Supporting Information: Zonal receptor distributions maximize olfactory information

Swati Sen and David Zwicker

I. MUTUAL INFORMATION

We here analyze the mutual information I, given by Eq. 12 in the main text, for uncorrelated inputs and noises. Assuming Gaussian input and noise distributions, $P_{\mathrm{env}}(\boldsymbol{c}) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}^{\mathrm{c}})$ and $P_\xi(\boldsymbol{\xi}) = \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}^{\xi})$, the distribution of the output $\boldsymbol{e}$, defined by the linear map given in Eq. 10 in the main text, is also Gaussian and reads $P(\boldsymbol{e}) = \mathcal{N}(\boldsymbol{A}\boldsymbol{\mu}^{\mathrm{c}}, \boldsymbol{\Sigma}^{\mathrm{e}})$ with $\boldsymbol{\Sigma}^{\mathrm{e}} = \boldsymbol{A}\boldsymbol{\Sigma}^{\mathrm{c}}\boldsymbol{A}^T + \boldsymbol{\Sigma}^{\xi}$ [1]. Similarly, we find a Gaussian joint distribution $P(\boldsymbol{c}, \boldsymbol{e}) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}^{\mathrm{c}} \\ \boldsymbol{A}\boldsymbol{\mu}^{\mathrm{c}} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}^{\mathrm{c}} & \boldsymbol{\Sigma}^{\mathrm{c}}\boldsymbol{A}^T \\ \boldsymbol{A}\boldsymbol{\Sigma}^{\mathrm{c}} & \boldsymbol{A}\boldsymbol{\Sigma}^{\mathrm{c}}\boldsymbol{A}^T + \boldsymbol{\Sigma}^{\xi} \end{bmatrix} . \tag{S1}$$

Hence, the mutual information reduces to [2]

$$I = \frac{1}{2}\log_2\left(\frac{|\boldsymbol{\Sigma}^{\mathrm{c}}||\boldsymbol{\Sigma}^{\mathrm{e}}|}{|\boldsymbol{\Sigma}|}\right) . \tag{S2}$$

For the simple case where the matrices $\boldsymbol{\Sigma}^{\mathrm{c}}$, $\boldsymbol{A}$, and $\boldsymbol{\Sigma}^{\xi}$ are diagonal, we have

$$I_{\mathrm{diag}} = \frac{1}{2}\sum_{i=1}^{N_{\mathrm{r}}}\log_2\left(1 + A_{ii}^2\frac{\Sigma^{\mathrm{c}}_{ii}}{\Sigma^{\xi}_{ii}}\right) . \tag{S3}$$

Using the definition of A_{ii} given in Eq. 11 in the main text, we find Eq. 13, where $\zeta_i = \alpha_i S_{ii}[\Sigma^{\mathrm{c}}_{ii}/\Sigma^{\xi}_{ii}]^{1/2}$ quantifies the signal-to-noise ratio of a single receptor.

We start by considering the case of high signal-to-noise ($\zeta_i \gg 1$). In this case, Eq. 13 in the main text becomes

$$I_{\mathrm{diag,H}} \approx \sum_{i=1}^{N_{\mathrm{r}}}\log_2\zeta_i + \sum_{i=1}^{N_{\mathrm{r}}}\log_2\left(\int_0^L n_i e^{-\frac{z}{\lambda_i}}\,\mathrm{d}z\right) , \tag{S4}$$

so we can maximize the information irrespective of ζ_i. In particular, we consider the situation where receptors are sorted into distinct zones and optimize $I_{\mathrm{diag,H}}$ with respect to the zone length. In the simplest case of $N_{\mathrm{r}} = 2$ types of receptors, we have

$$n_1(z) = \begin{cases} n_{\max} & 0 < z < L_1 \\ 0 & \text{otherwise} \end{cases} \tag{S5a}$$

$$n_2(z) = \begin{cases} n_{\max} & L_1 < z < L \\ 0 & \text{otherwise} \end{cases} \tag{S5b}$$

where we assumed $\lambda_1 < \lambda_2$ and $n_{\max} = N/L$. Optimizing $I_{\mathrm{diag,H}}$ with respect to L_1, we find the optimal lengths

$$L_1^*(\lambda_1, \lambda_2) = \frac{2\lambda_1\lambda_2(1 - e^{L/\lambda_2})}{\lambda_2 - (2\lambda_1 + \lambda_2)e^{L/\lambda_2}} \approx \frac{2\lambda_1\lambda_2}{2\lambda_1 + \lambda_2} \tag{S6}$$

and $L_2^* = L - L_1^*$.

In the opposite limit of low signal-to-noise ($\zeta_i N_i \ll 1$) we consider for simplicity $\lambda_i \to \infty$, so Eq. 13 in the main text becomes

$$I_{\mathrm{diag,L}} \approx \frac{1}{2\ln 2}\sum_i N_i^2\zeta_i^2 , \tag{S7}$$

where the total number of receptors is $N = \sum_i N_i$. We then ask when it is useful to add an additional receptor type to an existing array. For simplicity, we consider an array of N_{r} receptor types with equal $\zeta_i = \zeta$ for $i = 1, \ldots, N_{\mathrm{r}}$ and add a receptor type with potentially different signal-to-noise ζ_{add}. Using $N_i = (N - N_{\mathrm{add}})/N_{\mathrm{r}}$ for $i = 1, \ldots, N_{\mathrm{r}}$, we obtain the optimal count of the added receptor, $N^*_{\mathrm{add}}(N, \zeta, \zeta_{\mathrm{add}})$, from the optimality condition $\partial I_{\mathrm{diag,L}}/\partial N_{\mathrm{add}} = 0$,

$$N^*_{\mathrm{add}} = \frac{N_{\mathrm{r}} N\zeta^2({N_{\mathrm{r}}}^2 + N^2\zeta^2)}{N_{\mathrm{r}}\zeta^2({N_{\mathrm{r}}}^2 - N^2\zeta^2) + \zeta^2_{\mathrm{add}}({N_{\mathrm{r}}}^2 + N^2\zeta^2)^2} , \tag{S8}$$

which is valid for $\zeta_{\mathrm{add}} > \zeta^{\mathrm{th}}_{\mathrm{add}}$. The threshold value $\zeta^{\mathrm{th}}_{\mathrm{add}}$ follows from the pole of Eq. (S8) and reads

$$\zeta^{\mathrm{th}}_{\mathrm{add}} = \frac{\zeta\sqrt{N^2 N_{\mathrm{r}}\zeta^2 - {N_{\mathrm{r}}}^3}}{N^2\zeta^2 + {N_{\mathrm{r}}}^2} . \tag{S9}$$

Since $\zeta^{\mathrm{th}}_{\mathrm{add}}$ needs to be real, we require $N\zeta > N_{\mathrm{r}}$.

[1] Y. Tong, The Multivariate Normal Distribution, Research in Criminology (Springer New York, 1990).

[2] G. Forney and G. Ungerboeck, IEEE Transactions on Information Theory **44**, 2384 (1998).