Supporting information for Nirenberg and Latham (2003)
Proc. Natl. Acad. Sci. USA, 10.1073/pnas.1131895100

Supporting Text

This document is divided into three sections. In the first, we show that the cost function used by Dan et al. (1) and Oram et al. (2) to assess the importance of synchronous spikes is the extra number of yes/no questions associated with ignoring synchronous spikes; therefore, it is equal to ΔI for a specific code. In the second, we show that correlations can exist without being important for decoding. In the third, we show that there isn't a reliable relation between the importance of correlations and either $\Delta I_{\text{shuffled}}$ or $\Delta I_{\text{synergy}}$.

The Information Difference Between Two Codes is Equal to ΔI

The correlations that are most often proposed to carry extra information are those contained in synchronous spikes across pairs of neurons (3), where "synchronous" typically refers to spikes that occur within $1 - 10$ ms of each other. One way to test this proposal is to divide responses into finite time bins and compare two codes: one in which the neuronal responses are taken to be the number of spikes in each bin and another in which the neuronal responses are taken to be the number of spikes in each bin plus the number of synchronous spikes across each pair of bins (Fig. 4). This code comparison is typically done by computing information for the two cases (1, 2). Note that this approach is not completely general: first, synchronous spikes may not represent the sole source of correlations; second, if the bins are too large, information is lost to averaging, whereas if they're too small, the number of synchronous spikes is predicted by the number of spikes in each bin, and the difference in information vanishes. Nevertheless, it is a good starting point.

Let us denote the difference in information computed in this way as $\Delta I_{\text{synchronous}}$,

$$\Delta I_{\text{synchronous}} = I(s; \mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_s) - I(s; \mathbf{r}_1, \mathbf{r}_2)$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are vectors of spike counts for neurons 1 and 2, $\mathbf{r}_s$ is a vector of synchronous spikes across the two neurons and s is the stimulus. The components of the vectors correspond to different time bins (see Fig. 4).

How does $\Delta I_{\text{synchronous}}$ compare to ΔI? It turns out that $\Delta I_{\text{synchronous}}$ is the extra number of yes/no questions one would have to ask to guess the stimulus if one did not take

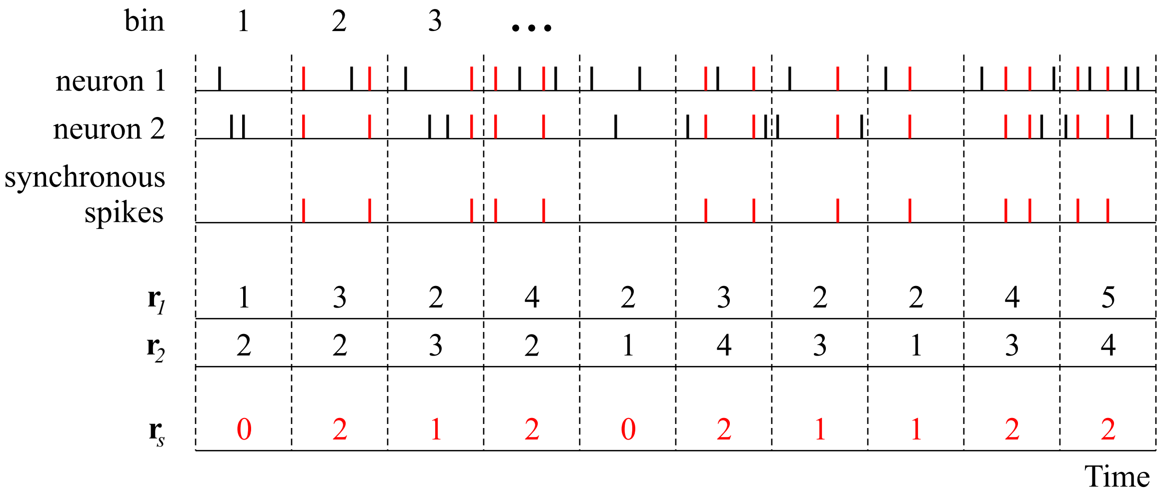

Figure 4: The method used by Dan et al. (1) and Oram et al. (2) to asses the importance of synchronous spikes. Spike trains from two neurons (top two rows) are binned at finite resolution. A third "pseudo-neuron" (third row) is constructed from coincident spikes (the red spikes in the first two rows). In experiments, spikes are considered coincident if their arrival times differ by some small amount, typically $1-10$ ms. Information is then computed for two codes: one consisting of spike counts from the two individual neurons ($\mathbf{r}_1$ and $\mathbf{r}_2$, fourth and fifth rows) and one consisting of these two spike counts plus the number of coincident spikes ($\mathbf{r}_s$, last row).

synchronous spikes into account. Thus, $\Delta I_{\text{synchronous}} = \Delta I$ when both are calculated using the above neural code. Rather than proving this for the specific case of synchronous spikes, we derive a more general result: We show that the difference in information computed using two neural codes is the extra number of yes/no questions it would take to guess the stimuli using one rather than the other.

When we say "code," we mean a choice for the neural response. For example, if we use a spike count code, the choice of response is the number of spikes in some interval; if we use a spike timing code, the choice of response is the arrival time of each spike. Letting $\mathbf{r}'$ and $\mathbf{r}''$ refer to two distinct codes, the difference in information they convey about the stimuli, denoted $\Delta\tilde{I}$, is given by

$$\Delta\tilde{I} \equiv I(s;\mathbf{r}') - I(s;\mathbf{r}'') = \sum_s p(s) \sum_{\mathbf{r}'} p(\mathbf{r}'|s) \log_2 \frac{p(\mathbf{r}'|s)}{p(\mathbf{r}')} - \sum_s p(s) \sum_{\mathbf{r}''} p(\mathbf{r}''|s) \log_2 \frac{p(\mathbf{r}''|s)}{p(\mathbf{r}'')} \,.$$

Using Bayes' theorem, this expression may be rewritten as

$$\Delta\tilde{I} = \sum_s p(s) \left[\sum_{\mathbf{r}'} p(\mathbf{r}'|s) \log_2 p(s|\mathbf{r}') - \sum_{\mathbf{r}''} p(\mathbf{r}''|s) \log_2 p(s|\mathbf{r}'') \right] . \tag{A1}$$

We would like to compare $\Delta\tilde{I}$ to the extra number of yes/no questions it would take to guess the stimulus if one observed $\mathbf{r}'$ instead of $\mathbf{r}''$. This difference, which we denote $\Delta\hat{I}$, can be found from the individual penalties, in yes/no questions, one pays for using $\mathbf{r}'$ and $\mathbf{r}''$ versus the true code, $\mathbf{r}$,

$$\Delta\hat{I} = \sum_s p(s) \sum_{\mathbf{r}} p(\mathbf{r}|s) \log_2 \frac{p(s|\mathbf{r})}{p(s|\mathbf{r}'')} - \sum_s p(s) \sum_{\mathbf{r}} p(\mathbf{r}|s) \log_2 \frac{p(s|\mathbf{r})}{p(s|\mathbf{r}')} .$$

Because $\mathbf{r}$ is the true code, both $\mathbf{r}'$ and $\mathbf{r}''$ must be functions of it, so in this expression $\mathbf{r}' = \mathbf{f}'(\mathbf{r})$ and $\mathbf{r}'' = \mathbf{f}''(\mathbf{r})$. Canceling the term $\log_2 p(s|\mathbf{r})$ that appears in both expressions and making it explicit that $\mathbf{r}'$ and $\mathbf{r}''$ depend on $\mathbf{r}$, we have

$$\Delta\hat{I} = \sum_s p(s) \left[\sum_{\mathbf{r}} p(\mathbf{r}|s) \sum_{\mathbf{r}'} \delta(\mathbf{r}' - \mathbf{f}'(\mathbf{r})) \log_2 p(s|\mathbf{r}') - \sum_{\mathbf{r}} p(\mathbf{r}|s) \sum_{\mathbf{r}''} \delta(\mathbf{r}'' - \mathbf{f}''(\mathbf{r})) \log_2 p(s|\mathbf{r}'') \right] .$$

Here δ is a Kronecker δ-like object: $\delta(\mathbf{r}' - \mathbf{f}'(\mathbf{r})) = 1$ if $\mathbf{r}' = \mathbf{f}'(\mathbf{r})$ and 0 otherwise, and similarly for $\delta(\mathbf{r}'' - \mathbf{f}''(\mathbf{r}))$. Had we been using continuous distributions, the sums would have been integrals and we would have replaced the Kronecker δ with the Dirac δ-function.

Interchanging the order of summation and rearranging terms slightly, we find that

$$\Delta\hat{I} = \sum_s p(s) \left[\sum_{\mathbf{r}'} \log_2 p(s|\mathbf{r}') \sum_{\mathbf{r}} \delta(\mathbf{r}' - \mathbf{f}'(\mathbf{r})) p(\mathbf{r}|s) - \sum_{\mathbf{r}''} \log_2 p(s|\mathbf{r}'') \sum_{\mathbf{r}} \delta(\mathbf{r}'' - \mathbf{f}''(\mathbf{r})) p(\mathbf{r}|s) \right] . \tag{A2}$$

This equation may be simplified by using the standard expression for transforming probability distributions: if, for any function $\mathbf{g}(\mathbf{r})$, $\mathbf{z} = \mathbf{g}(\mathbf{r})$, then

$$p(\mathbf{z}|s) = \sum_{\mathbf{r}} p(\mathbf{r}|s) \delta(\mathbf{z} - \mathbf{g}(\mathbf{r})) .$$

Using this expression, we see that Eq. **A2** is identical to Eq. **A1**. Thus, the difference in information using two codes is identical in the difference in the number of yes/no questions it would take to guess the stimulus using one code rather than the other.

Existence Versus Importance of Correlations

If neurons are not correlated, meaning $p_{\text{ind}}(\mathbf{r}|s) = p(\mathbf{r}|s)$, then $\Delta I = 0$. The converse, however, is not true: $\Delta I = 0 \not\Rightarrow p_{\text{ind}}(\mathbf{r}|s) = p(\mathbf{r}|s)$. In other words, it is possible for neurons to be correlated, meaning $p_{\text{ind}}(\mathbf{r}|s) \neq p(\mathbf{r}|s)$, and still have $\Delta I = 0$. To see why, we use Eqs. **1** and **3** of the main text to write

$$\frac{p_{\text{ind}}(s|\mathbf{r})}{p(s|\mathbf{r})} = \frac{p_{\text{ind}}(\mathbf{r}|s)}{p(\mathbf{r}|s)} \frac{p(\mathbf{r})}{p_{\text{ind}}(\mathbf{r})} \,. \tag{B1}$$

If $\Delta I = 0$, so that $p_{\text{ind}}(s|\mathbf{r}) = p(s|\mathbf{r})$, Eq. **B1** tells us only is that $p_{\text{ind}}(\mathbf{r}|s)/p(\mathbf{r}|s) = p_{\text{ind}}(\mathbf{r})/p(\mathbf{r})$; it does not tell us that $p_{\text{ind}}(\mathbf{r}|s) = p(\mathbf{r}|s)$. Thus, $\Delta I = 0$ does not imply that correlations do not exist.

Of course, Eq. **B1** does not tell us whether a distribution exists that is both correlated and has $\Delta I = 0$. However, it is not hard to find one. Consider an example like the one in Fig. 4, in which the neural code consists of three quantities: the number of spikes from neurons 1 and 2, denoted r_1 and r_2, respectively, and the number of synchronous spikes from the two neurons, denoted r_s. For definiteness, consider a single time bin, so r_1, r_2 and r_s are scalars. For this restricted code, the true distribution, $p(\mathbf{r}|s)$, is simply $p(r_1, r_2, r_s|s)$. It is convenient to write this in the form

$$p(\mathbf{r}|s) = p(r_s|r_1, r_2, s)p(r_1, r_2|s) \,, \tag{B2}$$

which follows immediately from Bayes' theorem. The independent distribution, $p_{\text{ind}}(\mathbf{r}|s)$, is the distribution derived without any knowledge of the correlations between the two neurons; it is thus given by

$$p_{\text{ind}}(\mathbf{r}|s) = p_{\text{ind}}(r_s|r_1, r_2)p(r_1|s)p(r_2|s) \tag{B3}$$

where $p_{\text{ind}}(r_s|r_1, r_2)$ is the probability of observing r_s synchronous spikes given that the neurons produced, independently, r_1 and r_2 spikes.

There is at least one case in which $p_{\text{ind}}(\mathbf{r}|s) \neq p(\mathbf{r}|s)$ even though $p_{\text{ind}}(s|\mathbf{r}) = p(s|\mathbf{r})$. That is the case in which the following two conditions hold: the number of synchronous spikes depends on r_1 and r_2 but not directly on s, meaning $p(r_s|r_1, r_2, s) = p(r_s|r_1, r_2)$, and the spike counts are independent, meaning $p(r_1, r_2|s) = p(r_1|s)p(r_2|s)$. Combining these conditions with Eq. **B2**, we see that

$$p(\mathbf{r}|s) = p(r_s|r_1, r_2)p(r_1|s)p(r_2|s)\,. \tag{B4}$$

To determine whether the neurons are correlated, we use Eqs. **B3** and **B4** to write

$$\frac{p_{\text{ind}}(\mathbf{r}|s)}{p(\mathbf{r}|s)} = \frac{p_{\text{ind}}(r_s|r_1, r_2)}{p(r_s|r_1, r_2)}\,.$$

As long as the synchronous spikes occur above chance, so that $p_{\text{ind}}(r_s|r_1, r_2) \neq p(r_s|r_1, r_s)$, then the neurons are correlated. However, the degree to which they are correlated is stimulus-independent, which implies, via Eq. **B1** and a small amount of algebra, that $p_{\text{ind}}(s|\mathbf{r}) = p(s|\mathbf{r})$, which in turn implies that $\Delta I = 0$. Thus, in this case, even though correlations exist, they are not important for decoding.

Evaluation of Other Measures for Decoding

In this section we show by construction that $\Delta I_{\text{shuffled}}$ and $\Delta I_{\text{synergy}}$ can be positive, negative, or zero both when ΔI is zero and when it is positive. This implies that neither $\Delta I_{\text{shuffled}}$ nor $\Delta I_{\text{synergy}}$ are reliable indicators of the importance of correlations for decoding.

Our program is to simply write down a probability distribution and compute the relevant information theoretic quantities. The probability distribution we use is given in Table 1, and the relevant information theoretic quantities are I, ΔI, $\Delta I_{\text{shuffled}}$ and $\Delta I_{\text{synergy}}$.

A feature of the distribution in Table 1 is that the stimuli and responses fall into disjoint classes: stimuli s_1 and s_2 produce responses between 0 and 2 but never above 2; stimuli $s_3 - s_5$ produce responses between 3 and 5 but never below 3. Consequently, we can write this distribution as

$$p(s, \mathbf{r}) = \rho p_1(s, \mathbf{r}) + (1 - \rho)p_2(s, \mathbf{r})$$

Table 1. Probability distribution used to show that $\Delta I_{\text{shuffled}}$ and $\Delta I_{\text{synergy}}$ are not directly related to ΔI. This distribution depends on three quantities, α, β, and ρ; the range of all of them is from 0 to 1.

stim	r_1	r_2	$p(s)$	$p(r_1, r_2 \vert s)$	$p_{\text{ind}}(r_1, r_2 \vert s)$	$p(s \vert r_1, r_2)$	$p_{\text{ind}}(s \vert r_1, r_2)$
s_1	0	0	$\rho/2$	$\alpha/2$	1/4	1	1
	0	1		$(1-\alpha)/2$	1/4	1	1
	1	0		$(1-\alpha)/2$	1/4	1	1
	1	1		$\alpha/2$	1/4	$\alpha/(\alpha+\beta)$	1/2
s_2	1	1	$\rho/2$	$\beta/2$	1/4	$\beta/(\alpha+\beta)$	1/2
	1	2		$(1-\beta)/2$	1/4	1	1
	2	1		$(1-\beta)/2$	1/4	1	1
	2	2		$\beta/2$	1/4	1	1
s_3	4	5	$(1-\rho)/3$	1/2	1/4	1	1
	5	4		1/2	1/4	1	1
s_4	5	3	$(1-\rho)/3$	1/2	1/4	1	1
	3	5		1/2	1/4	1	1
s_5	3	4	$(1-\rho)/3$	1/2	1/4	1	1
	4	3		1/2	1/4	1	1

where ρ is the probability of observing stimuli $s_1 - s_2$ and responses $0 - 2$, $1 - \rho$ is the probability of observing stimuli $s_3 - s_5$ and responses $3 - 5$, and both $p_1(s, \mathbf{r})$ and $p_2(s, \mathbf{r})$ are normalized to 1. For distributions of this form, it's not hard to show that

$$\begin{aligned} I &= \rho I_1 + (1-\rho) I_2 + h(\rho) && \text{(C1a)} \\ \Delta I &= \rho \Delta I_1 + (1-\rho) \Delta I_2 && \text{(C1b)} \\ \Delta I_{\text{shuffled}} &= \rho \Delta I_{\text{shuffled},1} + (1-\rho) \Delta I_{\text{shuffled},2} && \text{(C1c)} \\ \Delta I_{\text{synergy}} &= \rho \Delta I_{\text{synergy},1} + (1-\rho) \Delta I_{\text{synergy},2} - h(\rho) && \text{(C1d)} \end{aligned}$$

where $\Delta I_{\text{synergy,k}}$, ΔI_k, and I_k are computed using $p_k(s, \mathbf{r})$ $(k = 1, 2)$, and

$$h(\rho) \equiv -\rho \log_2 \rho - (1-\rho) \log_2 (1-\rho)$$

is the entropy of a binary variable with probability ρ.

Using Eq. **C1**, it is straightforward (although somewhat tedious) to show that

$$\begin{aligned}
I(\alpha,\beta,\rho) &= \rho\,\frac{4+\alpha\log_2\alpha+\beta\log_2\beta-(\alpha+\beta)\log_2(\alpha+\beta)}{4} \qquad \text{(C2a)}\\
&+ (1-\rho)\log_2 3+h(\rho)\\
\Delta I(\alpha,\beta,\rho) &= \rho\,\frac{\alpha+\beta+\alpha\log_2\alpha+\beta\log_2\beta-(\alpha+\beta)\log_2(\alpha+\beta)}{4} \qquad \text{(C2b)}\\
\Delta I_{\text{shuffled}}(\alpha,\beta,\rho) &= \rho\,\frac{1+\alpha\log_2\alpha+\beta\log_2\beta-(\alpha+\beta)\log_2(\alpha+\beta)}{4}+\frac{1-\rho}{2} \qquad \text{(C2c)}\\
\Delta I_{\text{synergy}}(\alpha,\beta,\rho) &= \rho\,\frac{\alpha\log_2\alpha+\beta\log_2\beta-(\alpha+\beta)\log_2(\alpha+\beta)}{4} \qquad \text{(C2d)}\\
&+ (1-\rho)\log_2(4/3)-h(\rho)\,.
\end{aligned}$$

We now look at several special cases.

Case 1: $\Delta I_{\text{shuffled}}$ can be positive, negative, or zero when $\Delta I = 0$. Let $\alpha=\beta$ and $\rho=1$. The quantities in Eq. **C2** then simplify to

$$\begin{aligned}
I(\alpha,\alpha,1) &= \frac{2-\alpha}{2} \qquad \text{(C3a)}\\
\Delta I(\alpha,\alpha,1) &= 0 \qquad \text{(C3b)}\\
\Delta I_{\text{shuffled}}(\alpha,\alpha,1) &= \frac{1-2\alpha}{4}\,. \qquad \text{(C3c)}
\end{aligned}$$

Because the range of α is from 0 to 1, $\Delta I_{\text{shuffled}}$ can range from $-1/4$ to $1/4$. In Fig. 5a we plot $\Delta I_{\text{shuffled}}/I$ versus α.

Case 2: $\Delta I_{\text{shuffled}}$ can be positive, negative, or zero when $\Delta I > 0$. Let $\beta=1$ and $\rho=1$. The quantities in Eq. **C2** then simplify to

$$\begin{aligned}
I(\alpha,1,1) &= \frac{4+\alpha\log_2\alpha-(1+\alpha)\log_2(1+\alpha)}{4} \qquad \text{(C4a)}\\
\Delta I(\alpha,1,1) &= \frac{1+\alpha+\alpha\log_2\alpha-(1+\alpha)\log_2(1+\alpha)}{4} \qquad \text{(C4b)}\\
\Delta I_{\text{shuffled}}(\alpha,1,1) &= \frac{1+\alpha\log_2\alpha-(1+\alpha)\log_2(1+\alpha)}{4}\,. \qquad \text{(C4c)}
\end{aligned}$$

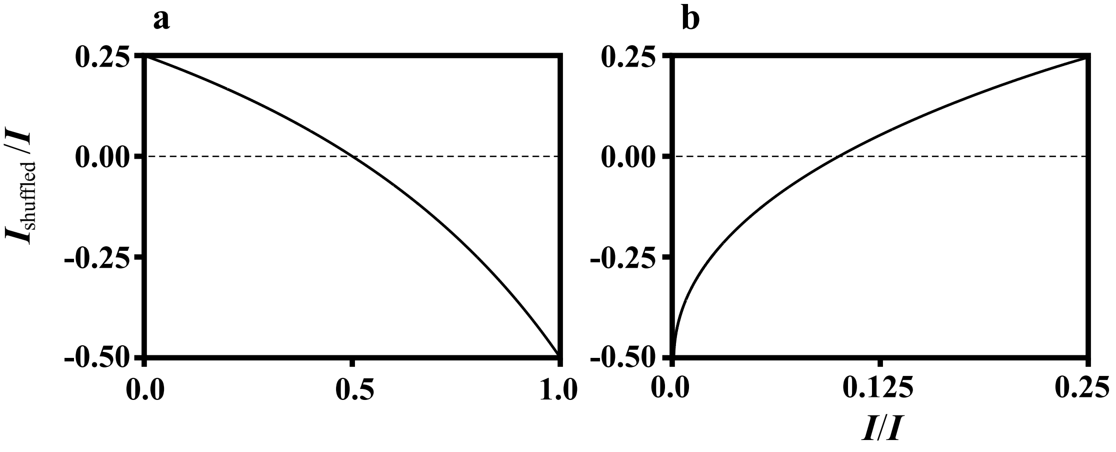

Figure 5: $\Delta I_{\text{shuffled}}$ isn't reliably related to ΔI. (*a*) Case 1: $\Delta I_{\text{shuffled}}/I$ versus α with $\beta = \alpha$ and $\rho = 1$, for which $\Delta I = 0$. (*b*) Case 2: $\Delta I_{\text{shuffled}}/I$ versus $\Delta I/I$ as α ranges from 0 to 1 with $\beta = \rho = 1$.

From these expressions, we see that when α ranges from 0 to 1, $\Delta I_{\text{shuffled}}$ ranges from 1/4 to $-1/4$, ΔI ranges from 1/4 to 0, and I ranges from 1 to 1/2. In Fig. 5b we plot $\Delta I_{\text{shuffled}}/I$ against $\Delta I/I$ versus α for α between 0 and 1. This plot shows that when $\Delta I > 0$, $\Delta I_{\text{shuffled}}$ can be positive, negative, or zero.

Case 3: $\Delta I_{\text{synergy}}$ can be positive, negative, or zero when $\Delta I = 0$. Let $\alpha = \beta = 1$. The quantities in Eq. **C2** then simplify to

$$\begin{aligned} I(1,1,\rho) &= \frac{\rho}{2} + (1-\rho)\log_2 3 + h(\rho) && \text{(C5a)} \\ \Delta I(1,1,\rho) &= 0 && \text{(C5b)} \\ \Delta I_{\text{synergy}}(1,1,\rho) &= -\frac{\rho}{2} + (1-\rho)\log_2(4/3) - h(\rho)\,. && \text{(C5c)} \end{aligned}$$

When $\rho = 0$, $\Delta I_{\text{synergy}} = \log_2(4/3) \approx 0.42$, whereas when $\rho = 1/2$, $\Delta I_{\text{synergy}} = \log_2(4/3)/2 - 5/4 \approx -1.0$; for the corresponding values of ρ, $I = \log_2 3 \approx 1.58$ and $(\log_2 3)/2 + 5/4 \approx 2.04$, respectively. In Fig. 6a we plot $\Delta I_{\text{synergy}}/I$ versus ρ for ρ between 0 and 1/2.

Case 4: $\Delta I_{\text{synergy}}$ can be positive, negative, or zero when $\Delta I > 0$. Let $\alpha = 1$ and $\beta = 0$. The quantities in Eq. **C2** then simplify to

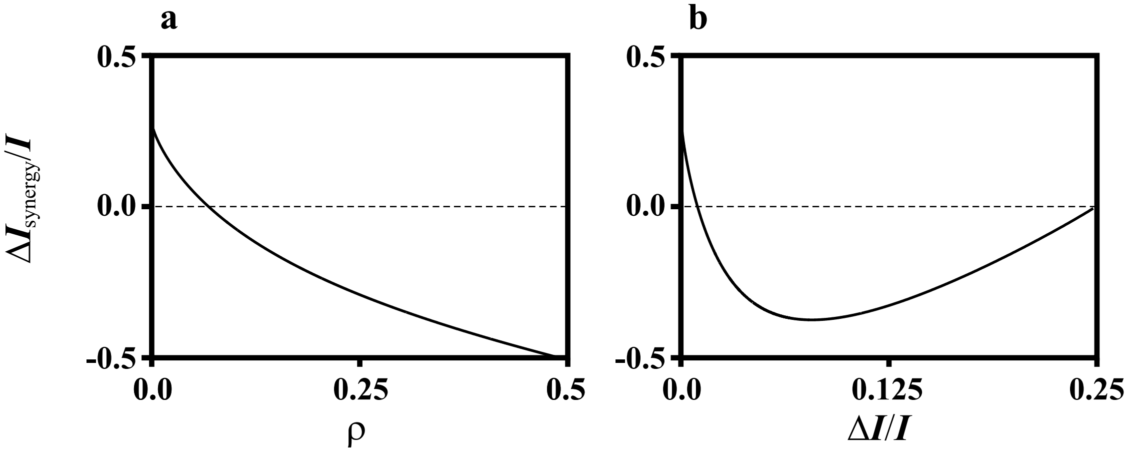

Figure 6: $\Delta I_{\text{synergy}}$ isn't reliably related to ΔI. (a) Case 3: $\Delta I_{\text{synergy}}/I$ versus ρ with $\alpha = \beta = 1$, for which $\Delta I = 0$. (b) Case 4: $\Delta I_{\text{synergy}}/I$ versus $\Delta I/I$ as ρ runs from 0 to 1 with $\alpha = 0$ and $\beta = 1$.

$$\begin{aligned} I(0,1,\rho) &= \rho + (1-\rho)\log_2 3 + h(\rho) && \text{(C6a)} \\ \Delta I(0,1,\rho) &= \frac{\rho}{4} && \text{(C6b)} \\ \Delta I_{\text{synergy}}(0,1,\rho) &= (1-\rho)\log_2(4/3) - h(\rho)\,. && \text{(C6c)} \end{aligned}$$

From these expressions, we see that when ρ ranges from 0 to 1, $\Delta I_{\text{synergy}}$ ranges (non-monotonically) from $\log_2(4/3)$ to 0, ΔI ranges from 0 to 1/4, and I ranges from $\log_2 3$ to 1. In Fig. 6b we plot $\Delta I_{\text{synergy}}/I$ against $\Delta I/I$ versus α for α between 0 and 1. This plot shows that when $\Delta I > 0$, $\Delta I_{\text{synergy}}$ can be positive, negative, or zero.

Taken together, Figs. 5 and 6 show that both $\Delta I_{\text{shuffled}}$ and $\Delta I_{\text{synergy}}$ can be positive, negative, or zero when $\Delta I = 0$ and when $\Delta I > 0$.

References

1. Dan, Y., Alonso, J. M., Usrey, W. M. & Reid, R. C. (1998) *Nat. Neurosci.* **1**, 501-507.
2. Oram, M. W., Hatsopoulos, N. G., Richmond, B. J. & Donoghue, J. P. (2001) *J. Neurophysiol.* **86**, 1700-1716.
3. For reviews, see *Neuron* (1999), **24**, 7-126.