# Appendix A – When do the stimulus and noise axes coincide?

This appendix demonstrates that, in a linear‑rate network, the stimulus (coding) axis **need not** align with the dominant noise (slow‑mode) axis and derives the algebraic conditions under which the two directions coincide.

## 1 Network and notation

The population dynamics (see Methods) are

where

| Symbol | Meaning |
| --- | --- |
|  | recurrent weight matrix (spectral radius < 1) |
|  | static neuronal gains |
|  | feed‑forward drive vector |
|  | stimulus |
|  | private noise |
|  | time constant (set to 1) |

This is equivalent to

As .

### 1.1 Feedback operator

with eigen-decomposition

* **Noise / slow‑mode axis:** .

### 1.2 Stimulus axis

The stimulus‑conditioned mean is

Hence the **stimulus axis** is

## 2 The Spectral expansion of the stimulus axis

Using the Neumann series ,

Expand in the eigenbasis :

Then

The factor amplifies slower modes ().

### **3 Necessary and sufficient condition for alignment**

#### Proposition

The stimulus axis and the noise axis  are collinear iff

equivalently .

#### Proof

##### Necessity

Assume .  
Then for some .

From the spectral expansion,

Because , this equality implies for all .  
Thus has no component outside , i.e.

which is exactly (A1).

##### Sufficiency

Assume (A1): with .

* .
* for (orthogonality).

Therefore

## 4 Illustrative two‑unit example

One finds

whose cosine is (angle ). Unequal gains tilt the stimulus axis toward the high‑gain neuron despite the feature‑aligned weight matrices.

## 5 Conclusion

Stimulus–recurrent alignment is **not automatic**.

Appendix B (optimality proof) then shows that, once alignment is achieved, projecting along the common axis maximizes single‑axis Fisher information.

# Appendix B  –  Optimality of the noise axis

We prove that the Fisher ratio

is maximized at , i.e. when the read‑out vector is **exactly the noise/coding axis**.

## 1 Geometry of the tuned rank‑one network

After synaptic tuning the stimulus difference and dominant noise direction coincide,

where is the unit eigen‑vector of with the largest eigenvalue. Choose an orthonormal basis and parameterise the read‑out axis as

Because the noise covariance shares the same eigen‑basis, write

with (the slow mode has largest variance) and living in the remaining orthogonal sub‑space.

## 2 Closed‑form expressions for and

Using the orthogonality relations:

because has no projection onto the “rest’’ sub‑space.

Hence

## 3 Stationary points of

Differentiate w.r.t. :

Expand and simplify (factor ):

A simpler route is to observe that

which holds for . Only (and the equivalent ) lies in the admissible range and produces a finite, non‑zero numerator.

## 4 Nature of the stationary point

Compute the second derivative at :

because . Hence attains a **strict local maximum** at . Since is ‑periodic and even, this local maximum is global.

## 5 Conclusion

The Fisher ratio is maximised when ; that is, the **optimal linear decoder aligns with the noise (slow‑mode) axis**, confirming the result shown empirically in Fig. 2E.

# Appendix C Aligning the correlated variability direction with the stimulus boosts single-axis Fisher information

**Goal.** Show that, in the linear‑rate network defined in Appendices A and B, tuning the recurrent matrix so that the stimulus‑conditioned mean

becomes collinear with the dominant noise direction **increases** the single‑axis Fisher information whenever the decoder is constrained to one dimension.

## 1 Network definition and notation *(recap)*

We study the linear-rate network introduced in Appendix A:

(C.1)

where

* is the firing-rate vector,
* is the binary stimulus,
* are *feed-forward* weights,
* are static gains,
* is the *recurrent* weight matrix,
* is i.i.d. external noise, and
* is the spectral radius of .

Denote the eigendecomposition

(C.2)

and write the feed-forward drive in that basis

(C.3)

*Slow/Noise mode.* The dominant eigen-vector is the slowest dynamical mode and therefore the principal axis of internally generated correlated variability (Appendix A).

*Stimulus axis.* The difference in stimulus-conditioned means is

(C.4)

and its unit vector is called the *stimulus axis*.

*Noise covariance.* Because external noise is private, the covariance of population responses is

(C.5)

## 2 Single-axis Fisher information

For any unit *decoder* the linear Fisher information is

(C.6)

Choosing isolates the -th dynamical mode.  
With Eqs. (C.3)–(C.5) one finds

(C.7)

Hence—*holding fixed*—the signal-to-noise ratio along a mode is a strictly increasing function of its eigenvalue .

## 3 A general rank-one retuning of

Let the *baseline network* have eigen-pair and coefficients .  
Assume that **may or may not hold**; perhaps all feed-forward drive lands on a faster mode.

We now apply a **rank-one perturbation**

(C.8)

chosen so that

1. All eigenvalues are preserved: .
2. The dominant eigen-vector *rotates* onto the stimulus drive:  
   .

Condition 1 is feasible because adding a projector in the direction of changes only its eigen-vector, not the whole spectrum (classical rank-one perturbation theory).  
Condition 2 guarantees that after retuning

(C.9)

In words: *the stimulus axis is now perfectly aligned with the noise (slow) axis.*

## 4 Information gain of alignment

### Proposition C.1

Let be the Fisher information along the slow mode **before** retuning and **after** retuning as defined above. Then

(C.10)

where indexes the mode that carried the stimulus before alignment  
(i.e. ).

#### Proof

1. *Baseline.* Equation (C.7) gives
2. *Aligned network.* After retuning, Eq. (C.7) with and Eq. (C.9) give
3. *Constant signal power.* Because retuning (C.8) does **not** alter , . Taking the ratio yields Eq. (C.10).

Finally, since , both factors in the numerator exceed the corresponding factors in the denominator, proving the inequality. ∎

### Corollary C.2

The gain factor (C.10) is monotone in both eigenvalues; maximal gain is achieved when the stimulus is retuned from the *fastest* mode () to the *slowest* mode ().

## 5 Intuition

Equation (C.7) can be rewritten as

(C.11)

Increasing amplifies the *mean* as **faster** than it amplifies the *variance* (only ).  
Aligning the stimulus with the slowest mode therefore yields a net benefit proportional to .

## Where the extra comes from

For each mode:

* Mean amplification:
* Variance amplification:

Because decays **slower** than , the coherent mean gains an **extra factor**  relative to the noise power.

## 6 Summary

A purely *rank-one* adjustment of the recurrent matrix that **rotates (but does not rescale) the noise mode** can always bring the stimulus axis onto that mode. Because mean amplification outpaces variance amplification, this maneuver **strictly increases** single-axis Fisher information. The result holds for arbitrary feed-forward weights.