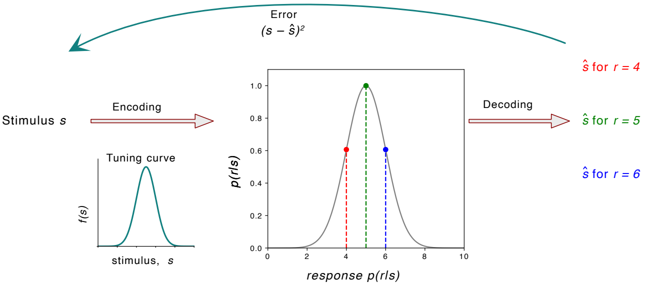
Neural Coding

6. Correlations

Anna Levina

May 21, 2025

Previous setting



We considered the response r of a *single neuron* to a stimulus s, modeled as a noisy function of s, with a tuning curve $f(s)$ and trial-to-trial variability described by $p(r|s)$.
In the previously described coding setup, what shall we do with many neurons? One option is to have a large number of them coding for the same stimulus and summing them, thus decreasing the noise in spike generation. It will work straightforwardly if the neurons are independent of each other (depending only on input).

General setup, correlated populations

But real neurons are not independent. What happens when the variability of different neurons is *correlated*? In a simplified form, we assume that the response is:

$$r \sim \mathcal{N}\left(f(\mathrm{s}), \Sigma(\mathrm{s})\right),$$

where *covariance matrix*

$$\Sigma(\mathrm{s}) = \mathrm{cov}[\mathrm{r}(\mathrm{s})] = E\left[(\mathrm{r} - E[\mathrm{r}])(\mathrm{r} - E[\mathrm{r}])^T\right] = E[\mathrm{r}\mathrm{r}^T] - E[\mathrm{r}]E[\mathrm{r}]^T$$

For individual entries, we have:

$$\Sigma_{ij} = \Sigma_{ji} = E[(r_i - E[r_i])(r_j - E[r_j])]$$

Covariance matrix. Properties

If the input is $1 \times n$-dimensional random vector x, A is $n \times m$ matrix and a is $1 \times m$ constant, then

$$\mathrm{cov}(\mathrm{Ax} + \mathrm{a}) = \mathrm{A}\,\mathrm{cov}(\mathrm{x})\mathrm{A}^{\mathrm{T}} = \mathrm{A}\Sigma(\mathrm{x})\mathrm{A}^{\mathrm{T}}$$

$\Sigma(\mathrm{x})$ is positive semi-definite (and each symmetric positive semi-definite matrix is a covariance matrix for some random variable)

$$M \text{ positive semi-definite} \iff \quad \mathrm{x}^\top M \mathrm{x} \geq 0 \text{ for all } \mathrm{x} \in \mathbb{R}^n \backslash \{0\}$$

Reminder: Positive semi-definite matrices have all eigenvalues greater than or equal to zero.

Correlated variability and accuracy of a population code

If N neurons code for the same thing (having the same $f(s)$ and the same noise in spike generation). Simple population code will be average of their rates: $\hat{r}_N = 1/N \sum r_i$, we want $\hat{r}_N \to r = f(s)$. And $\mathrm{var}[\hat{r}_N] \to 0$. If the neurons are independent, we have:

$$\mathrm{var}\left[\frac{1}{N}\sum_{i=1}^{N} r_i\right] = \frac{1}{N^2}\mathrm{var}\left[\sum_{i=1}^{N} r_i\right] = \frac{1}{N^2} N\,\mathrm{var}[r_i] = \frac{\mathrm{var}[r_i]}{N}.$$

$$\mathrm{var}\left(\sum_{i=1}^{N} X_i\right) = \sum_{i=1}^{N}\mathrm{var}(X_i) + \sum_{i\neq j}\mathrm{cov}(X_i, X_j)$$

New setting: include correlations. Neurons have rates r_i, individual mean rates f_i and same variance σ^2, but

$$\langle (r_i - f_i)(r_j - f_j)\rangle = \sigma^2\left[\delta_{ij} + c\,(1-\delta_{ij})\right],$$

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. Correlation coefficient c: $0 \leq c < 1$

Proof: Variance of sign-alternating population code

Using linearity and covariance structure:

$$\mathrm{var}[\tilde{R}] = \frac{1}{N^2} \sum_{i,j} (-1)^i (-1)^j \, \mathrm{cov}(r_i, r_j)$$

Split into diagonal and off-diagonal terms:

$$= \frac{1}{N^2} \left(N\sigma^2 + \sum_{i \neq j} (-1)^i (-1)^j c\sigma^2 \right)$$

If N is even, then: $\sum_{i=1}^{N} (-1)^i = 0$ This implies:

$$\left(\sum_{i=1}^{N} (-1)^i \right)^2 = 0 = \sum_{i=1}^{N} (-1)^i (-1)^i + \sum_{i \neq j} (-1)^i (-1)^j \Rightarrow$$

$$0 = N + \sum_{i \neq j} (-1)^i (-1)^j \quad \Rightarrow \quad \sum_{i \neq j} (-1)^i (-1)^j = -N$$

$$\Rightarrow \mathrm{var}[\tilde{R}] = \frac{\sigma^2}{N} (1 - c)$$

Fisher Information for Gaussian population codes

We follow Abbott and Dayan ("The effect of correlated variability on the accuracy of a population code"). Each neuron's response is modeled as:

$$r_i = f_i(s) + \eta_i, \quad \text{with noise } \boldsymbol{\eta} \sim \mathcal{N}(0, Q(s))$$

Then the likelihood of observing response vector r is:

$$P[\mathrm{r}|s] = \frac{1}{\sqrt{(2\pi)^N \det Q(s)}} \exp\left(-\frac{1}{2}[\mathrm{r} - \mathrm{f}(s)]^T Q^{-1}(s)[\mathrm{r} - \mathrm{f}(s)]\right)$$

The Fisher information consists of two terms:

$$I_F(s) = \underbrace{\mathrm{f}'(s)^T Q^{-1}(s)\mathrm{f}'(s)}_{\text{Mean sensitivity}} + \underbrace{\frac{1}{2}\,\mathrm{Tr}\left[Q'(s)Q^{-1}(s)Q'(s)Q^{-1}(s)\right]}_{\text{Noise change (ignored if } Q \text{ is constant)}}$$

Additive correlated noise

Assume additive noise: Q is independent of s. Then

$$I_F(s) = \mathrm{f}'(s)^T Q^{-1} \mathrm{f}'(s)$$

Let all neurons have variance σ^2 and pairwise correlation c. Then:

$$Q_{ij} = \sigma^2 \left[\delta_{ij} + c(1-\delta_{ij})\right], \quad Q_{ij}^{-1} = \frac{\delta_{ij}(Nc+1-c) - c}{\sigma^2(1-c)(Nc+1-c)}$$

Define:

$$F_1(s) = \frac{1}{N}\sum_i \left(f_i'(s)\right)^2, \quad F_2(s) = \left(\frac{1}{N}\sum_i f_i'(s)\right)^2, \quad \text{then}$$

$$I_F(s) = \frac{cN^2\left[F_1(s) - F_2(s)\right] + (1-c)NF_1(s)}{\sigma^2(1-c)(Nc+1-c)}, \Rightarrow\ I_F(s) \to \frac{N[F_1 - F_2]}{\sigma^2(1-c)}$$

The more *heterogeneous* the derivatives $f_i'(s)$ are (i.e., $F_1 > F_2$), the more information is preserved. But also, Fisher information is growing with N and c.

Limited-range correlations

Now consider a more realistic case: *limited-range correlations* that decay with distance between neurons.

$$Q_{ij} = \sigma^2 \rho^{|i-j|}, \quad \text{with} \quad \rho = \exp(-\Delta/L)$$

- Δ is the spacing between peaks of adjacent tuning curves
- L is a characteristic correlation length

In the limit $N \to \infty$, the Fisher information becomes:

$$I_F(s) \to \frac{N(1-\rho)F_1(s)}{\sigma^2(1+\rho)}$$

- I_F still grows linearly with N
- Higher correlation ($\rho \to 1$) reduces I_F: shared noise limits how much signal averaging helps
- Correlations decay with distance, so only nearby neurons interfere

Summary: Correlations reduce the *efficiency* of information scaling, but do not prevent the Fisher information from growing linearly with the number of encoding neurons.

Noise and signal covariance

Law of total covariance:

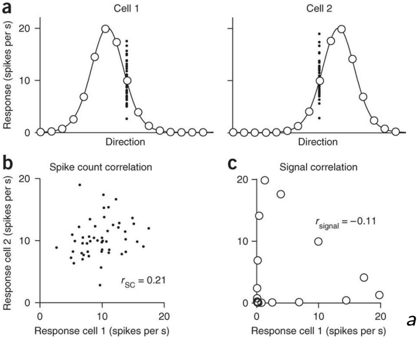
$$\text{cov}[\text{r}] = \text{cov}[E[\text{r}|s] + E[\text{cov}[\text{r}|s]],$$

Here $\text{cov}[E[\text{r}|s]]$ is called signal covariance and $E[\text{cov}[\text{r}|s]]$ noise covariance.
For the neurons and inputs in our typical setup: $E[\text{r}|s] = f(s)$ so we have:

$$\text{cov}_{\text{signal}} = E_{\text{x}}[f(s)f(s)^T] - E[f(s)]E[f(s)]^T$$

$$\text{cov}_{\text{noise}} = E_s[\Sigma(s)]$$

Noise correlations in more details

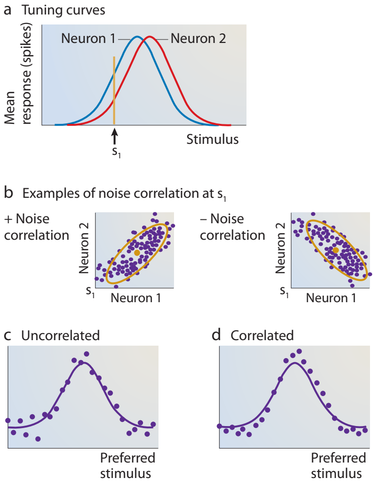


[a]

(a) Tuning curves for two neurons. Open circles show mean responses and small points show responses to individual presentations of a stimulus at a particular direction. (b) Noise correlations measures the correlation between fluctuations in responses to the same stimulus. (c) Signal correlation measures the correlation between the two cells' mean responses to different stimuli. Each point represents the mean response to a given direction of motion. Because the responses of cell 2 increase of a range of motion directions in which the responses of cell 1 decline, signal correlation is negative.

[a] Cohen and Kohn, "Measuring and interpreting neuronal correlations".

Noise correlations again, from a more theoretical view

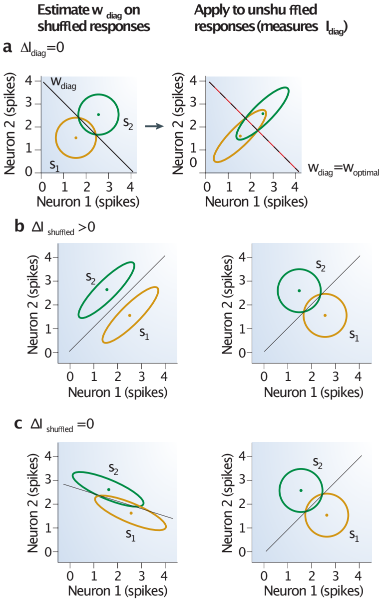


[a]

c and d: the response of a population of neurons. The x-axis - preferred orientation of the neuron, the response - on the y-axis. Neurons in both panels c and d exhibit noise fluctuations but on individual trials the responses of nearby neurons in panel c are uncorrelated (fluctuating up and down independently), in panel d - correlated (fluctuating up and down together). Nearby neurons in panel d are positively correlated (as in panel b, left), those that are far apart are negatively correlated (as in panel b, right).

[a] Averbeck, Latham, and Pouget, "Neural correlations, population coding and computation".

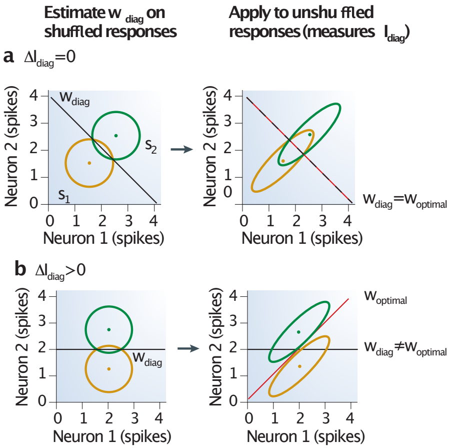
Intuition for the effect of correlation on encoding accuracy



Consider two neurons and two stimuli that we want to encode (such that they can be discriminated later). To understand the effect of the correlation we create a shuffled data, where correlations of responses are removed. $\Delta I_{\text{shuffled}} = I - I_{\text{shuffled}}$, here I is not formalized information. Line is the best discrimination threshold for ideal observer (who knows about the correlations)[a]

[a] Averbeck, Latham, and Pouget, "Neural correlations, population coding and computation".

Intuition for the naive decoding



Again two neurons and two stimuli. Neuronal responses are correlated. But imagine, that the decoder does not know about it and creates the decision lines based on the shuffled (correlation-free) responses. It can keep the performance (as in a), or destroy it (as in b)

Summary

- Correlations are inevitable in the responses of the neurons
- In the repetitive trials, we can split noise and signal correlations
- Noise correlations can have a detrimental but sometimes also helpful effect

Next time:

- Examples of noise correlations studies
- Neural manifolds

Bibliography

- Abbott, L. F. and Peter Dayan. "The effect of correlated variability on the accuracy of a population code". In: *Neural Computation* 11.1 (1999), pp. 91–101.
- Averbeck, Bruno B, Peter E Latham, and Alexandre Pouget. "Neural correlations, population coding and computation". In: *Nature reviews neuroscience* 7.5 (2006), pp. 358–366.
- Cohen, Marlene R. and Adam Kohn. "Measuring and interpreting neuronal correlations". In: *Nature Neuroscience* 14.7 (2011), pp. 811–819. arXiv: NIHMS150003.
- Shadlen, Michael N and William T Newsome. "Noise, neural codes and cortical organization". In: *Current Opinion in Neurobiology* 4.4 (Aug. 1994), pp. 569–579. ISSN: 09594388.