Neuronal learning rules

- Skewness 1 Deviation from symmetry:

$$S_1 = E[c^3]/E^{1.5}[c^2].$$

$$\begin{aligned} \nabla S_1 &= \frac{1}{{\theta_M}^{1.5}} E\left[c\left(c - E[c^3]/E[c^2]\right)\sigma'\mathbf{d}\right] \\ &= \frac{1}{{\theta_M}^{1.5}} E\left[c\left(c - E[c^3]/\theta_M\right)\sigma'\mathbf{d}\right] \end{aligned}$$

where Θ_m is defined as $E[c^2]$.

- Skewness 2 (additive):

$$S_2 = E[c^3] - E^{1.5}[c^2].$$

$$\begin{aligned} \nabla S_2 &= 3E\left[c^2 - c\sqrt{E[c^2]}\right] \\ &= 3E\left[c\left(c - \sqrt{\theta_M}\right)\sigma'\mathbf{d}\right], \end{aligned}$$

subject to the constraint $\| \mathbf{m} \| = 1$.

- Kurtosis 1 Emphasizes the tails:

$$K_1 = E[c^4]/E^2[c^2] - 3.$$

$$\begin{aligned}\nabla K_1 &= \frac{1}{{\theta_M}^2} E\left[c\left(c^2 - E[c^4]/E[c^2]\right)\sigma'\mathbf{d}\right] \\ &= \frac{1}{{\theta_M}^2} E\left[c\left(c^2 - E[c^4]/\theta_M\right)\sigma'\mathbf{d}\right].\end{aligned}$$

- Kurtosis 2 (additive):

$$K_2 = E[c^4] - 3E^2[c^2].$$

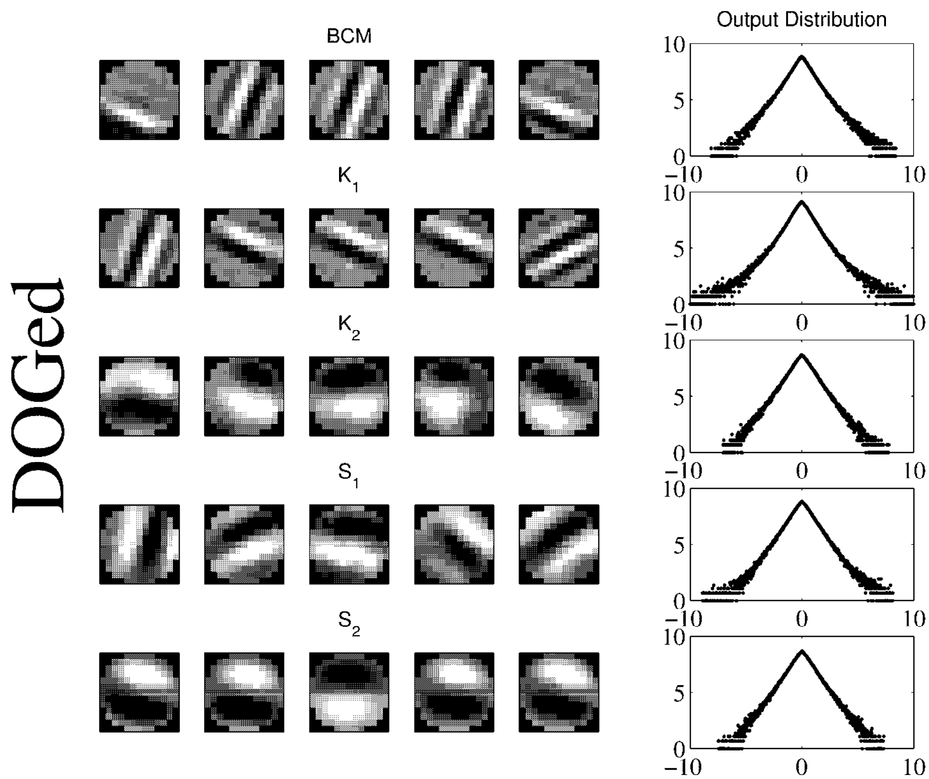
$$\begin{aligned}\nabla K_2 &= 4E\left[\left(c^3 - 3cE[c^2]\right)\sigma'\mathbf{d}\right] \\ &= 4E\left[c(c^2 - 3\theta_M)\sigma'\mathbf{d}\right], \quad \| \mathbf{m} \| = 1.\end{aligned}$$

- QBCM

$$\text{QBCM} = \frac{1}{3}E[c^3] - \frac{1}{4}E^2[c^2].$$

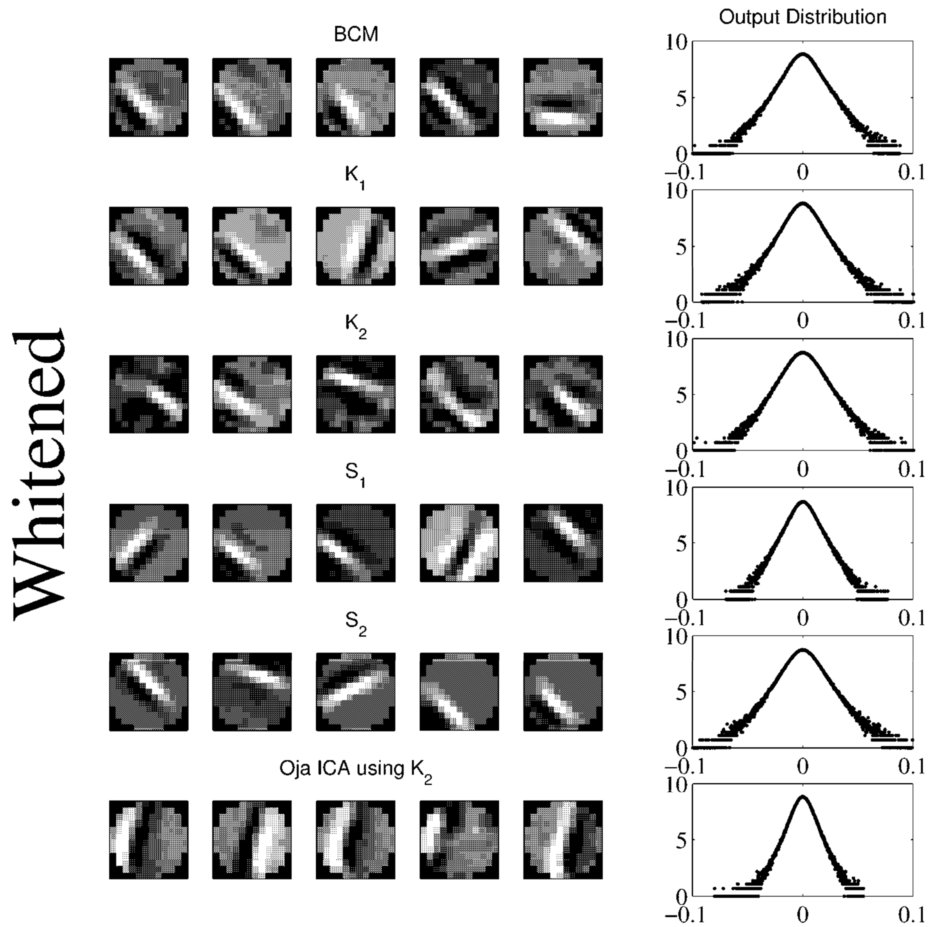
$$\begin{aligned}\nabla \text{QBCM} &= E\left[c^2 - cE[c^2]\right] \\ &= E[c(c - \theta_M)\sigma'\mathbf{d}].\end{aligned}$$

Projections (RFs) from Natural Scenes (DOGed)



Top to bottom: QBCM, K_1, K_2, S_1, S_2. Shown are five examples from each learning rule as well as the log of the normalized output distribution, before the application of the rectifying sigmoid.

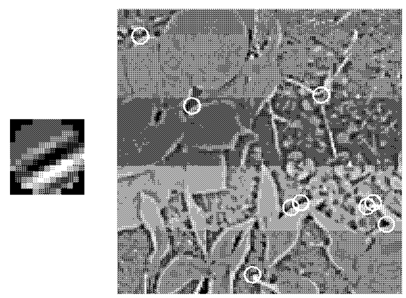
Projections from Natural Scenes (Sphered)



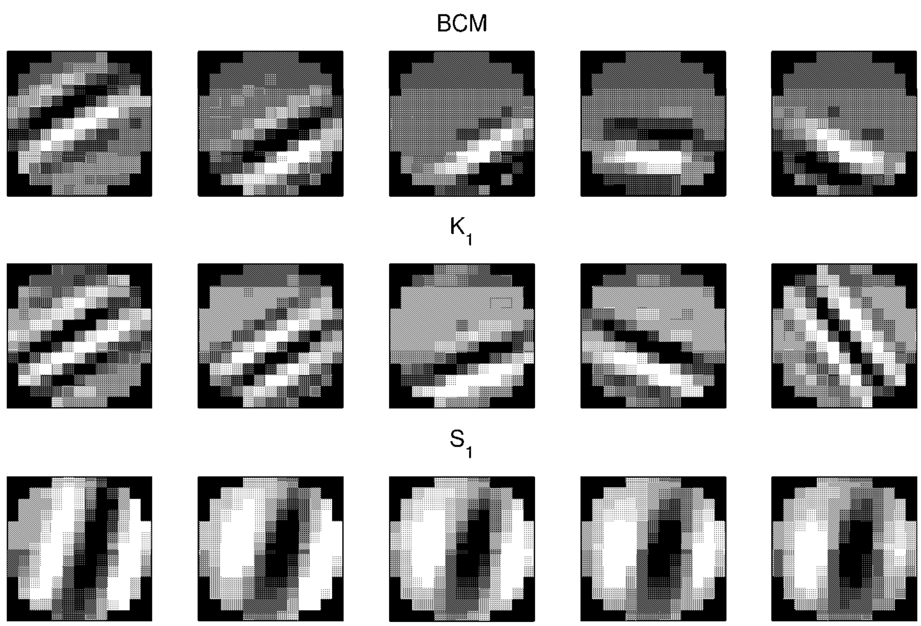
Top to bottom: QBCM, K_1, K_2, S_1, S_2. Shown are five examples from each learning rule as well as the log of the normalized output distribution, before the application of the rectifying sigmoid.

Structure Removal
(Sensitivity to outliers)

Patterns leading to high response

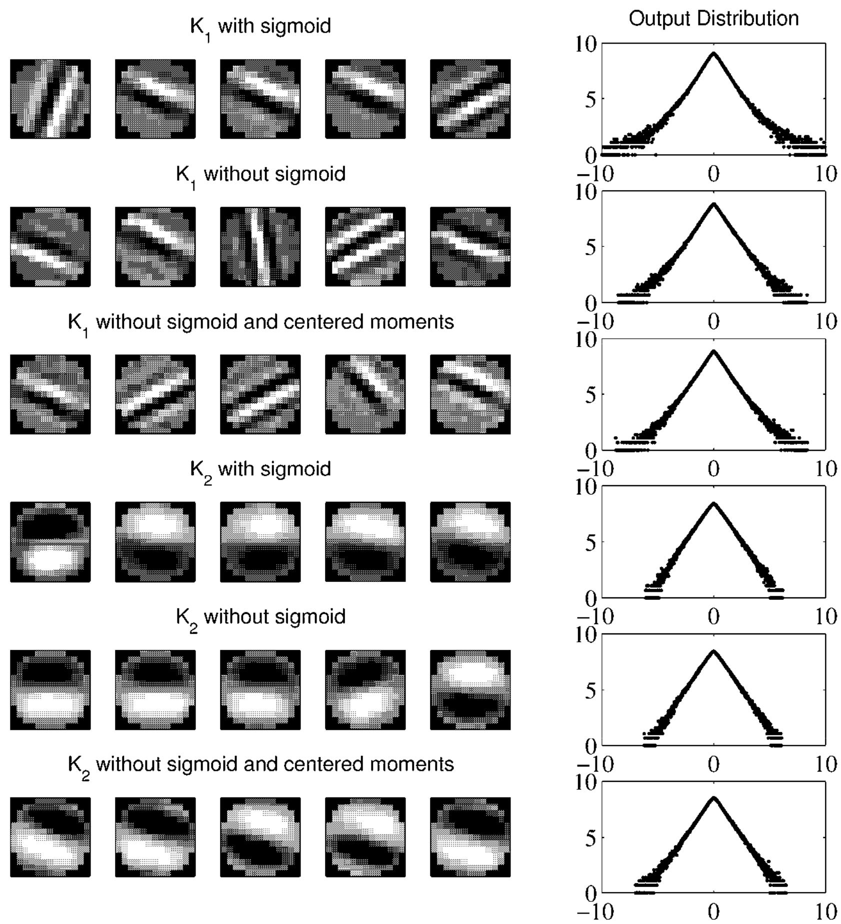


Effect of removal or top 1% response patterns



Top to bottom: QBCM; Kurtosis (K_1); Skewness (S_1).

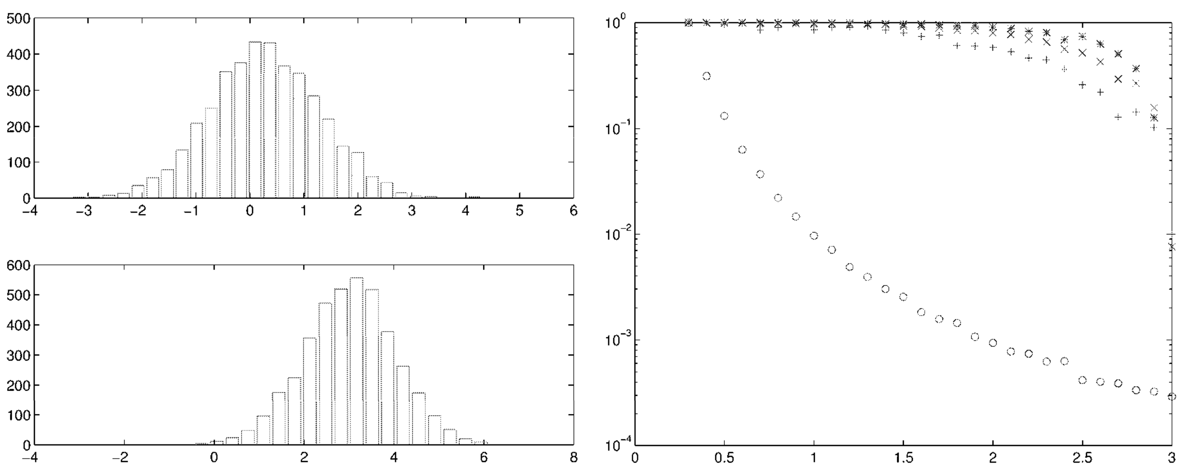
Variants on the Kurtosis rules



DOGed images.

multiplicative, rectified outputs, non-rectified outputs, non-rectified outputs with centered moments,
additive with rectified outputs, non-rectified outputs, non-rectified outputs with centered moments.

Measuring Bi-modality



kurtosis: 'x',
Friedman's deviation from uniformity: '+',
approximation of the negative entropy: '*',
QBCM index: 'o'.

Projection Pursuit Regression

- Presented by Friedman and Stuetzle (1981)
- Let (X, Y) be a pair of random variables, $X \in R^d$, and $Y \in R$.
- We seek an approximation to the d dimensional surface
$$f(x) = E[Y|X = x]$$
from n observations $(x_1, y_1), \ldots, (x_n, y_n)$.
- PPR tries to approximate a function f by a sum of ridge functions
$$f(x) \simeq \sum_{j=1}^{m} g_j(a_j^T x).$$
- **Neural Networks: The function g_j is a sigmoidal**