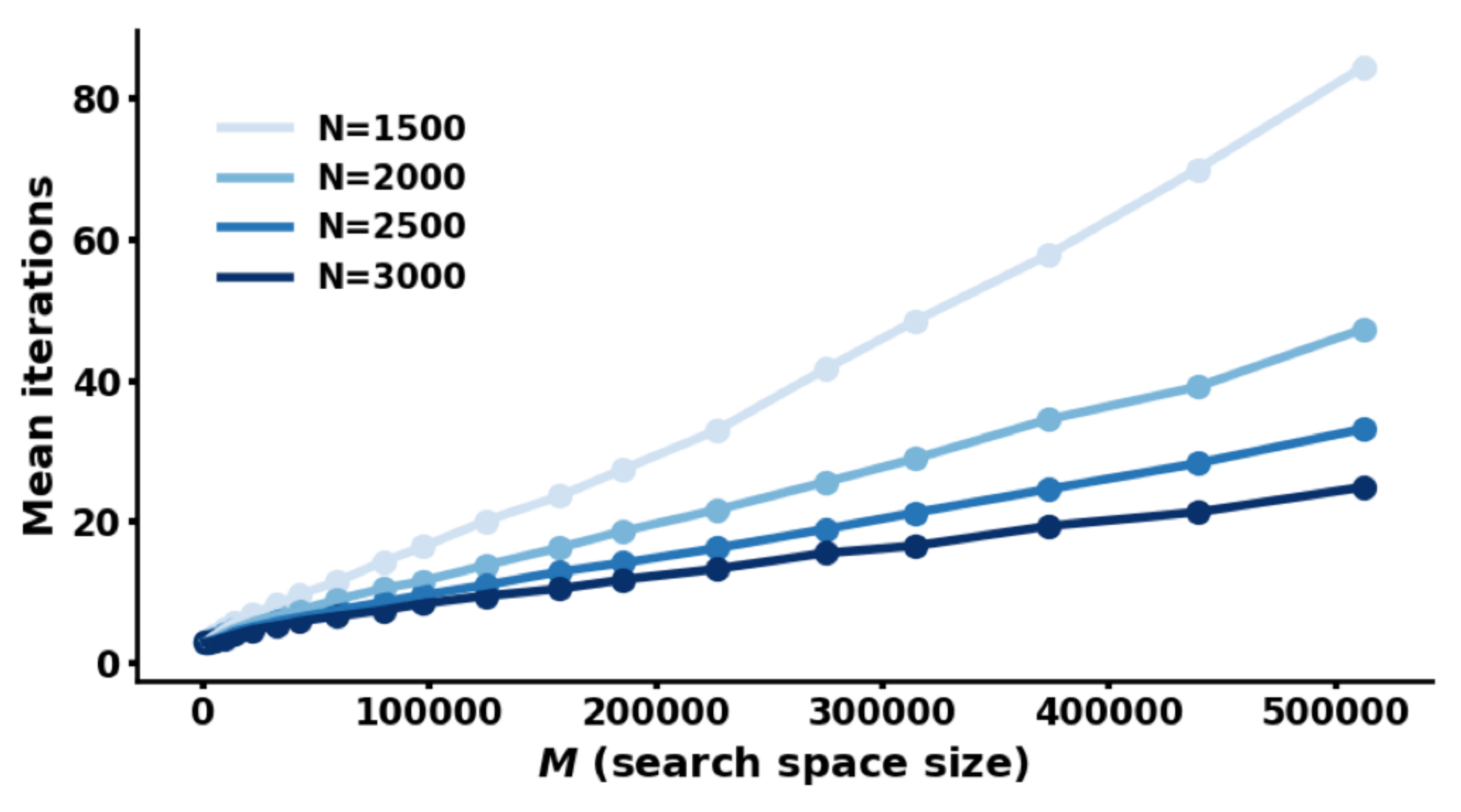
Three factors, $F = 3$

Operational capacity far exceeds gradient-based and other standard optimization methods

(Spencer Kent)



Search efficiency



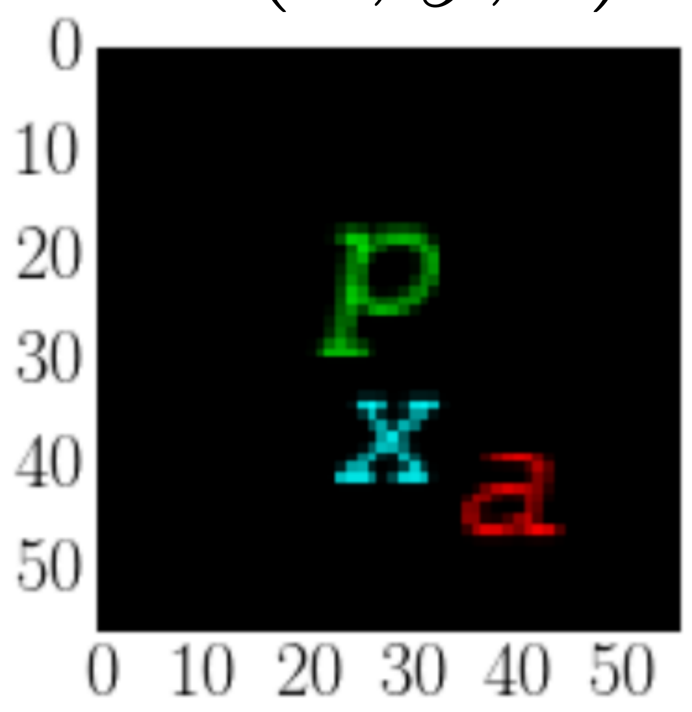
Visual scene analysis via factorization of HD vectors
(Paxon Frady)

$\mathbf{U}^{x_i}$ = horizontal position x_i

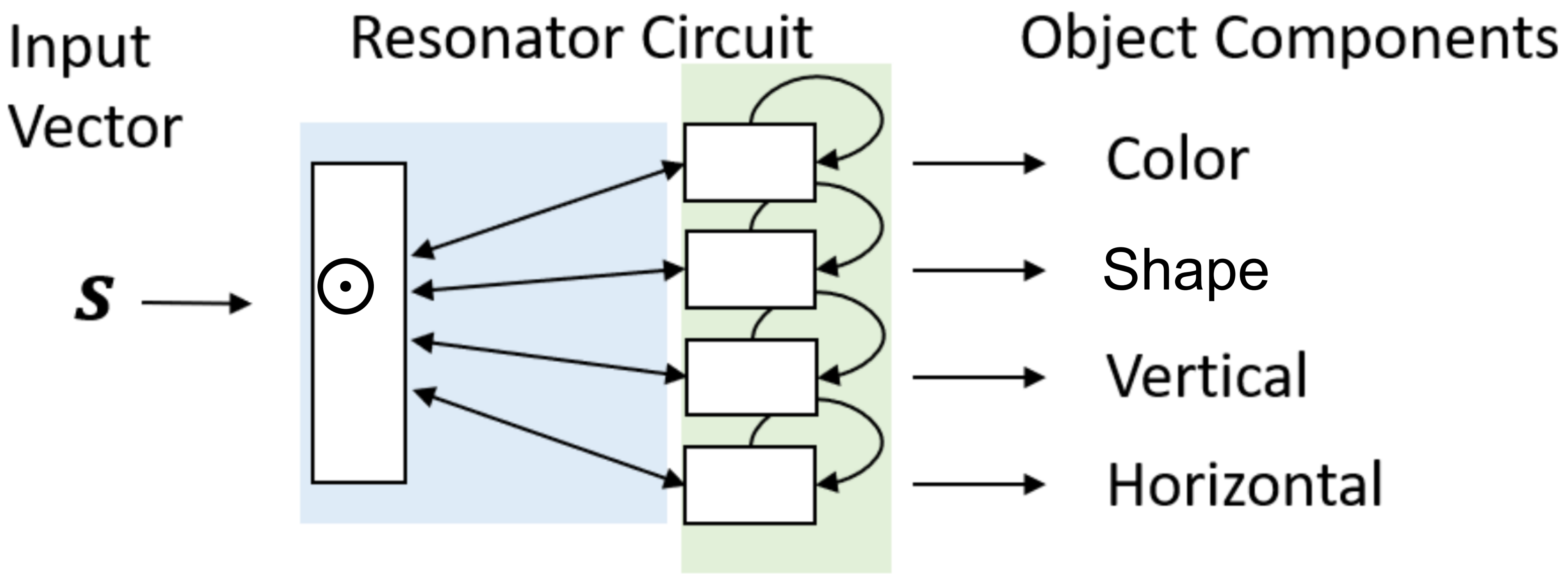
$\mathbf{V}^{y_j}$ = vertical position y_j

$\mathbf{W}_c$ = color channel c

$I(x, y, c)$

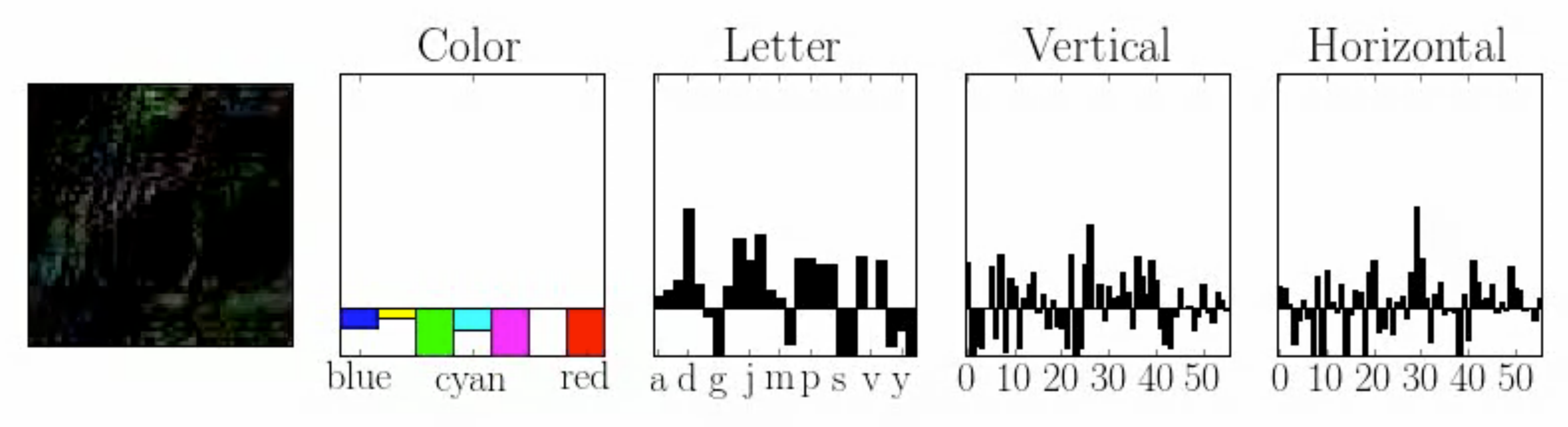


$$\longrightarrow \quad \mathbf{s} = \sum_{i,j,c} I(x_i, y_j, c)\, \mathbf{U}^{x_i}\, \mathbf{V}^{y_j}\, \mathbf{W}_c$$



Visual scene analysis via factorization of HD vectors

(Paxon Frady)



Main points

- A common set of design principles may be used to understand brains and to engineer intelligent machines:
  - probabilistic memory and computation
  - holistic representation and computation

- Emerging memory (PCM/RRAM) may be most efficiently utilized as analog devices for storing analog-valued data.

- High-dimensional representation combined with an algebra of operators opens the door to combine and factorize data representations in new ways that enable us to solve problems in a manner that is not only tractable but also robust.