Part 3: Coding with Noisy Populations

Problem 1: Real neurons are "noisy"

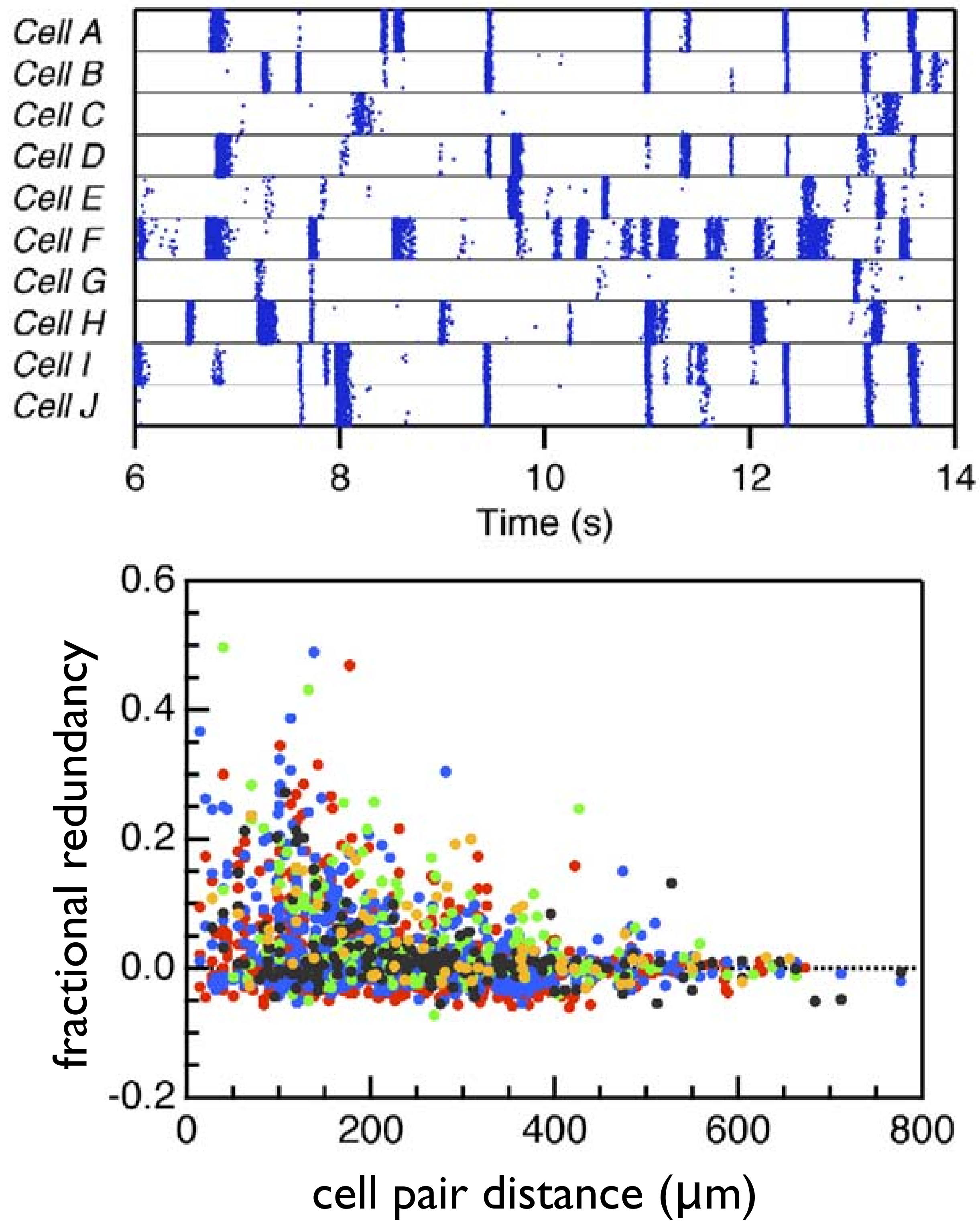
Estimates of neural information capacity

system (area)	**stimulus**	**bits / sec**	**bits / spike**
fly visual (H1)	motion	64	~1
monkey visual (MT)	motion	5.5 - 12	0.6 - 1.5
frog auditory (auditory nerve)	noise & call	46 & 133	1.4 & 7.8
Salamander visual (ganglinon cells)	rand. spots	3.2	1.6
cricket cercal (sensory afferent)	mech. motion	294	3.2
cricket cercal (sensory afferent)	wind noise	75 - 220	0.6 - 3.1
cricket cercal (10-2 and 10-3)	wind noise	8 - 80	avg. = 1
Electric fish (P-afferent)	amp. modulation	0 - 200	0 - 1.2

After Borst and Theunissen, 1999

Redundancy plays an important role in neural coding

Response of salamander retinal ganglion cells to natural movies



From Puchalla et al, 2005

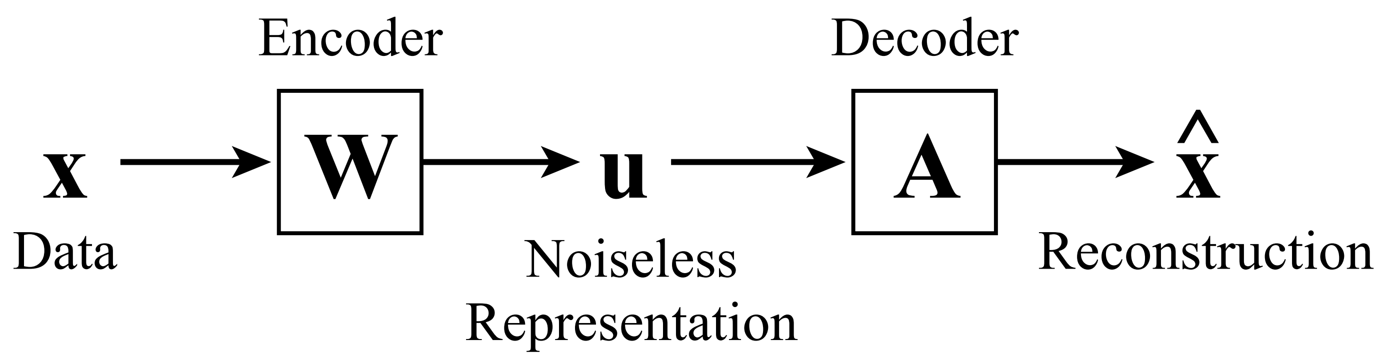
Representation in spikes limits neural capacity

- Graded potentials can convey 5 times as much information as spikes (de Ruyter van Steveninck & Laughlin, 1996)
- The spike generator limits the information in a retinal ganglion cell (Dhingra & Smith, 2004)

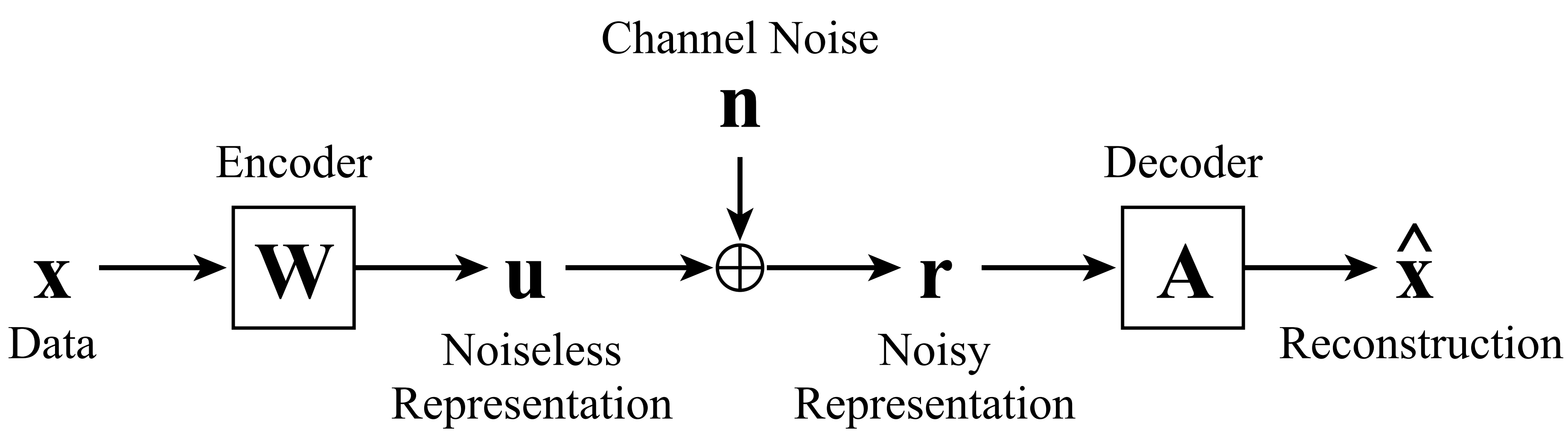
$\boxed{\text{Limited capacity} \Rightarrow \text{neural codes need to be } \textit{robust.}}$

(This is also a desirable property in real systems)

Defining the problem: A standard linear encoder/decoder



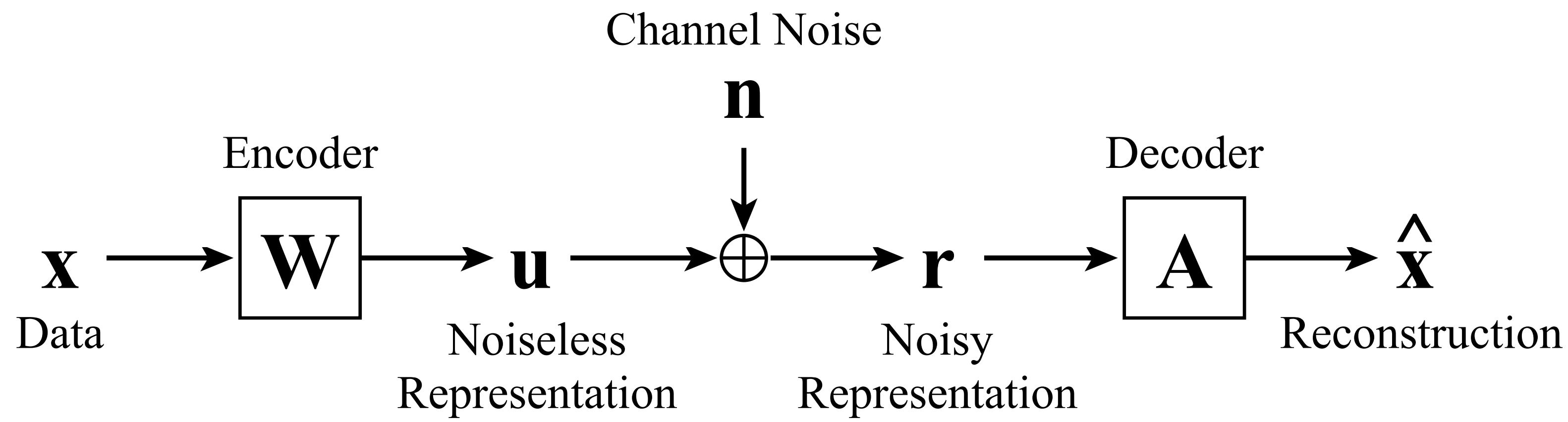
Model limited precision using additive channel noise:



$$\mathbf{r} = \mathbf{W}\mathbf{x} + \mathbf{n} = \mathbf{u} + \mathbf{n}.$$
$$\hat{\mathbf{x}} = \mathbf{A}\mathbf{r} = \mathbf{A}\mathbf{W}\mathbf{x} + \mathbf{A}\mathbf{n}.$$

Caveat:
It's linear, but we can evaluate it for different noise levels.

Traditional codes are *not* robust



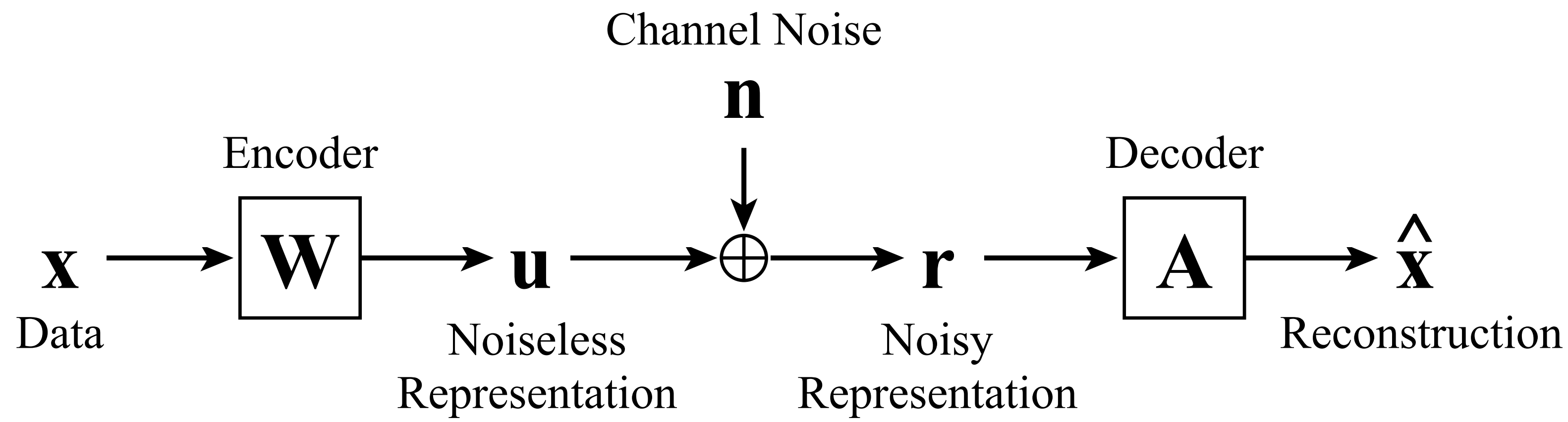
Original

wavelet coding

Add channel noise equivalent to 1 bit precision

reconstruction error: 34.8%

Traditional codes are *not* robust

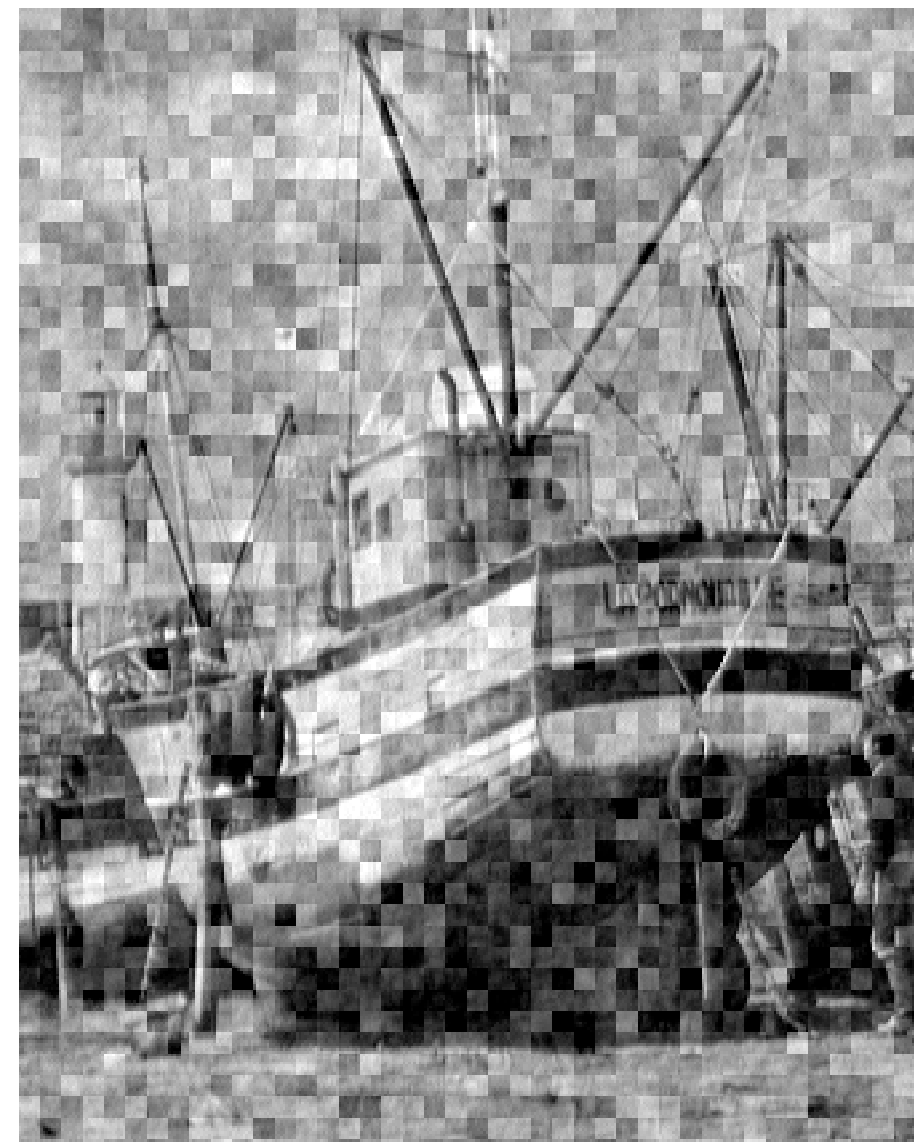


Original

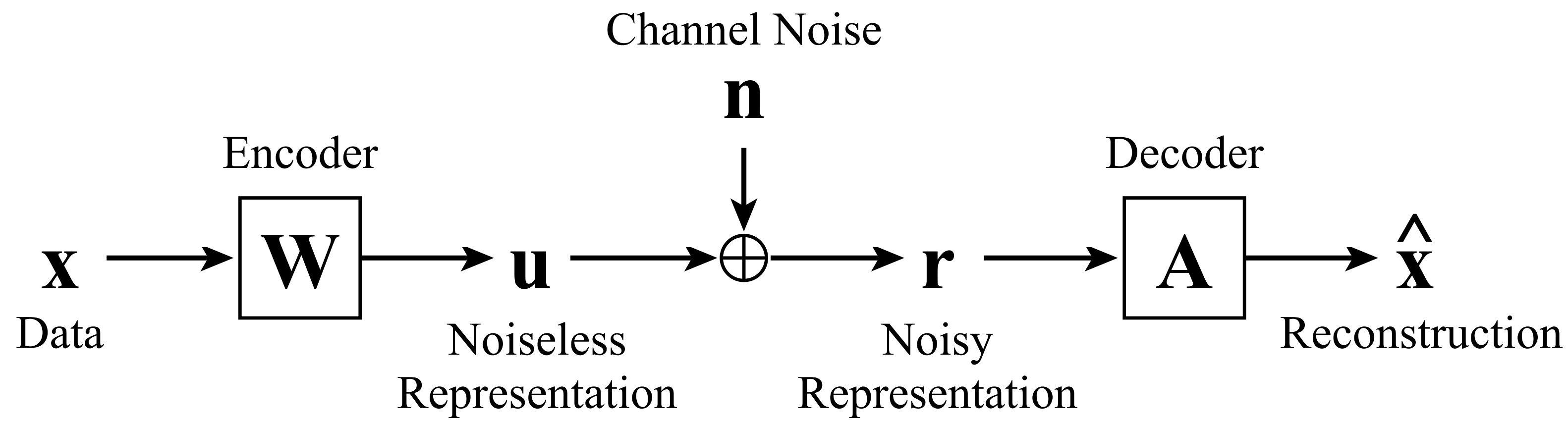
ICA coding

Add channel noise equivalent to 1 bit precision

reconstruction error: 34.8%



Robustness is not dimensionality reduction



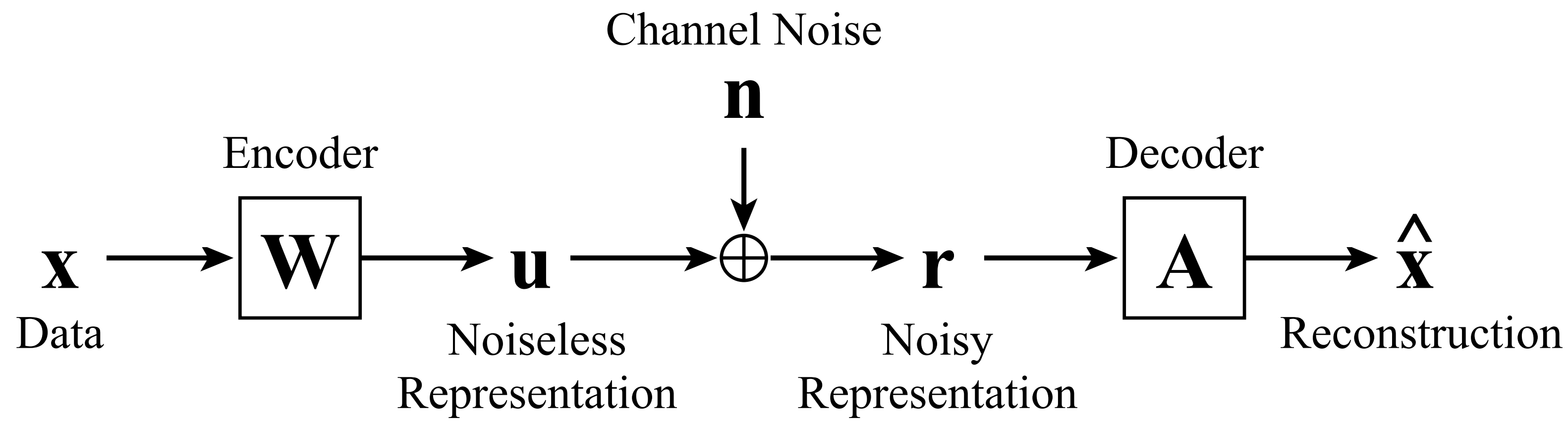
Reduce dimensionality by ½ using PCA

Original

No channel noise

reconstruction error: 4.5%

Robustness is not dimensionality reduction



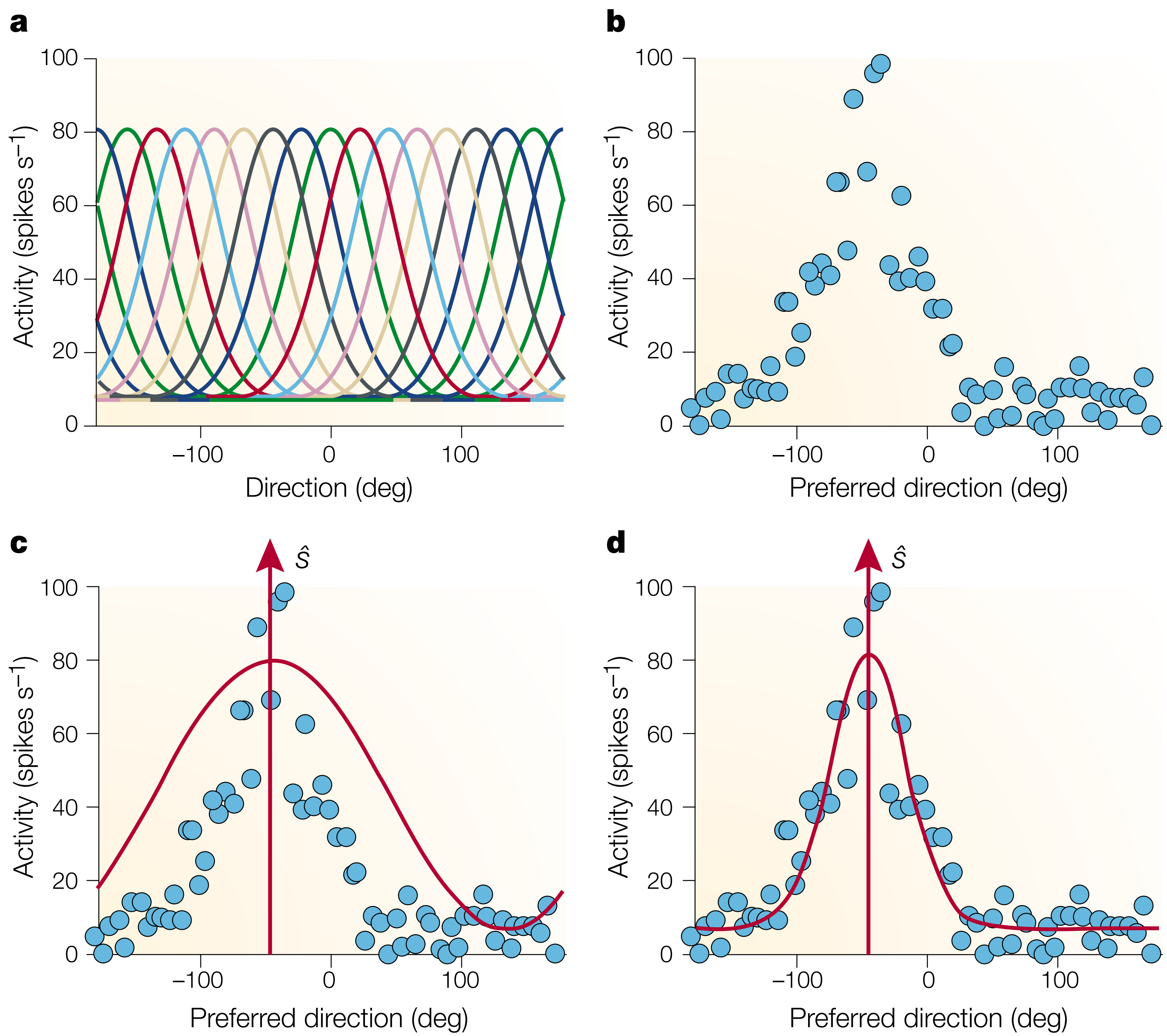
Reduce dimensionality by ½ using PCA

Original

channel noise equal
to 1 bit precision

reconstruction error: 36.4%

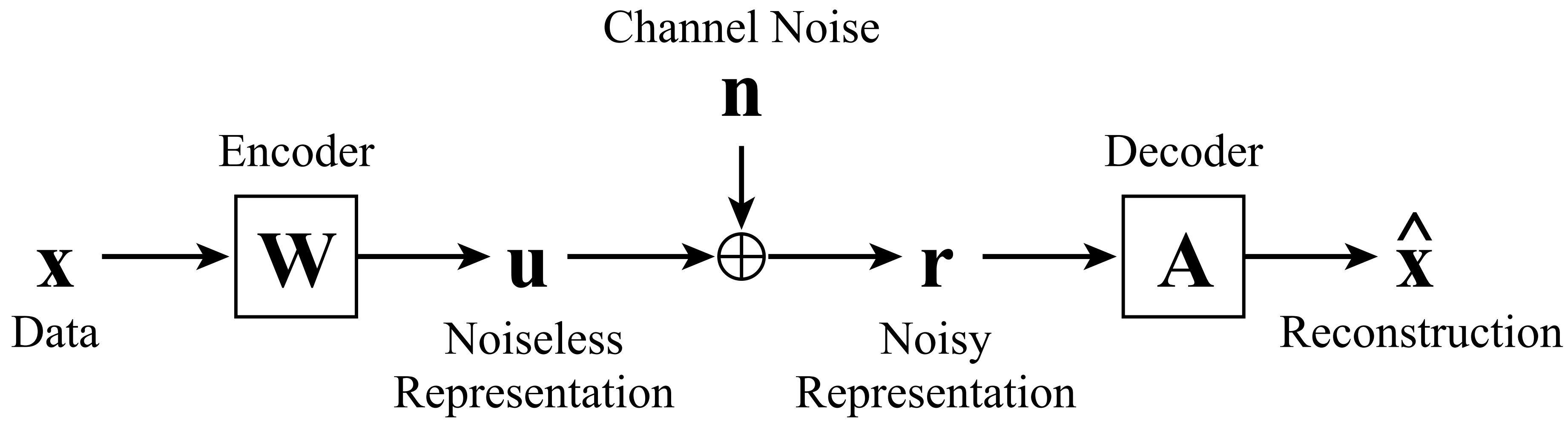
This is distinct from “reading” noisy populations



From Pouget et al, 2001

Here, we want to *learn* an optimal image code using noisy neurons.

How do we learn robust codes?



Objective:

Find $\mathbf{W}$ and $\mathbf{A}$ that minimize reconstruction error.

Residual error: $\boldsymbol{\epsilon} \;=\; \mathbf{x} - \hat{\mathbf{x}} \;=\; (\mathbf{I}_N - \mathbf{AW})\mathbf{x} - \mathbf{An}$

Squared error: $\mathcal{E} = \langle \boldsymbol{\epsilon}^T \boldsymbol{\epsilon} \rangle = \mathrm{tr}\left(\langle \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T \rangle\right)$

$$\mathcal{E}(\mathbf{A},\mathbf{W}) \;=\; \mathrm{tr}\{(\mathbf{I}_N - \mathbf{AW})\boldsymbol{\Sigma}_{\mathbf{x}}(\mathbf{I}_N - \mathbf{AW})^T\} + \sigma_n^2\,\mathrm{tr}\{\mathbf{AA}^T\}$$

We also need to constrain capacity.

Constraining neural capacity

- Channel capacity of the i[th] neuron:

$$C_i = \frac{1}{2}\ln(\mathrm{SNR}_i + 1)$$

- To limit capacity, we fix the coefficient variance:

$$\mathrm{SNR}_i = \frac{\langle u_i^2 \rangle}{\sigma_n^2}$$

- This implies $\mathbf{W}$ must satisfy:

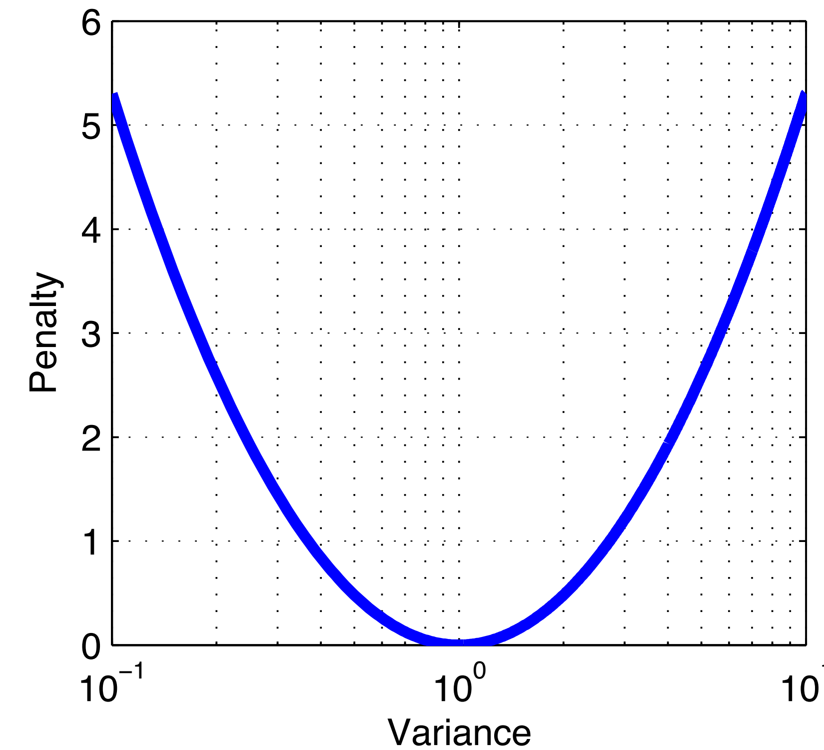
$$\mathrm{diag}(\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{W}^T) \;=\; \sigma_u^2 \; \mathbf{1}_M$$

Now robust coding is formulated as a *constrained optimization* problem.

Solving the robust coding problem

- Analytically: can compute exact solutions for 1-D and 2-D cases.

- Numerically:



$$(\text{Cost}) \quad = \quad (\text{Error}) \; + \; \lambda \; (\text{Var. Const.})$$

$$(\text{Var. Const.}) \quad = \quad \sum_{i=1}^{M} \left[\ln \left(\frac{\langle u_i^2 \rangle}{\sigma_u^2} \right) \right]^2$$

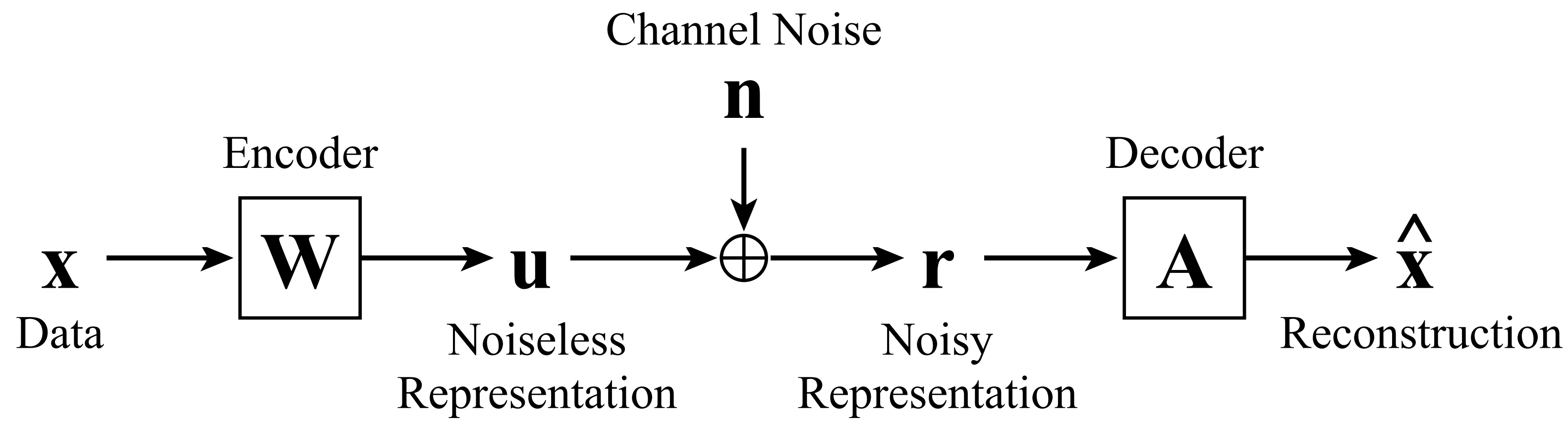
$$\begin{aligned}
\Delta \mathbf{W} \;\; &\propto \;\; -\frac{\partial}{\partial \mathbf{W}} (\text{Cost}) \\
&= \; 2 \; \mathbf{A}^T (\mathbf{I}_N - \mathbf{A}\mathbf{W}) \boldsymbol{\Sigma}_{\mathbf{x}} \; - \; \lambda \; \frac{4}{M} \; \text{diag} \left[\frac{\ln\{\text{diag}(\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{W}^T)/c\}}{\text{diag}(\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{W}^T)} \right] \mathbf{W}\boldsymbol{\Sigma}_{\mathbf{x}}
\end{aligned}$$

$$\mathbf{A} \quad = \quad \boldsymbol{\Sigma}_{\mathbf{x}} \mathbf{W}^T (\sigma_n^2 \mathbf{I}_M + \mathbf{W}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{W}^T)^{-1} \quad \Leftrightarrow \quad \frac{\partial}{\partial \mathbf{A}} (\text{Cost}) \quad = \quad \mathbf{O}$$

Robust coding of natural images: methods

- 8x8 pixel blocks from 62 natural images (Kyoto natural image dataset)
- No preprocessing except for DC subtraction for each sample.
- Precision is set as 1 bit for each neuron, comparable to neural data.

Comparing robust coding to traditional codes



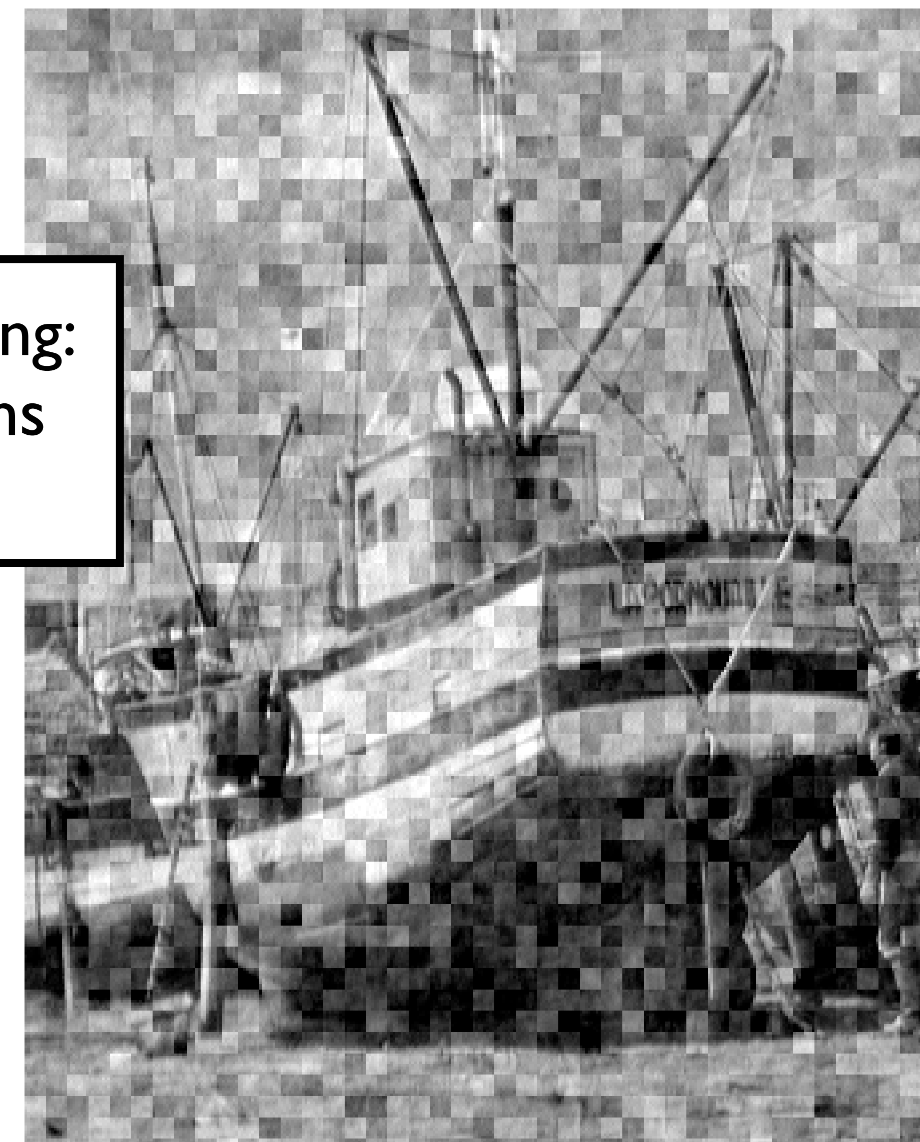
Original

ICA coding

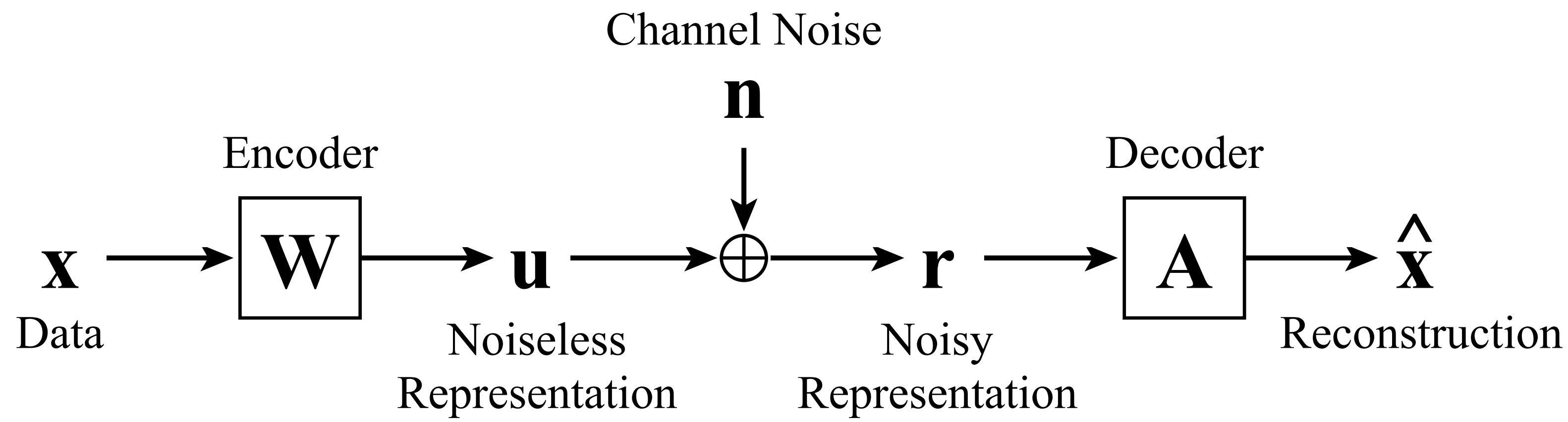
1 bit precision

Now use 1x robust coding:
same number of neurons
same precision

reconstruction error: 34.8%



Robust coding of natural images greatly improves reconstruction



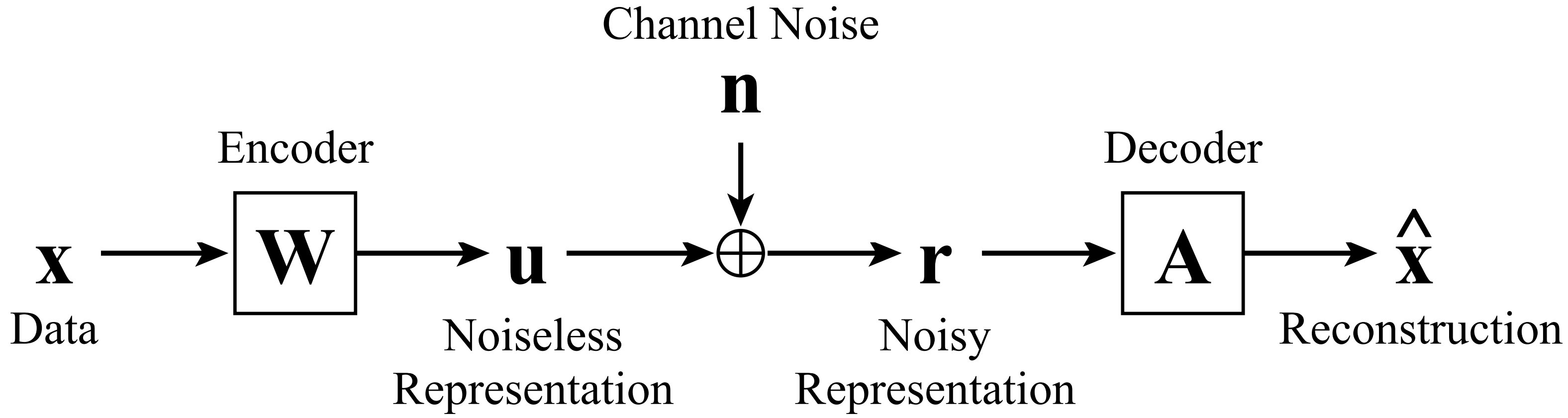
Original

1x robust coding

1 bit precision

reconstruction error: 4.5%

Reconstruction error can be reduced arbitrarily by adding neurons



Original

8x robust coding

1 bit precision

reconstruction error: 1.5%

Generalizing minimum squared error bound for N-D

2-D Anisotropic

$$\mathcal{E} = \frac{1}{\frac{M}{2} \cdot \mathrm{SNR} + 1} \frac{(\sqrt{\lambda_1} + \sqrt{\lambda_2})^2}{2} \qquad \text{if } \mathrm{SNR} \geq \mathrm{SNR}_c$$

λ_i - ith eigenvalue of the data.

M - # of coding units (neurons)

conjecture:

N-D Anisotropic

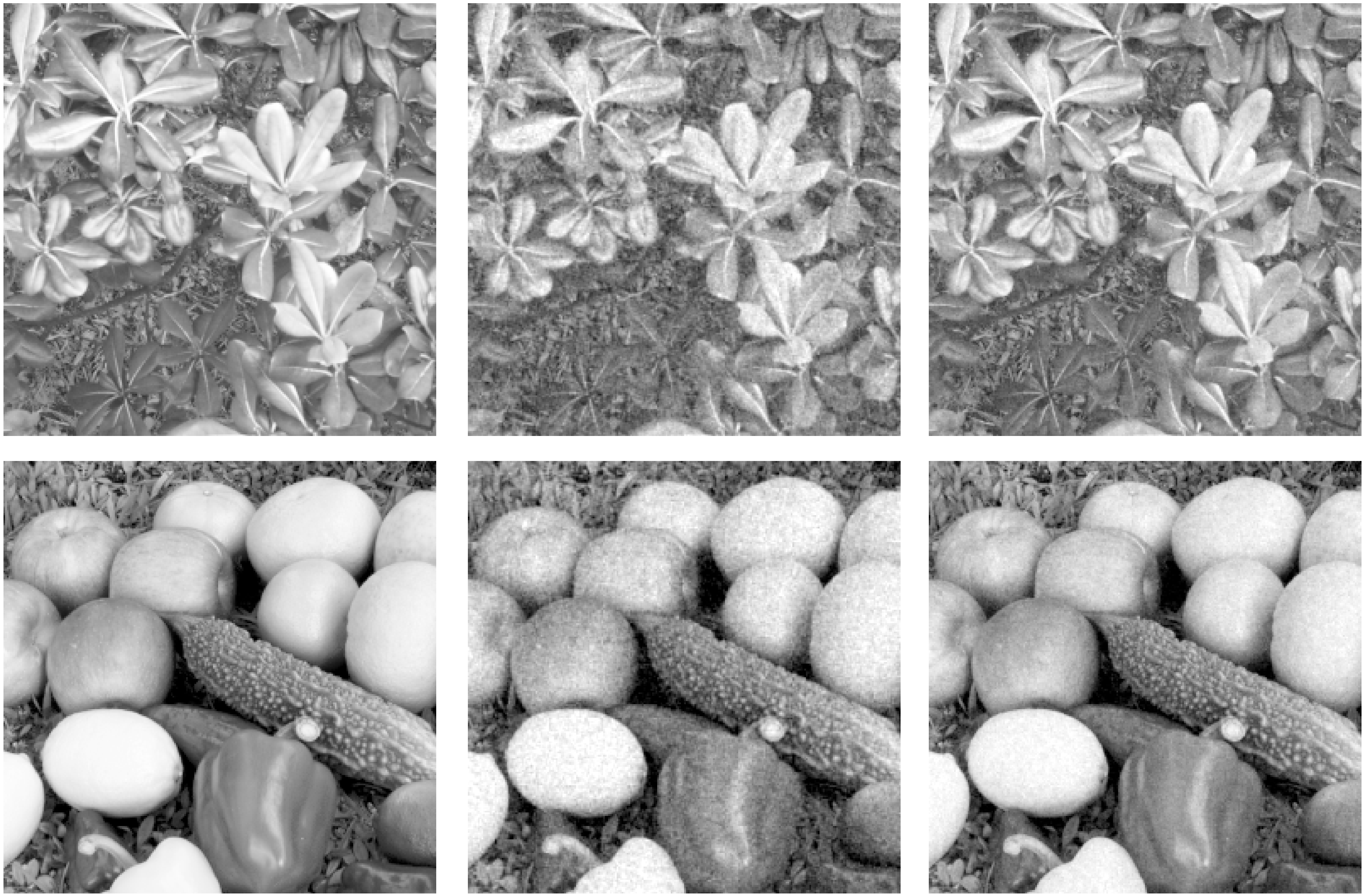
$$\mathcal{E} = \frac{1}{\frac{M}{N} \cdot \mathrm{SNR} + 1} \frac{1}{N} \left[\sum_{i=1}^{N} \sqrt{\lambda_i} \right]^2 \qquad \text{if } \mathrm{SNR} \geq \mathrm{SNR}_c$$

Predicted vs actual error

	Results	Conjecture
0.5x	19.9 %	20.3 %
1x	12.4 %	12.5 %
8x	2.0 %	2.0 %

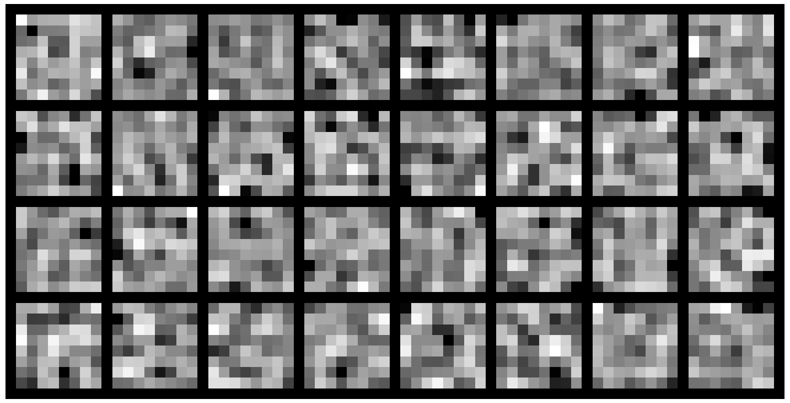
Adding precision to 1x robust coding

original images | 1 bit: 12.5% error | 2 bit: 3.1% error

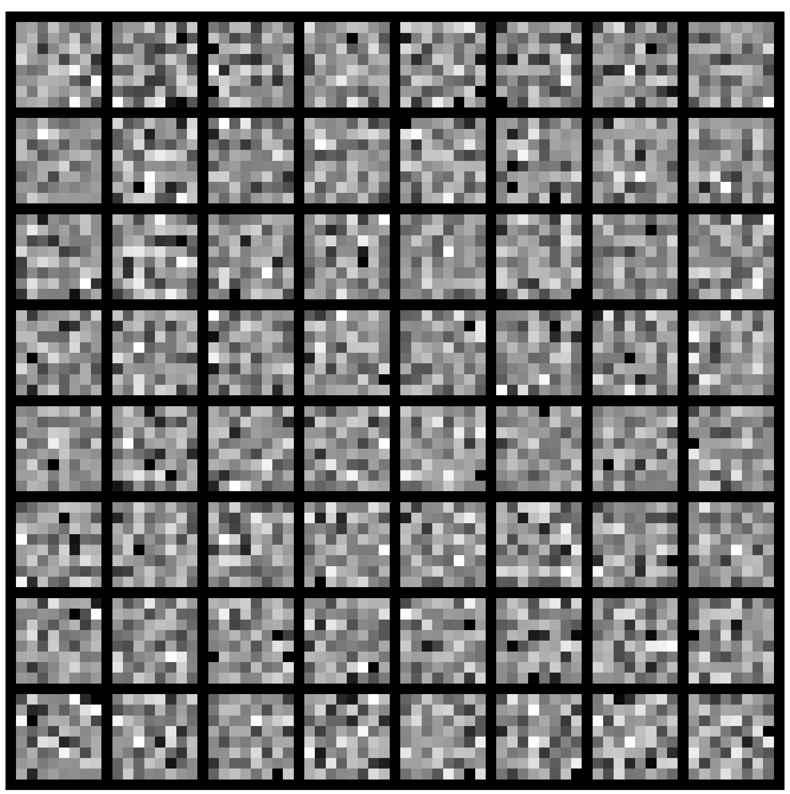


Robust Coding: encoding vectors ($\mathbf{W}$)

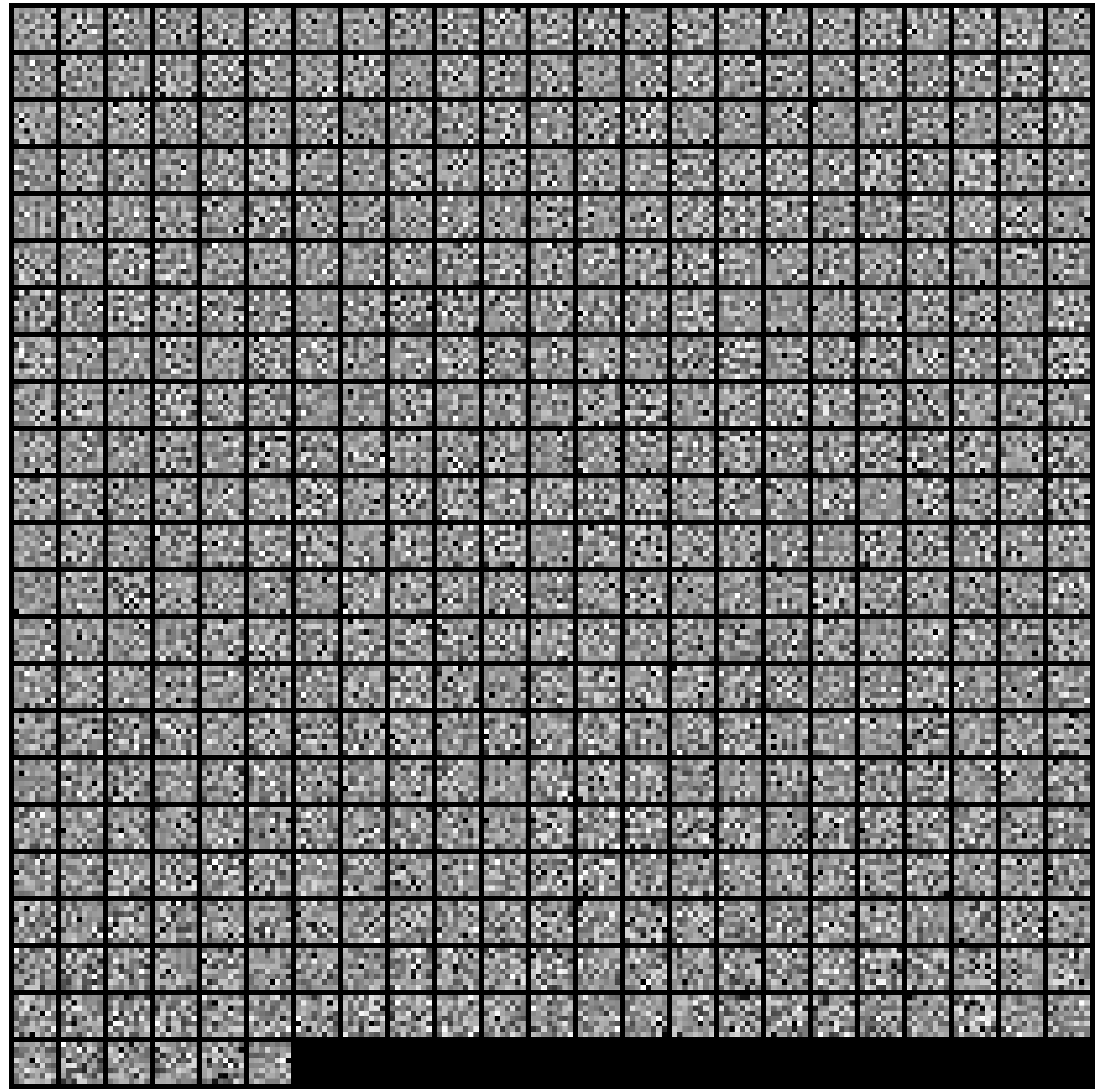
0.5x



1x



8x



Making robust codes efficient

- Current robust coding objective minimizes error subject to capacity constraint:

$$\text{(Cost)} \quad = \quad \text{(Error)} \; + \; \lambda \; \text{(Var. Const.)}$$

- We can add an additional constraint to improve coding efficiency:

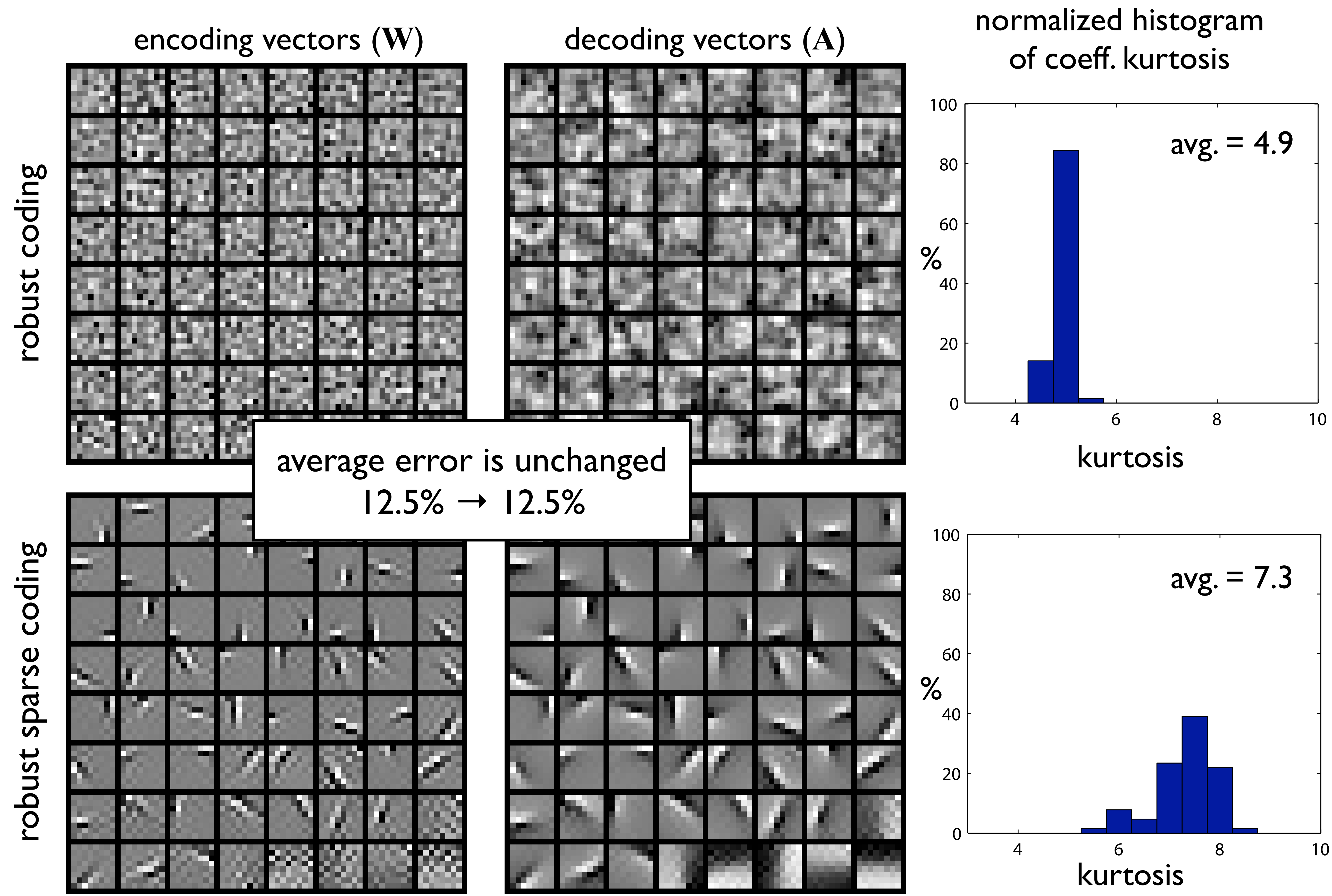
$$\text{(Cost)} \quad = \quad \text{(Error)} \; + \; \lambda \; \text{(Var. Const.)} \; - \; \lambda' \; \text{(Sparseness)}$$

- Using a Cauchy sparseness term, the new update rule is:

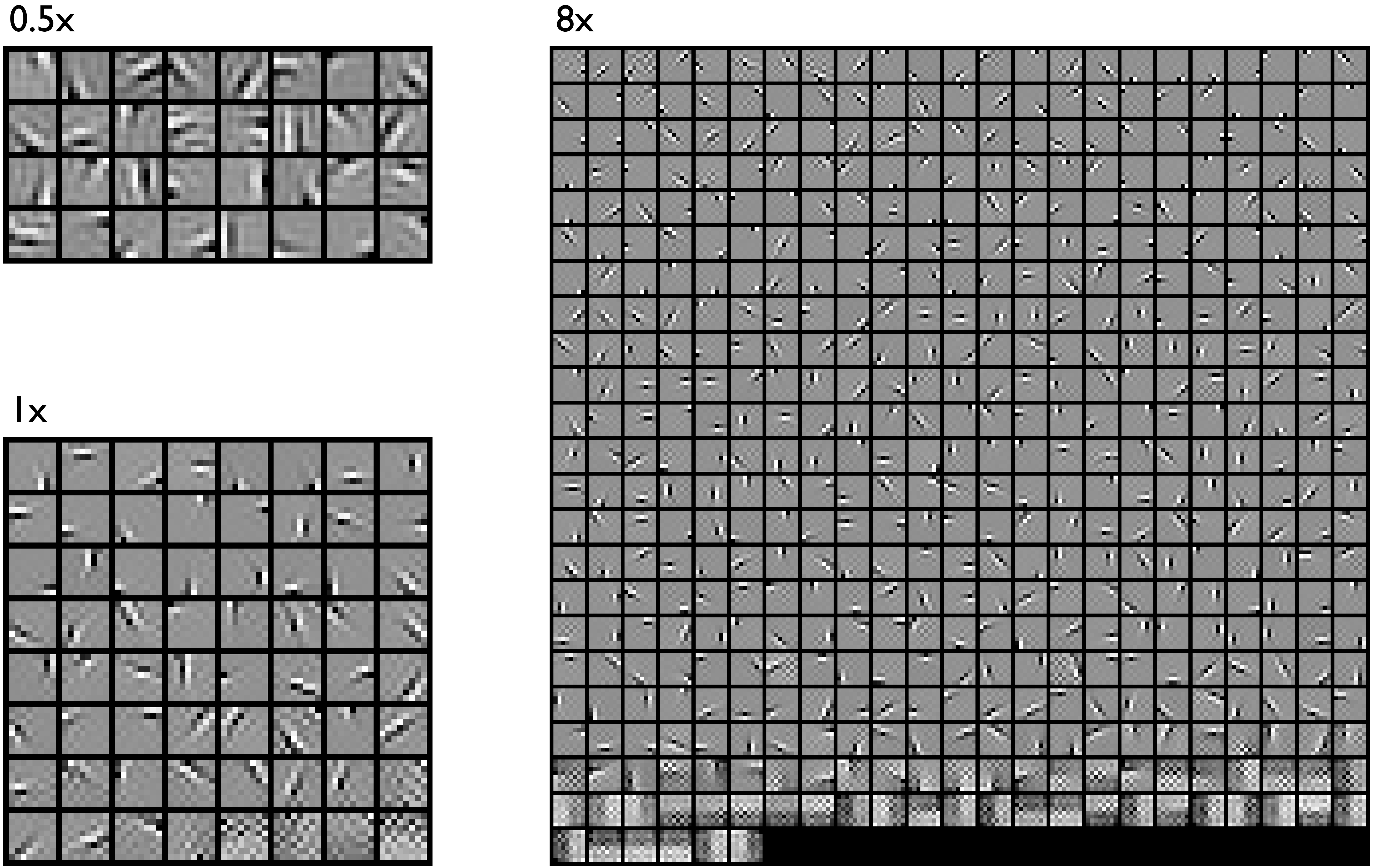
$$\begin{aligned}\Delta\mathbf{W} \;=\; & 2\;\mathbf{A}^T(\mathbf{I}_N - \mathbf{A}\mathbf{W})\mathbf{\Sigma}_\mathbf{x} \;-\; \lambda\,\frac{4}{M}\;\mathrm{diag}\left[\frac{\ln\{\mathrm{diag}(\mathbf{W}\mathbf{\Sigma}_\mathbf{x}\mathbf{W}^T)/c\}}{\mathrm{diag}(\mathbf{W}\mathbf{\Sigma}_\mathbf{x}\mathbf{W}^T)}\right]\mathbf{W}\mathbf{\Sigma}_\mathbf{x} \\ & +\lambda'\;\frac{2\beta}{M\sigma^2}\;\frac{\mathbf{W}\mathbf{x}}{1+\mathrm{diag}(\mathbf{W}\langle\mathbf{x}\mathbf{x}^T\rangle\mathbf{W}^T)/\sigma^2}\;\mathbf{x}^T\end{aligned}$$

$$\mathbf{A} \quad = \quad \mathbf{\Sigma}_\mathbf{x}\mathbf{W}^T(\sigma_n^2\mathbf{I}_M + \mathbf{W}\mathbf{\Sigma}_\mathbf{x}\mathbf{W}^T)^{-1} \quad \Leftrightarrow \quad \frac{\partial}{\partial\mathbf{A}}\text{(Cost)} \quad = \quad \mathbf{O}$$

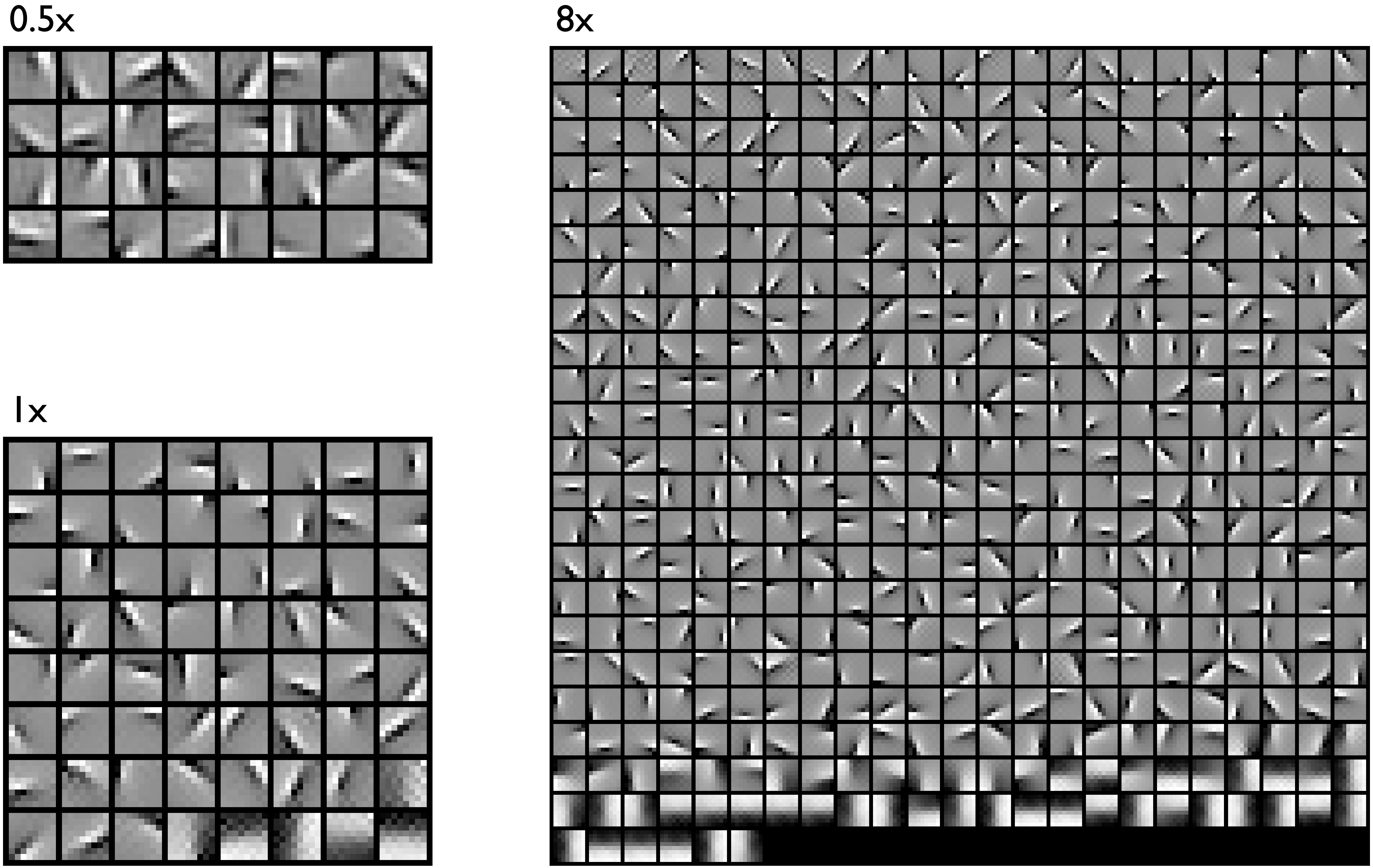
Sparseness localizes the vectors and increases coding efficiency



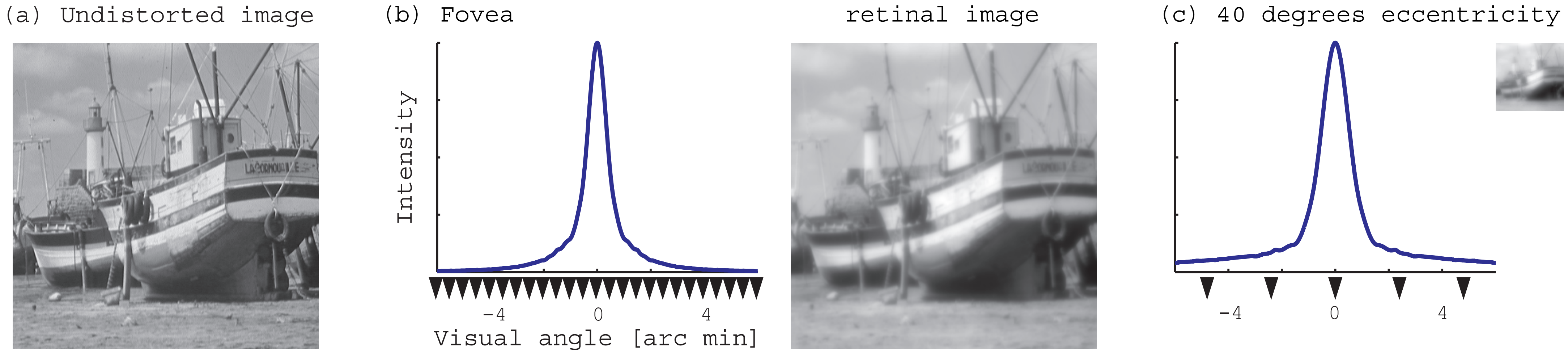
Robust *Sparse* Coding: encoding vectors (**W**)

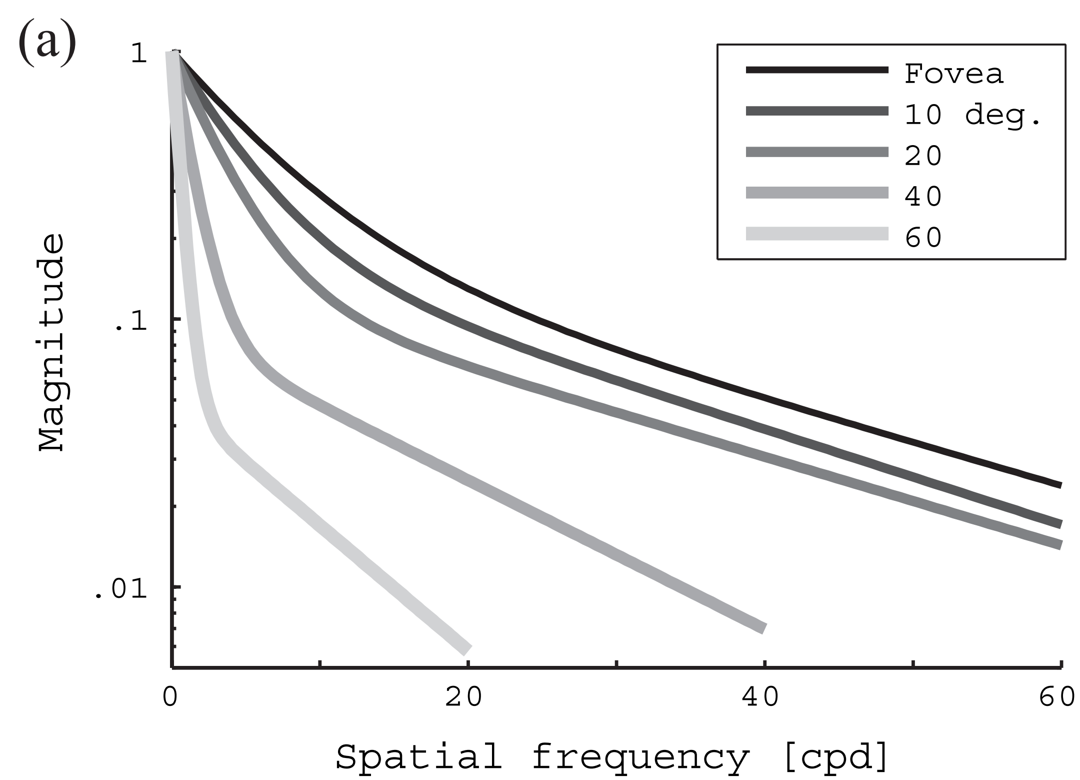


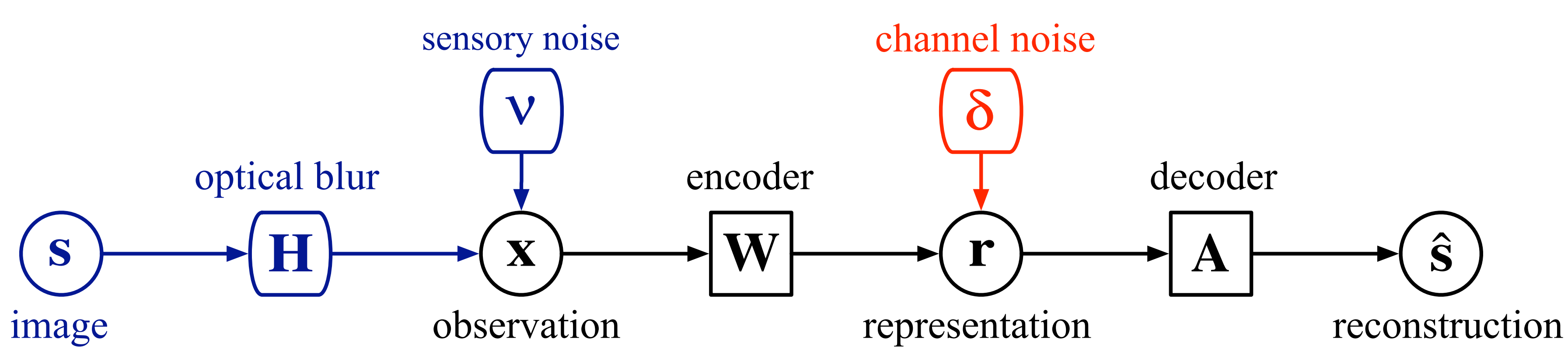
Robust *Sparse* Coding: decoding vectors ($\mathbf{A}$)



Incorporating sensory noise







Optimal receptive fields for different noise levels and eccentricities