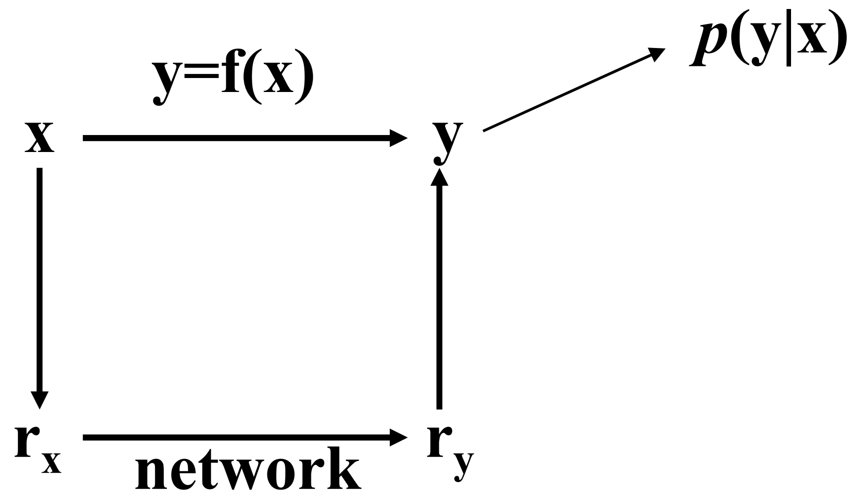
Probabilistic inference in networks of spiking neurons

Peter Latham
Gatsby Computational Neuroscience Unit, UCL

Kavli Institute for Theoretical Physics
October 7, 2010

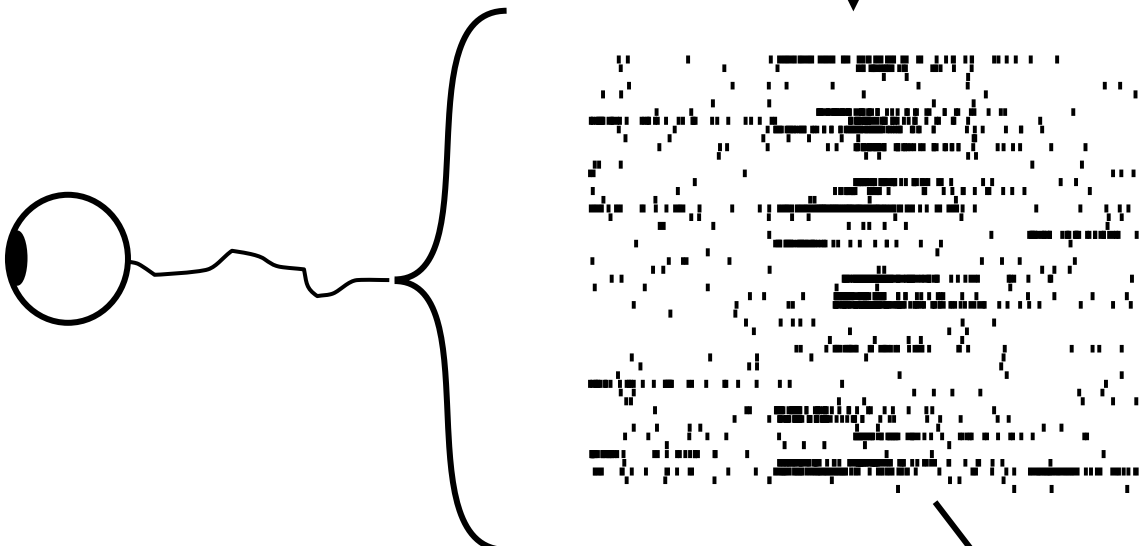
learn to

Goal: understand how networks of neurons compute.



$\mathbf{r_x}$

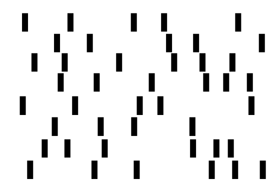
x = image



y = sunset

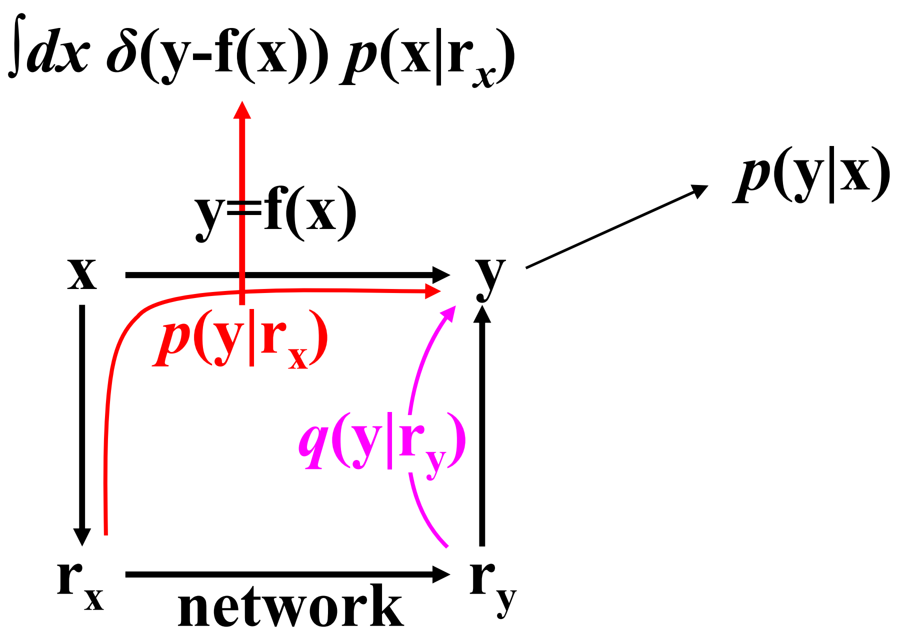
network

true probability: $p(y|\mathbf{r_x})$



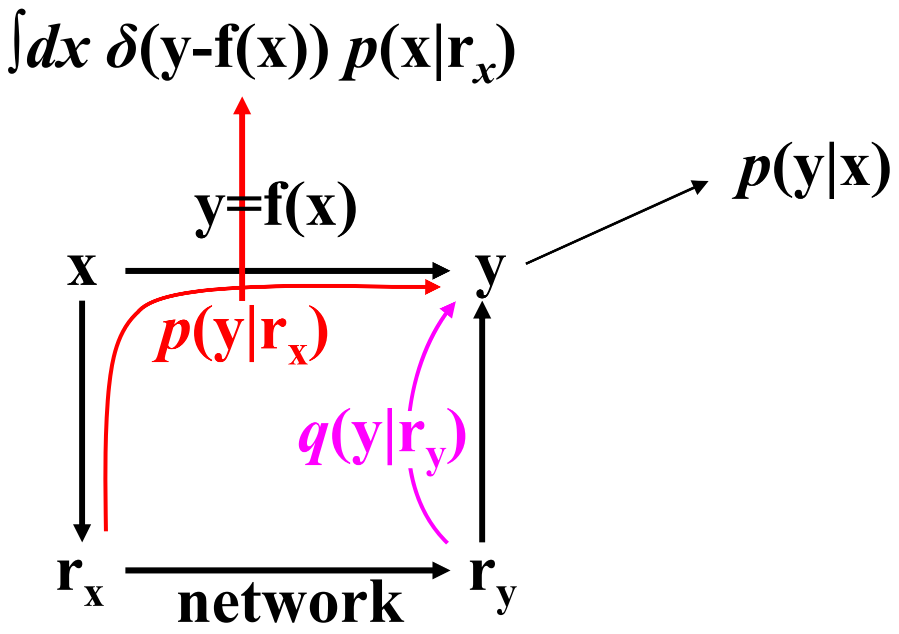
$q(y|\mathbf{r}_y)$ ← $\mathbf{r}_y$ →

goal: choose the network so these are as close as possible, and $\mathbf{r}_y$ codes only for sunsets.



Goal: choose network so that $p(y|r_x)$ is close to $q(y|r_y)$ and r_y is a "pure code" for y.

Quantitatively: minimize $D_{KL}(p(y|r_x)||q(y|r_y))$



minimize $D_{KL}(p(y|r_x)||q(y|r_y))$ with respect to:

1. **Parameters of the encoding model (e.g., $p(r_x|x)$, $q(r_y|y)$).**
2. **Parameters of the <u>biologically plausible</u> network.**

Three examples:

- **multisensory integration**
- $z = f(x, y)$
- **a hard problem**

minimize $D_{KL}(p(y|r_x)||q(y|r_y))$ **with respect to:**

1. **Parameters of the encoding model (e.g.,** $p(r_x|x)$, $q(r_y|y)$**).**
2. **Parameters of the biologically plausible network.**

Take home message

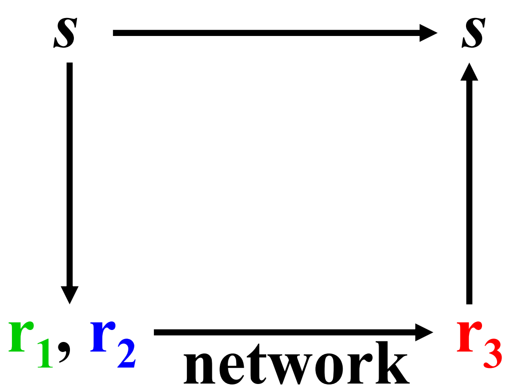
In some ways this talk is largely technical – I'm going to tell you how biologically plausible networks could perform some (relatively) simple computations.

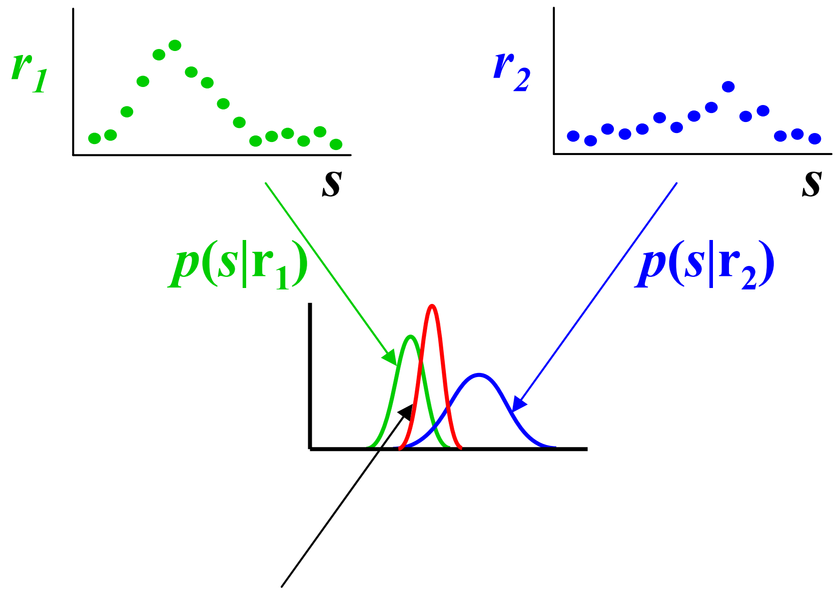
There are no deep insights into how the brain works.

Two things to pay attention to:

1. our methodology,
2. the problems we don't solve.

1. multisensory integration





$$p(s|r_1, r_2) \propto p(s|r_1)p(s|r_2)$$

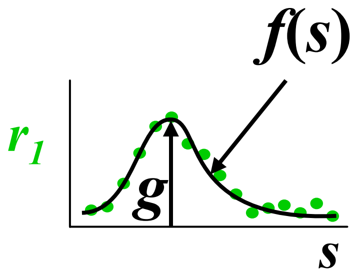
$$\text{minimize } D_{KL}(p(s|r_1, r_2)||q(s|r_3))$$

encoding model: exponential family with linear sufficient statistics

independent Poisson:

$$p(\mathbf{r}|s, g) = \prod_i \frac{1}{r_i!} [g f_i(s)]^{r_i} \exp[-g f_i(s)]$$

$$= \varphi(\mathbf{r}, g) \exp[\mathbf{h}(s) \cdot \mathbf{r}]$$



$$h_i(s) = \log f_i(s)$$

$$\varphi(\mathbf{r}, g) = \exp\left[-\sum_i \log r_i! + r_i \log g - g f_i(s)\right]$$

$$\exp[r_i \log g + r_i \log f_i(s)]$$

encoding model: <u>exponential family with linear sufficient statistics</u>

$$p(\mathbf{r}|s, g) = \varphi(\mathbf{r}, g) \exp[\mathbf{h}(s)\cdot\mathbf{r}]$$

$$p(s|\mathbf{r}) \propto \exp[\mathbf{h}(s)\cdot\mathbf{r}]$$

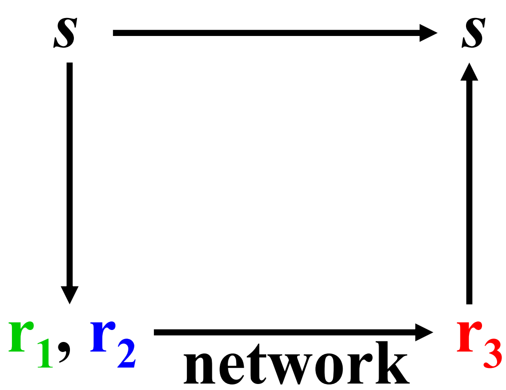
the gain parameter doesn't appear in the posterior!!!!

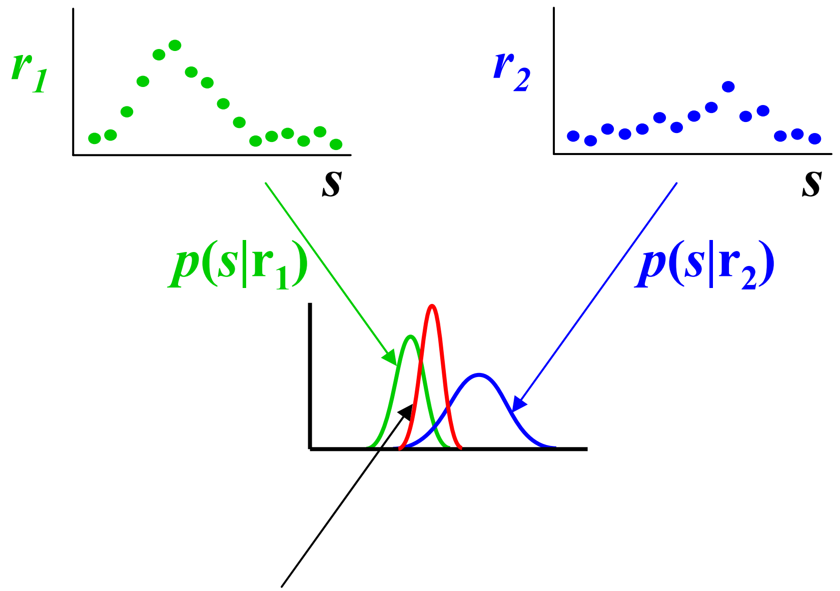
encoding model: linear PPC

$$p(\mathbf{r}|s, g) = \varphi(\mathbf{r}, g) \exp[\mathbf{h}(s)\cdot\mathbf{r}]$$

$$p(s|\mathbf{r}) \propto \exp[\mathbf{h}(s)\cdot\mathbf{r}]$$

the gain parameter doesn't appear in the posterior!!!!





$$p(s|\mathrm{r}_1, \mathrm{r}_2) \propto p(s|\mathrm{r}_1)p(s|\mathrm{r}_2)$$

$$\text{minimize } D_{KL}(p(s|\mathrm{r}_1, r_2)||q(s|\mathrm{r}_3))$$

minimize $D_{KL}(p(s|r_1, r_2)||q(s|r_3))$

$p(s|r_1, r_2) \propto p(s|r_1)p(s|r_2)$

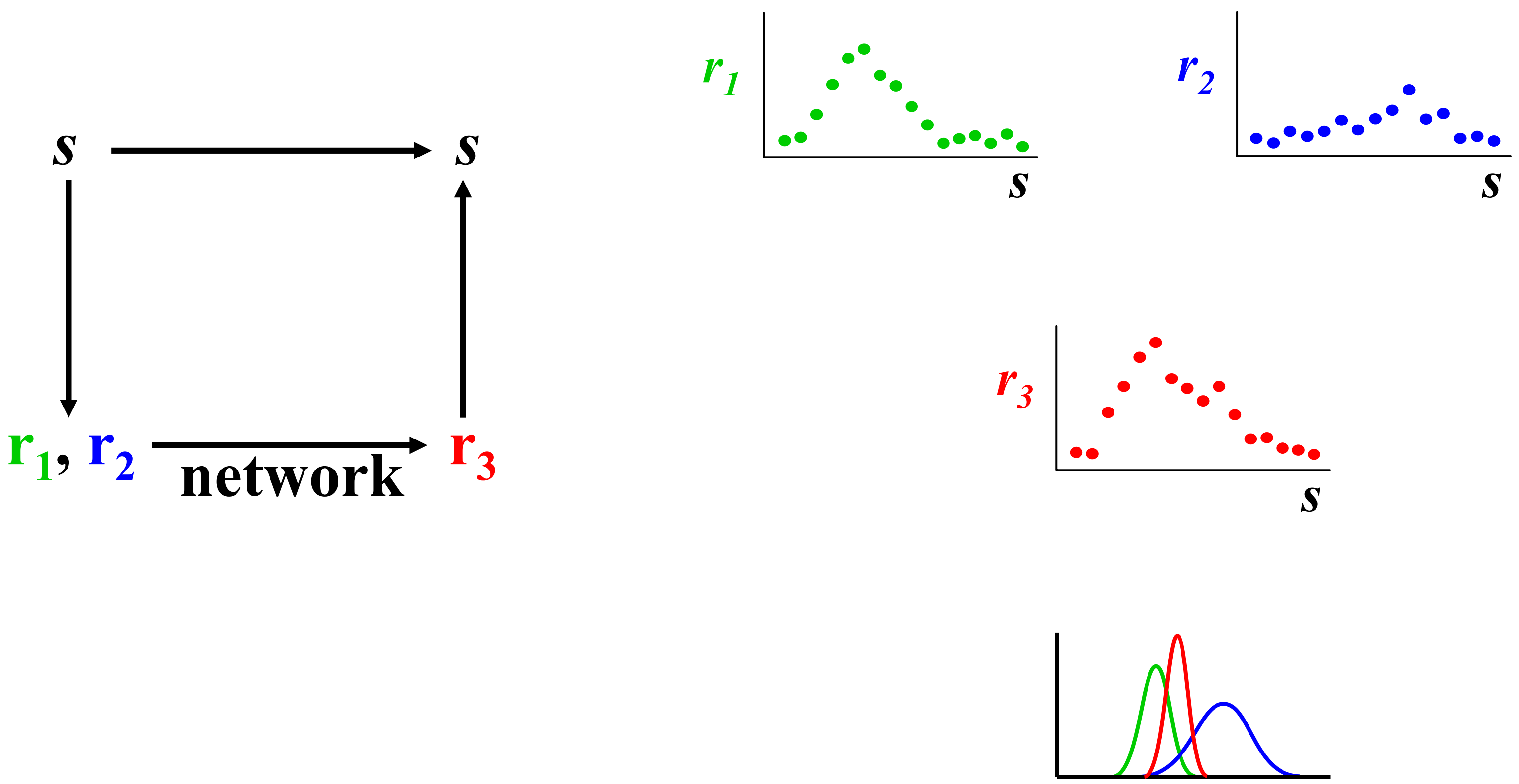
$p(s|r_1) \propto \exp[h(s)\cdot r_1]$

$p(s|r_2) \propto \exp[h(s)\cdot r_2]$

$q(s|r_3) \propto \exp[h(s)\cdot r_3]$

$p(s|r_1, r_2) \propto p(s|r_1)p(s|r_2) \propto \exp[h(s)\cdot(r_1+r_2)]$

network: $r_3 = r_1+r_2$



Ma et al., *Nature Neurosci.* (2006)

The encoding model matters!

Gaussian encoding model:

$$p(r_1|s) \propto \exp[-(f(s)-r_1)^2/2\sigma_1^2]$$
$$p(r_2|s) \propto \exp[-(f(s)-r_2)^2/2\sigma_2^2]$$

these will vary from trial to trial

$$p(s|r_1, r_2) \propto p(s|r_1)p(s|r_2) \propto \exp[-(f(s)-ar_1-br_2)^2/2\sigma^2]$$

$$1/\sigma^2 = 1/\sigma_1^2 + 1/\sigma_2^2$$
$$a = \sigma^2/\sigma_1^2$$
$$b = \sigma^2/\sigma_2^2$$

network: $r_3 = ar_1 + br_2$

depend on σ_1 and σ_2, which depend, probabilistically, on r_1 and r_2.

The encoding model matters!

linear PPC encoding:

$$p(\mathbf{r}_1|s) \propto \varphi(\mathbf{r}_1, \mathbf{g}_1) \exp[\mathbf{h}(s)\cdot\mathbf{r}_1]$$
$$p(\mathbf{r}_2|s) \propto \varphi(\mathbf{r}_2, \mathbf{g}_2) \exp[\mathbf{h}(s)\cdot\mathbf{r}_2]$$

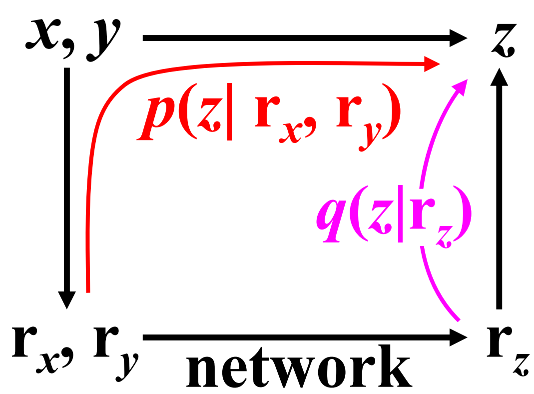
these will vary from trial to trial

network: $\mathbf{r}_3 = \mathbf{r}_1 + \mathbf{r}_2$

This was a simple example, but it has the main ingredients:

1. **The encoding model matters.**
2. **Once you specify the encoding model, the network follows.**

2. computing functions $z=f(x, y)$



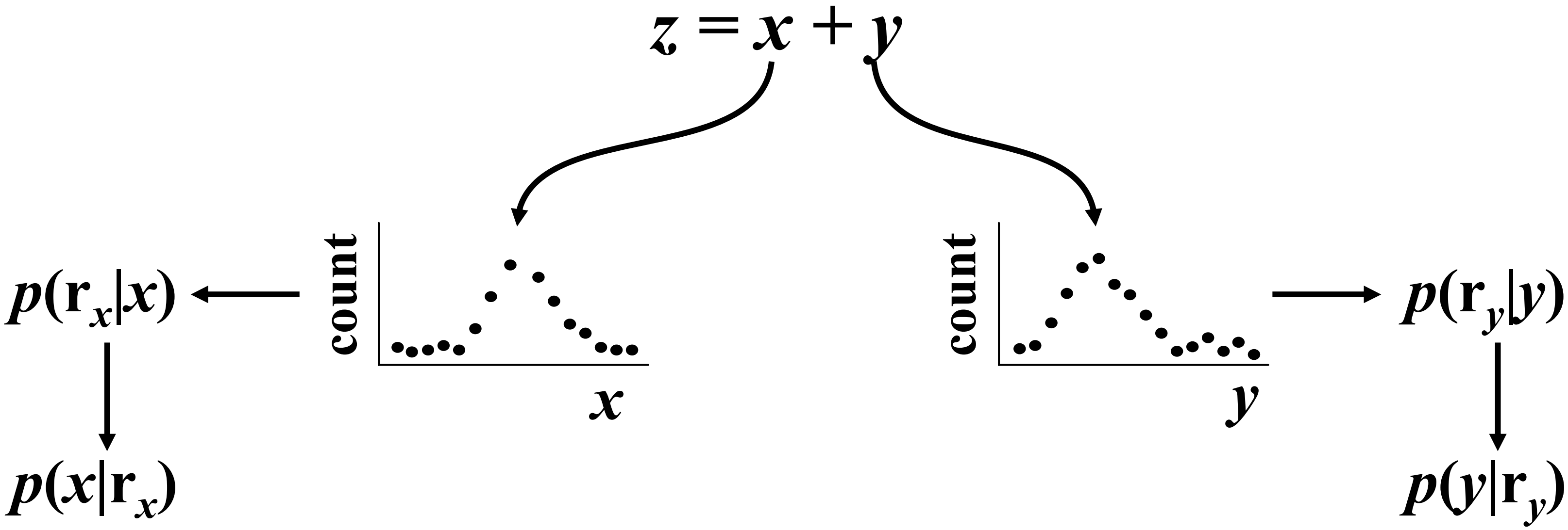
minimize $D_{KL}(p(z|r_x, r_y)||q(z|r_z))$ with respect to:

1. **Parameters of the encoding model.**
2. **Parameters of the network.**

typically, we can't find the optimal network

Find the optimal network for linear transformations and Gaussian posterior distributions;

use that network architecture for nonlinear transformation and non-Gaussian posteriors.



$$p(z|\mathrm{r}_x, \mathrm{r}_y) = \int dx dy\ \delta(z-(x+y))\ p(x|\mathrm{r}_x)\ p(y|\mathrm{r}_y)$$

hard integral!

$$z = x + y$$

The Gaussian case

$$p(z|\mathbf{r}_x, \mathbf{r}_y) = \int dxdy\, \delta(z-(x+y))\, p(x|\mathbf{r}_x)\, p(y|\mathbf{r}_y)$$

$p(z|\mathbf{r}_x, \mathbf{r}_y)$: $N(\mu_x+\mu_y, \sigma_x^2 + \sigma_y^2)$

$p(x|\mathbf{r}_x)$: $N(\mu_x,\sigma_x^2)$

$p(y|\mathbf{r}_y)$: $N(\mu_y,\sigma_y^2)$

Easier problem:

1. parameterize the mean and variance in a linear PPC;
2. find a network such that

$$\mu_z = \mu_x + \mu_y$$
$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2$$

An (important) aside: representation matters Suppose

$p(r_x|x) \sim \exp[-(x-a \cdot r_x)^2/2b \cdot r_x] \sim p(x|r_x)$

$p(r_y|y) \sim \exp[-(y-a \cdot r_y)^2/2b \cdot r_y] \sim p(y|r_y)$

$p(r_z|z) \sim \exp[-(z-a \cdot r_z)^2/2b \cdot r_z] \sim p(z|r_z)$

$\mu_x = a \cdot r_x$
$\sigma_x^2 = b \cdot r_x$

$\mu_y = a \cdot r_y$
$\sigma_y^2 = b \cdot r_y$

$\mu_z = a \cdot r_z$
$\sigma_z^2 = b \cdot r_z$

An (important) aside: representation matters Suppose

$$p(\mathrm{r}_x|x) \sim \exp[-(x-\mathrm{a}\cdot\mathrm{r}_x)^2/2\mathrm{b}\cdot\mathrm{r}_x] \sim p(x|\mathrm{r}_x)$$

$$p(\mathrm{r}_y|y) \sim \exp[-(y-\mathrm{a}\cdot\mathrm{r}_y)^2/2\mathrm{b}\cdot\mathrm{r}_y] \sim p(y|\mathrm{r}_y)$$

$$p(\mathrm{r}_z|z) \sim \exp[-(z-\mathrm{a}\cdot\mathrm{r}_z)^2/2\mathrm{b}\cdot\mathrm{r}_z] \sim p(z|\mathrm{r}_z)$$

$$\mu_x = \mathrm{a}\cdot\mathrm{r}_x$$
$$\sigma_x^2 = \mathrm{b}\cdot\mathrm{r}_x$$

$$\mu_y = \mathrm{a}\cdot\mathrm{r}_y$$
$$\sigma_y^2 = \mathrm{b}\cdot\mathrm{r}_y$$

$$\mu_z = \mathrm{a}\cdot\mathrm{r}_z = \mu_x + \mu_y$$
$$\sigma_z^2 = \mathrm{b}\cdot\mathrm{r}_z = \sigma_x^2 + \sigma_y^2$$

goal

An (important) aside: representation matters. Suppose

$$p(r_x|x) \sim \exp[-(x - a \cdot r_x)^2/2b \cdot r_x] \sim p(x|r_x)$$

$$p(r_y|y) \sim \exp[-(y - a \cdot r_y)^2/2b \cdot r_y] \sim p(y|r_y)$$

$$p(r_z|z) \sim \exp[-(z - a \cdot r_z)^2/2b \cdot r_z] \sim p(z|r_z)$$

$$\mu_x = a \cdot r_x$$
$$\sigma_x^2 = b \cdot r_x$$

$$\mu_y = a \cdot r_y$$
$$\sigma_y^2 = b \cdot r_y$$

$$\mu_z = a \cdot r_z = \mu_x + \mu_y = a \cdot r_x + a \cdot r_y$$
$$\sigma_z^2 = b \cdot r_z = \sigma_x^2 + \sigma_y^2 = b \cdot r_x + b \cdot r_y$$

goal

optimal network:

$$r_z = r_x + r_y$$

If brains had evolved to perform linear transformations ($z=ax+by$), and noise was Gaussian, then encoding probably would have looked like

$$p(\mathbf{r}_x|x) \sim \exp[-(x-\mathbf{a}\cdot\mathbf{r}_x)^2/2\mathbf{b}\cdot\mathbf{r}_x].$$

The fact that the encoding model does not look like this is probably a clue to what the brain has evolved to compute.

So far this is just a clue – we haven't made sense of it.

An (important) aside: representation matters! Suppose

$$p(\mathbf{r}_x|x) \sim \exp[-(x-\mathbf{a}\cdot\mathbf{r}_x)^2/2\mathbf{b}\cdot\mathbf{r}_x] \sim p(x|\mathbf{r}_x)$$

$$p(\mathbf{r}_y|y) \sim \exp[-(y-\mathbf{a}\cdot\mathbf{r}_y)^2/2\mathbf{b}\cdot\mathbf{r}_y] \sim p(y|\mathbf{r}_y)$$

$$p(\mathbf{r}_z|z) \sim \exp[-(z-\mathbf{a}\cdot\mathbf{r}_z)^2/2\mathbf{b}\cdot\mathbf{r}_z] \sim p(z|\mathbf{r}_z)$$

$$\mu_x = \mathbf{a}\cdot\mathbf{r}_x$$
$$\sigma_x^2 = \mathbf{b}\cdot\mathbf{r}_x$$

$$\mu_y = \mathbf{a}\cdot\mathbf{r}_y$$
$$\sigma_y^2 = \mathbf{b}\cdot\mathbf{r}_y$$

$$\mu_z = \mathbf{a}\cdot\mathbf{r}_z = \mu_x + \mu_y = \mathbf{a}\cdot\mathbf{r}_x + \mathbf{a}\cdot\mathbf{r}_y$$
$$\sigma_z^2 = \mathbf{b}\cdot\mathbf{r}_z = \sigma_x^2 + \sigma_y^2 = \mathbf{b}\cdot\mathbf{r}_x + \mathbf{b}\cdot\mathbf{r}_y$$

goal

optimal network:

$$\mathbf{r}_z = \mathbf{r}_x + \mathbf{r}_y$$

The real thing: a linear PPC

$$p(\mathbf{r}_x|x) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_x(x - \mathbf{b}\cdot\mathbf{r}_x/\mathbf{a}\cdot\mathbf{r}_x)^2/2]$$

$$p(\mathbf{r}_y|y) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_y(x - \mathbf{b}\cdot\mathbf{r}_y/\mathbf{a}\cdot\mathbf{r}_y)^2/2]$$

$$p(\mathbf{r}_z|z) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_z(x - \mathbf{b}\cdot\mathbf{r}_z/\mathbf{a}\cdot\mathbf{r}_z)^2/2]$$

this is a linear PPC:

$$= \exp[-\mathbf{a}\cdot\mathbf{r}_x x^2/2 + \mathbf{b}\cdot\mathbf{r}_x x + (\mathbf{b}\cdot\mathbf{r}_x)^2/2\mathbf{a}\cdot\mathbf{r}_x]$$

$$= \varphi(\mathbf{r}_x) \exp[(-(x^2/2)\mathbf{a} + x\mathbf{b}) \cdot\mathbf{r}_x]$$

$\mathbf{h}(x)$

The real thing: a linear PPC

$$p(\mathbf{r}_x|x) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_x(x - \mathbf{b}\cdot\mathbf{r}_x/\mathbf{a}\cdot\mathbf{r}_x)^2/2] \sim p(x|\mathbf{r}_x)$$

$$p(\mathbf{r}_y|y) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_y(x - \mathbf{b}\cdot\mathbf{r}_y/\mathbf{a}\cdot\mathbf{r}_y)^2/2] \sim p(y|\mathbf{r}_y)$$

$$p(\mathbf{r}_z|z) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_z(x - \mathbf{b}\cdot\mathbf{r}_z/\mathbf{a}\cdot\mathbf{r}_z)^2/2] \sim p(z|\mathbf{r}_z)$$

$$\mu_x = \mathbf{b}\cdot\mathbf{r}_x/\mathbf{a}\cdot\mathbf{r}_x$$
$$\sigma_x^2 = 1/\mathbf{a}\cdot\mathbf{r}_x$$

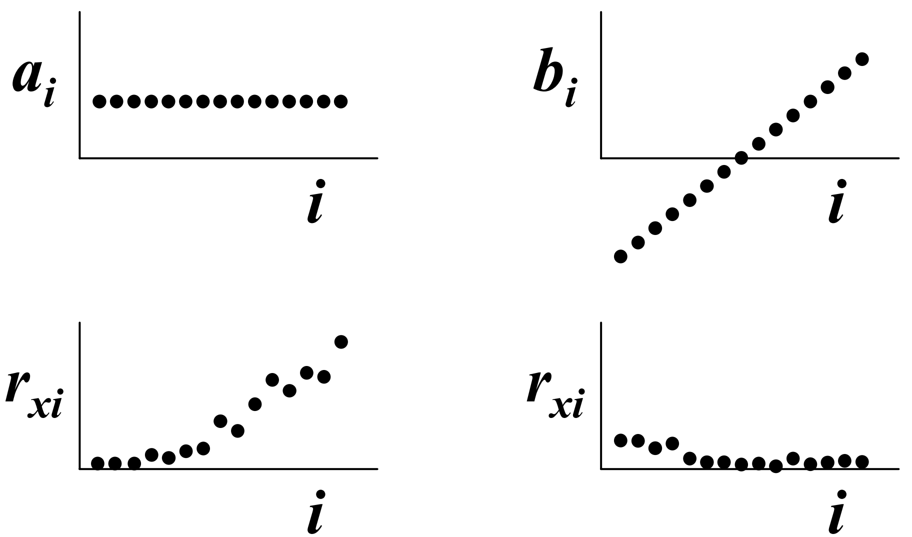
The real thing: a linear PPC

$$p(\mathbf{r}_x|x) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_x(x - \mathbf{b}\cdot\mathbf{r}_x/\mathbf{a}\cdot\mathbf{r}_x)^2/2] \sim p(x|\mathbf{r}_x)$$

$$p(\mathbf{r}_y|y) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_y(x - \mathbf{b}\cdot\mathbf{r}_y/\mathbf{a}\cdot\mathbf{r}_y)^2/2] \sim p(y|\mathbf{r}_y)$$

$$p(\mathbf{r}_z|z) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_z(x - \mathbf{b}\cdot\mathbf{r}_z/\mathbf{a}\cdot\mathbf{r}_z)^2/2] \sim p(z|\mathbf{r}_z)$$

$$\mu_x = \mathbf{b}\cdot\mathbf{r}_x/\mathbf{a}\cdot\mathbf{r}_x$$
$$\sigma_x^2 = 1/\mathbf{a}\cdot\mathbf{r}_x$$



positive mean,
low variance

negative mean,
high variance

The real thing: a linear PPC

$$p(\mathbf{r}_x|x) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_x(x - \mathbf{b}\cdot\mathbf{r}_x/\mathbf{a}\cdot\mathbf{r}_x)^2/2] \sim p(x|\mathbf{r}_x)$$

$$p(\mathbf{r}_y|y) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_y(x - \mathbf{b}\cdot\mathbf{r}_y/\mathbf{a}\cdot\mathbf{r}_y)^2/2] \sim p(y|\mathbf{r}_y)$$

$$p(\mathbf{r}_z|z) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_z(x - \mathbf{b}\cdot\mathbf{r}_z/\mathbf{a}\cdot\mathbf{r}_z)^2/2] \sim p(z|\mathbf{r}_z)$$

$$\mu_x = \mathbf{b}\cdot\mathbf{r}_x/\mathbf{a}\cdot\mathbf{r}_x$$
$$\sigma_x^2 = 1/\mathbf{a}\cdot\mathbf{r}_x$$

$$\mu_y = \mathbf{b}\cdot\mathbf{r}_y/\mathbf{a}\cdot\mathbf{r}_y$$
$$\sigma_y^2 = 1/\mathbf{a}\cdot\mathbf{r}_y$$

$$\mu_z = \mathbf{b}\cdot\mathbf{r}_z/\mathbf{a}\cdot\mathbf{r}_z$$
$$\sigma_z^2 = 1/\mathbf{a}\cdot\mathbf{r}_z$$

The real thing: a linear PPC

$$p(\mathbf{r}_x|x) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_x(x - \mathbf{b}\cdot\mathbf{r}_x/\mathbf{a}\cdot\mathbf{r}_x)^2/2] \sim p(x|\mathbf{r}_x)$$

$$p(\mathbf{r}_y|y) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_y(x - \mathbf{b}\cdot\mathbf{r}_y/\mathbf{a}\cdot\mathbf{r}_y)^2/2] \sim p(y|\mathbf{r}_y)$$

$$p(\mathbf{r}_z|z) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_z(x - \mathbf{b}\cdot\mathbf{r}_z/\mathbf{a}\cdot\mathbf{r}_z)^2/2] \sim p(z|\mathbf{r}_z)$$

$$\mu_x = \mathbf{b}\cdot\mathbf{r}_x/\mathbf{a}\cdot\mathbf{r}_x$$
$$\sigma_x^2 = 1/\mathbf{a}\cdot\mathbf{r}_x$$

$$\mu_y = \mathbf{b}\cdot\mathbf{r}_y/\mathbf{a}\cdot\mathbf{r}_y$$
$$\sigma_y^2 = 1/\mathbf{a}\cdot\mathbf{r}_y$$

$$\mu_z = \mathbf{b}\cdot\mathbf{r}_z/\mathbf{a}\cdot\mathbf{r}_z = \mu_x + \mu_y$$
$$\sigma_z^2 = 1/\mathbf{a}\cdot\mathbf{r}_z = \sigma_x^2 + \sigma_y^2$$

goal

The real thing: a linear PPC

$$p(\mathbf{r}_x|x) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_x(x - \mathbf{b}\cdot\mathbf{r}_x/\mathbf{a}\cdot\mathbf{r}_x)^2/2] \sim p(x|\mathbf{r}_x)$$

$$p(\mathbf{r}_y|y) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_y(x - \mathbf{b}\cdot\mathbf{r}_y/\mathbf{a}\cdot\mathbf{r}_y)^2/2] \sim p(y|\mathbf{r}_y)$$

$$p(\mathbf{r}_z|z) \sim \exp[-\mathbf{a}\cdot\mathbf{r}_z(x - \mathbf{b}\cdot\mathbf{r}_z/\mathbf{a}\cdot\mathbf{r}_z)^2/2] \sim p(z|\mathbf{r}_z)$$

$$\mu_x = \mathbf{b}\cdot\mathbf{r}_x/\mathbf{a}\cdot\mathbf{r}_x$$
$$\sigma_x^2 = 1/\mathbf{a}\cdot\mathbf{r}_x$$

$$\mu_y = \mathbf{b}\cdot\mathbf{r}_y/\mathbf{a}\cdot\mathbf{r}_y$$
$$\sigma_y^2 = 1/\mathbf{a}\cdot\mathbf{r}_y$$

$$\mu_z = \mathbf{b}\cdot\mathbf{r}_z/\mathbf{a}\cdot\mathbf{r}_z = \mu_x + \mu_y = \mathbf{b}\cdot\mathbf{r}_x/\mathbf{a}\cdot\mathbf{r}_x + \mathbf{b}\cdot\mathbf{r}_y/\mathbf{a}\cdot\mathbf{r}_y$$
$$\sigma_z^2 = 1/\mathbf{a}\cdot\mathbf{r}_z = \sigma_x^2 + \sigma_y^2 = 1/\mathbf{a}\cdot\mathbf{r}_x + 1/\mathbf{a}\cdot\mathbf{r}_y$$

goal

$$\mathbf{r}_z = \frac{\mathbf{a}}{\mathbf{a}\cdot\mathbf{a}}\,\frac{\mathbf{a}\cdot\mathbf{r}_x\ \mathbf{a}\cdot\mathbf{r}_y}{\mathbf{a}\cdot\mathbf{r}_x + \mathbf{a}\cdot\mathbf{r}_y} + \frac{\mathbf{b}}{\mathbf{b}\cdot\mathbf{b}}\,\frac{\mathbf{a}\cdot\mathbf{r}_x\ \mathbf{b}\cdot\mathbf{r}_y + \mathbf{a}\cdot\mathbf{r}_y\ \mathbf{b}\cdot\mathbf{r}_x}{\mathbf{a}\cdot\mathbf{r}_x + \mathbf{a}\cdot\mathbf{r}_y} + \mathbf{c}\ f(\mathbf{r}_x, \mathbf{r}_y)$$

$\mathbf{a}\cdot\mathbf{b} = 0$

$\mathbf{a}\cdot\mathbf{c} = \mathbf{b}\cdot\mathbf{c} = 0$

$$\mathbf{a}\cdot\mathbf{r}_z = \frac{\mathbf{a}\cdot\mathbf{r}_x\ \mathbf{a}\cdot\mathbf{r}_y}{\mathbf{a}\cdot\mathbf{r}_x + \mathbf{a}\cdot\mathbf{r}_y}$$

$$\mathbf{b}\cdot\mathbf{r}_z = \frac{\mathbf{a}\cdot\mathbf{r}_x\ \mathbf{b}\cdot\mathbf{r}_y + \mathbf{a}\cdot\mathbf{r}_x\ \mathbf{b}\cdot\mathbf{r}_y}{\mathbf{a}\cdot\mathbf{r}_x + \mathbf{a}\cdot\mathbf{r}_y}$$

$$\mathbf{b}\cdot\mathbf{r}_z/\mathbf{a}\cdot\mathbf{r}_z = \mathbf{b}\cdot\mathbf{r}_x/\mathbf{a}\cdot\mathbf{r}_x + \mathbf{b}\cdot\mathbf{r}_y/\mathbf{a}\cdot\mathbf{r}_y$$

$$1/\mathbf{a}\cdot\mathbf{r}_z = 1/\mathbf{a}\cdot\mathbf{r}_x + 1/\mathbf{a}\cdot\mathbf{r}_y$$

goal

$$\mathbf{r}_z = \frac{\mathbf{a}}{\mathbf{a}\cdot\mathbf{a}} \frac{\mathbf{a}\cdot\mathbf{r}_x \; \mathbf{a}\cdot\mathbf{r}_y}{\mathbf{a}\cdot\mathbf{r}_x + \mathbf{a}\cdot\mathbf{r}_y} + \frac{\mathbf{b}}{\mathbf{b}\cdot\mathbf{b}} \frac{\mathbf{a}\cdot\mathbf{r}_x \; \mathbf{b}\cdot\mathbf{r}_y + \mathbf{a}\cdot\mathbf{r}_y \; \mathbf{b}\cdot\mathbf{r}_x}{\mathbf{a}\cdot\mathbf{r}_x + \mathbf{a}\cdot\mathbf{r}_y} + \mathbf{c}\, \mathbf{f}(\mathbf{r}_x, \mathbf{r}_y)$$

quadratic nonlinearity: $\sum_{jk} w_{ijk}\, r_{xj}\, r_{yk}$

$$\mathbf{r}_z = \frac{\mathbf{a}}{\mathbf{a}\cdot\mathbf{a}} \frac{\mathbf{a}\cdot\mathbf{r}_x\,\mathbf{a}\cdot\mathbf{r}_y}{\mathbf{a}\cdot\mathbf{r}_x + \mathbf{a}\cdot\mathbf{r}_y} + \frac{\mathbf{b}}{\mathbf{b}\cdot\mathbf{b}} \frac{\mathbf{a}\cdot\mathbf{r}_x\,\mathbf{b}\cdot\mathbf{r}_y + \mathbf{a}\cdot\mathbf{r}_y\,\mathbf{b}\cdot\mathbf{r}_x}{\mathbf{a}\cdot\mathbf{r}_x + \mathbf{a}\cdot\mathbf{r}_y} + \mathbf{c}\, \mathbf{f}(\mathbf{r}_x, \mathbf{r}_y)$$

quadratic nonlinearity: $\sum_{jk} w_{ijk}\, r_{xj}\, r_{yk}$

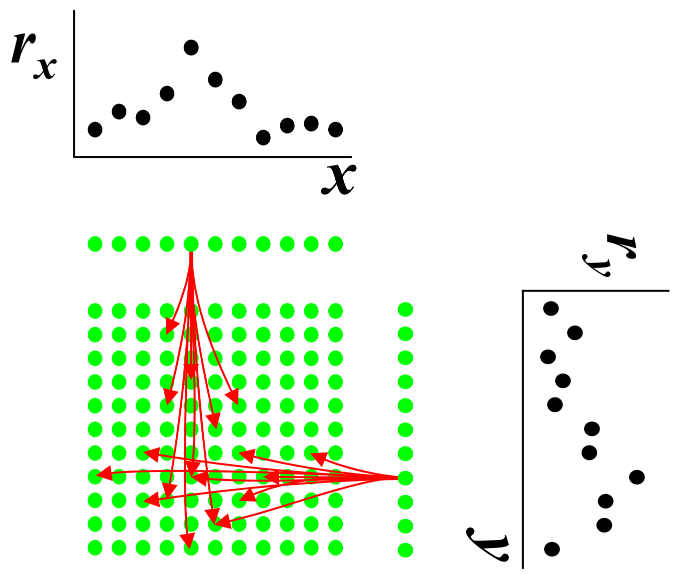
divisive normalization

$$\mathbf{r}_z = \frac{\mathbf{a}}{\mathbf{a}\cdot\mathbf{a}}\,\frac{\mathbf{a}\cdot\mathbf{r}_x\;\mathbf{a}\cdot\mathbf{r}_y}{\mathbf{a}\cdot\mathbf{r}_x + \mathbf{a}\cdot\mathbf{r}_y} + \frac{\mathbf{b}}{\mathbf{b}\cdot\mathbf{b}}\,\frac{\mathbf{a}\cdot\mathbf{r}_x\;\mathbf{b}\cdot\mathbf{r}_y + \mathbf{a}\cdot\mathbf{r}_y\;\mathbf{b}\cdot\mathbf{r}_x}{\mathbf{a}\cdot\mathbf{r}_x + \mathbf{a}\cdot\mathbf{r}_y} + c\,\mathbf{f}(\mathbf{r}_x, \mathbf{r}_y)$$

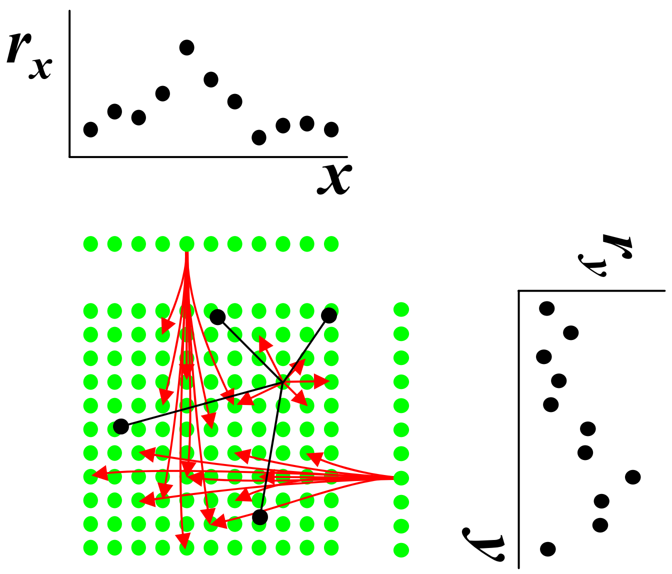
quadratic nonlinearity: $\sum_{jk} w_{ijk}\, r_{xj}\, r_{yk}$

divisive normalization

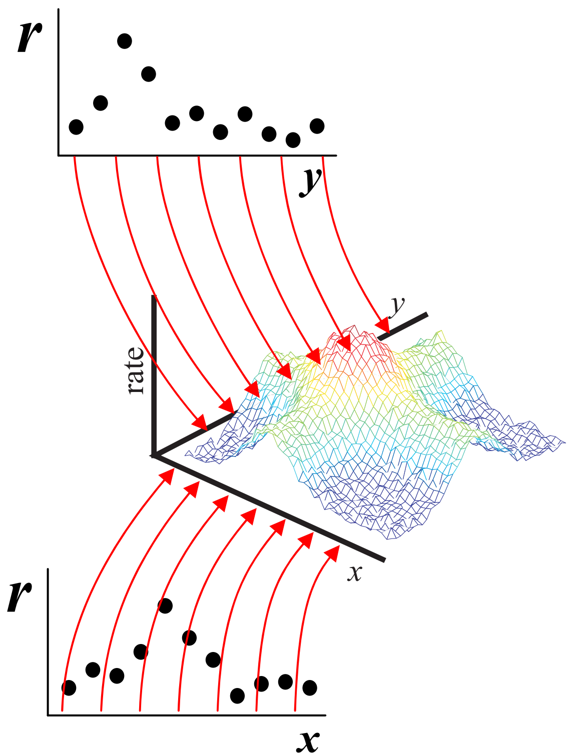
The optimal network has a quadratic nonlinearity and divisive normalization.

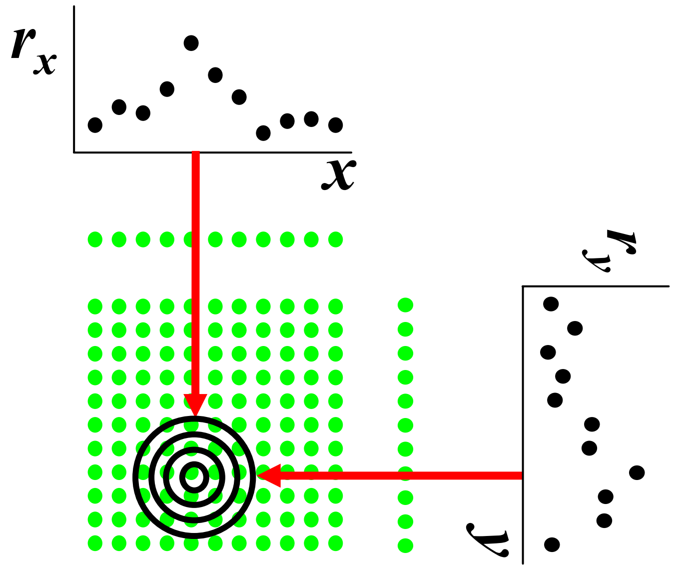


r_x
x
r_y
y

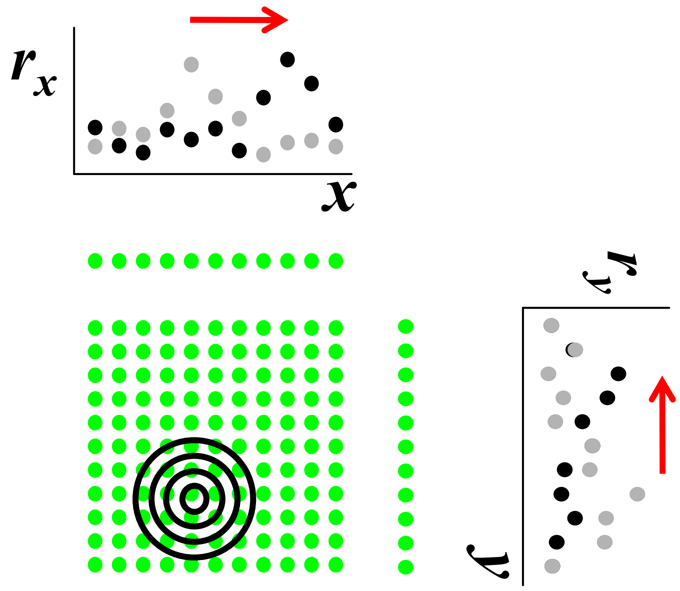


r_x
x
r_y
y

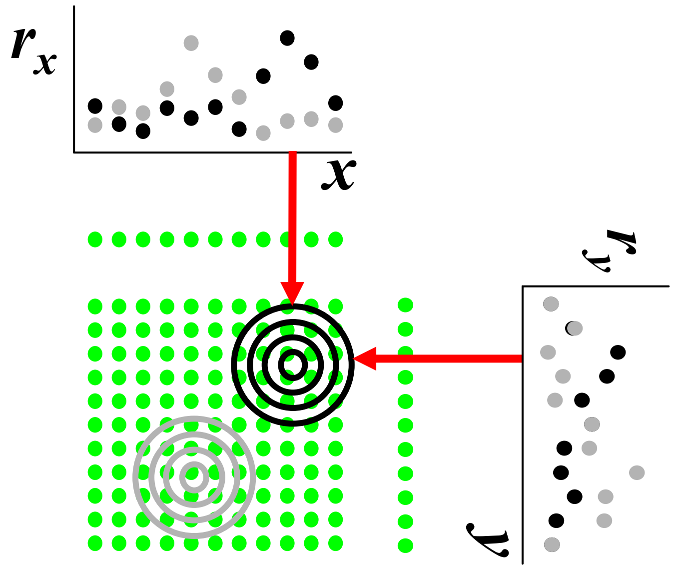

r
y
rate
y
x
r
x



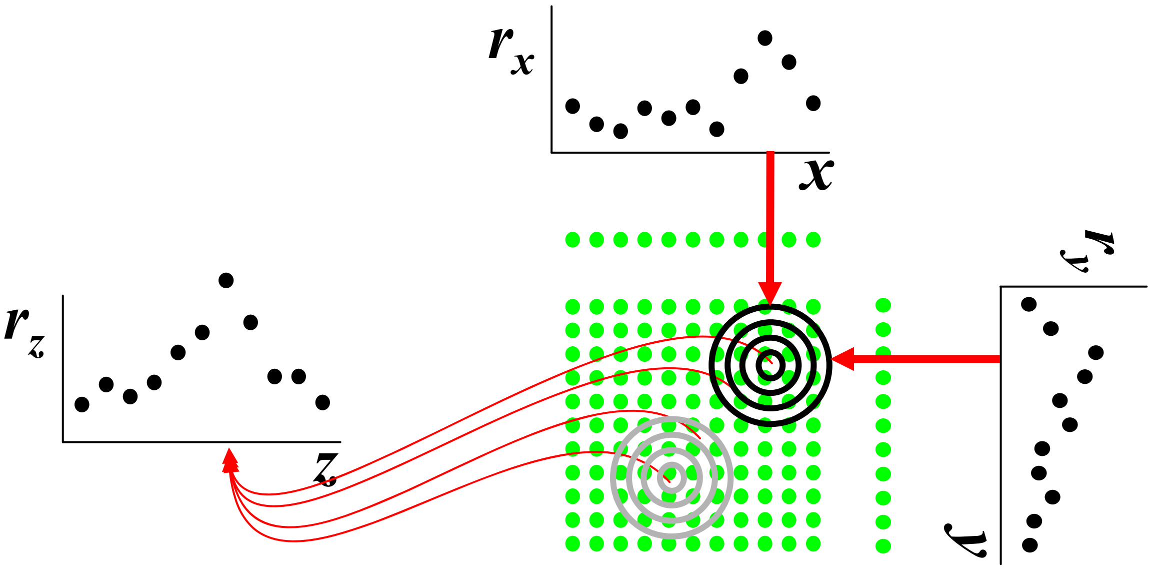
r_x
x
r_y
y



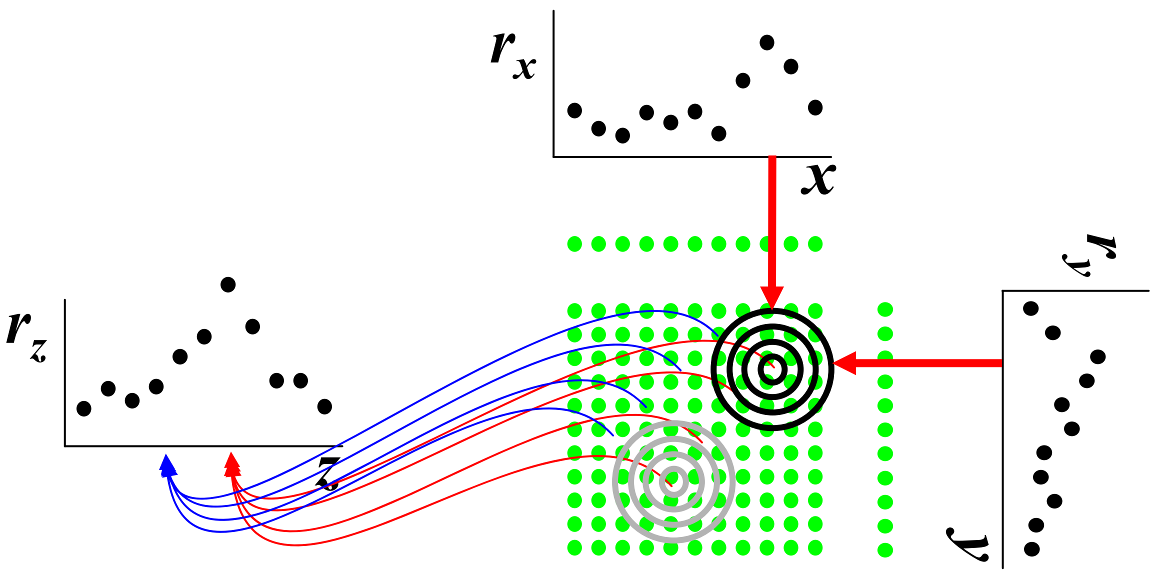
r_x
x
r_y
y



r_x
x
r_y
y



r_x
x
r_y
y
r_z
z

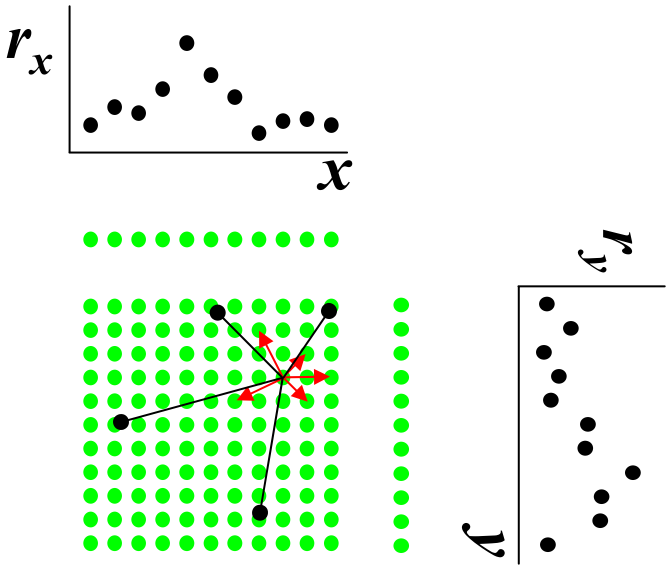


Deneve et al., *Nature Neurosci.* (2001)

the new analysis tells us exactly what the recurrent connections need to do:

they need to produce a hill of activity whose amplitude scales as

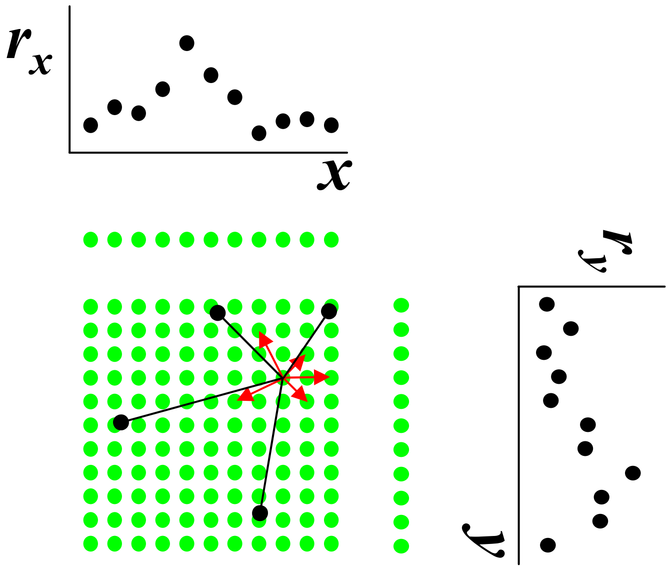
$$\frac{\mathbf{r}_{xi}\,\mathbf{r}_{yj}}{\sum_i \mathbf{a}_i\,\mathbf{r}_{xi} + \mathbf{a}_i\,\mathbf{r}_{yi}}$$

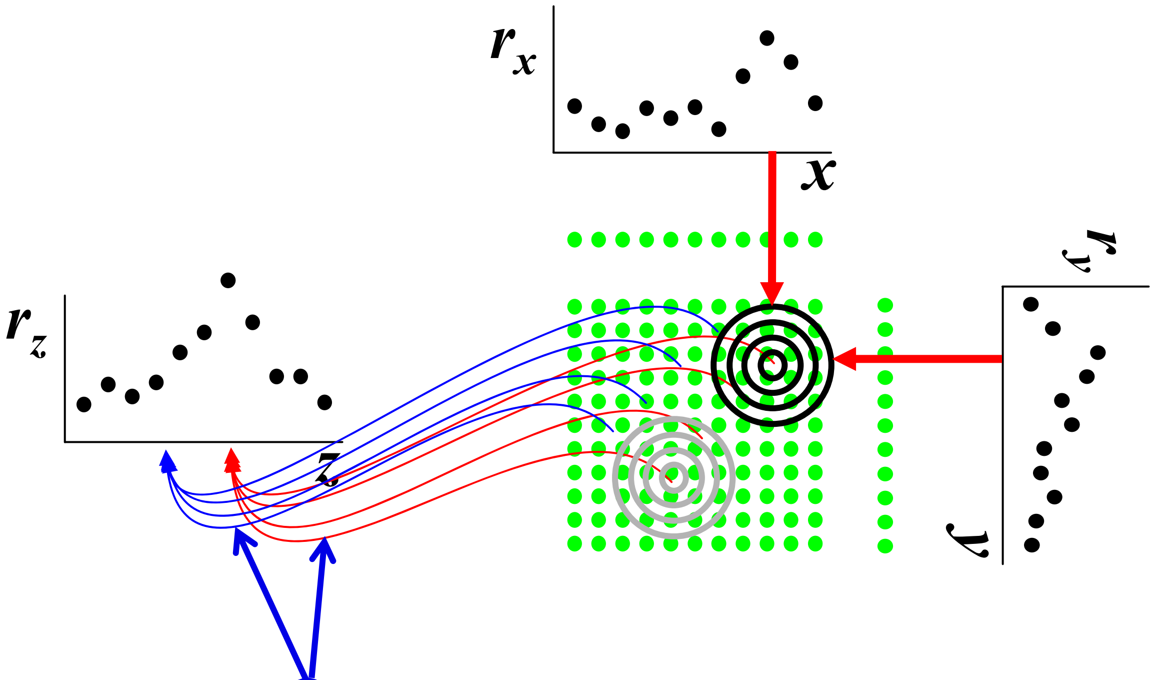


the new analysis tells us exactly what the recurrent connections need to do:

they need to produce a hill of activity whose amplitude scales as

$$\frac{\alpha r_{xi} + \beta r_{yi} + \gamma r_{xi}\, r_{yj}}{\sum_i a_i\, r_{xi} + a_i\, r_{yi}}$$





$$\frac{\sum_{jk} w_{ijk}\, r_{xj}\, r_{yk}}{\sum_i \mathbf{a}_i\, \mathbf{r}_{xi} + \mathbf{a}_i\, \mathbf{r}_{yi}}$$

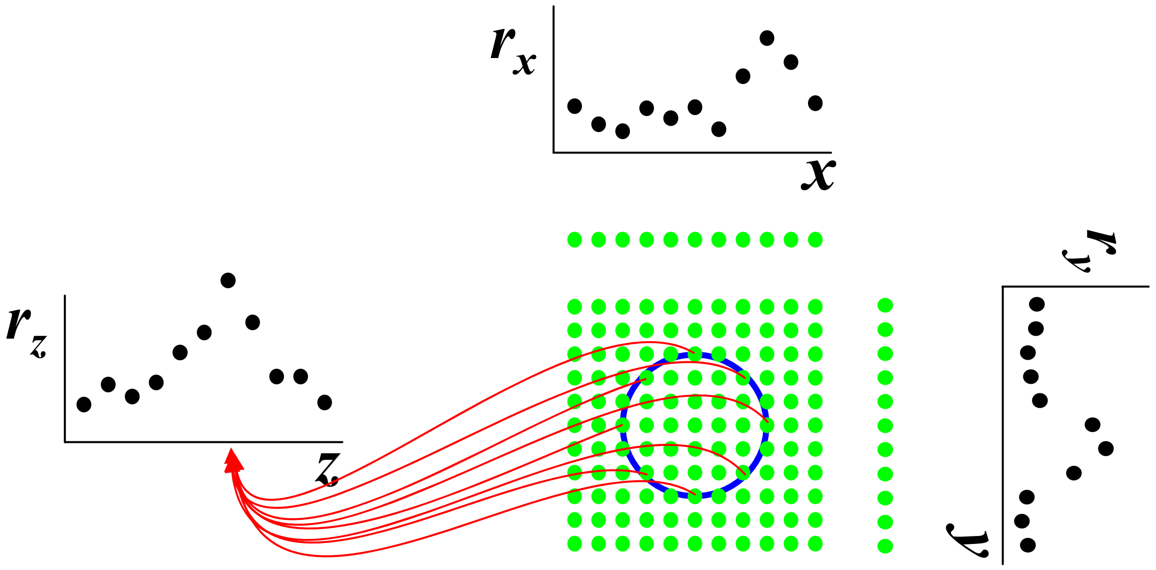
An aside: our network is deterministic,

$$\mathbf{r}_z = \frac{\mathbf{a}}{\mathbf{a}\cdot\mathbf{a}} \frac{\mathbf{a}\cdot\mathbf{r}_x\ \mathbf{a}\cdot\mathbf{r}_y}{\mathbf{a}\cdot\mathbf{r}_x + \mathbf{a}\cdot\mathbf{r}_y} + \frac{\mathbf{b}}{\mathbf{b}\cdot\mathbf{b}} \frac{\mathbf{a}\cdot\mathbf{r}_x\ \mathbf{b}\cdot\mathbf{r}_y + \mathbf{a}\cdot\mathbf{r}_y\ \mathbf{b}\cdot\mathbf{r}_x}{\mathbf{a}\cdot\mathbf{r}_x + \mathbf{a}\cdot\mathbf{r}_y} + \mathbf{c}\ \mathbf{f}(\mathbf{r}_x, \mathbf{r}_y)$$

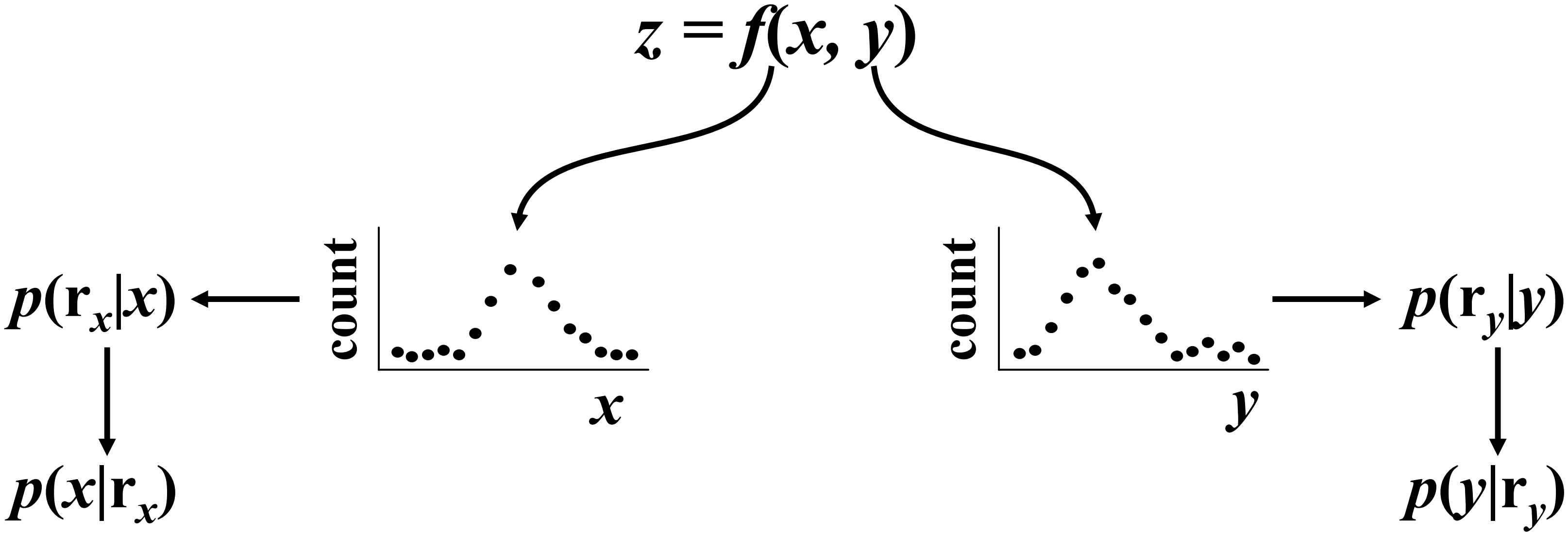
If neurons are sufficiently uncorrelated, we can replace rates with spikes, and the information loss will scale as 1/*N*.

This information loss is negligible compared to the loss associated with approximate computations.

can handle nonlinear functions (e.g., $z=x^2+y^2$)



the hard part is choosing the recurrent connectivity properly.



$$p(z|\mathrm{r}_x, \mathrm{r}_y) = \int dxdy\ \delta(z - f(x, y))\ p(x|\mathrm{r}_x)\ p(y|\mathrm{r}_y)$$

hard integral!

$$p(z|\mathbf{r}_x, \mathbf{r}_y) = \int dx\,dy\; \delta(z - f(x, y))\; p(x|\mathbf{r}_x)\; p(y|\mathbf{r}_y)$$

hard integral!

$$\mathbf{r}_z = \mathbf{F}(\mathbf{r}_x, \mathbf{r}_y) \leftarrow \textbf{network}$$

$$p(z|\mathbf{r}_z) = p(z|\mathbf{F}(\mathbf{r}_x, \mathbf{r}_y)) = p(z|\mathbf{r}_x, \mathbf{r}_y)$$

goal

approach: minimize

$$D_{KL}\left(p(z|\mathbf{r}_x, \mathbf{r}_y) \| p(z|\mathbf{r}_z)\right)$$

$$\mathbf{r}_z = \mathbf{F}(\mathbf{r}_x, \mathbf{r}_y)$$

network

network: $\mathbf{r}_z = \mathrm{F}(\mathbf{r}_x, \mathbf{r}_y; \Theta)$

minimize: $\mathrm{D}_{\mathrm{KL}}\left(p(z|\mathbf{r}_x, \mathbf{r}_y) \,||\, q(z|\mathbf{r}_z)\right)$

with respect to network parameters and parameters of encoding model

problem #1: there's a trivial solution,

$$\mathbf{r}_z = (\mathbf{r}_x, \mathbf{r}_y)$$

solution: demand that

$$p(z|\mathbf{r}_z) \sim \exp[\mathbf{h}(z)\cdot\mathbf{r}_z]$$

pure code!

minimize

$$D_{KL}\left(p(z|\mathbf{r}_x, \mathbf{r}_y)\,||\,\exp[\mathbf{h}(z)\cdot\mathbf{r}_z]/Z\right)$$

problem #2: what class of networks do we consider?

$$r_{zi} = \frac{\sum_j w_{ij}^x r_{xj} + \sum_j w_{ij}^y r_{yj} + \sum_{jk} w_{ijk} r_{xj} r_{yk}}{\alpha_i + w_i \sum_j \mathbf{a}_j \mathbf{r}_{xj} + \mathbf{a}_j \mathbf{r}_{yj}}$$

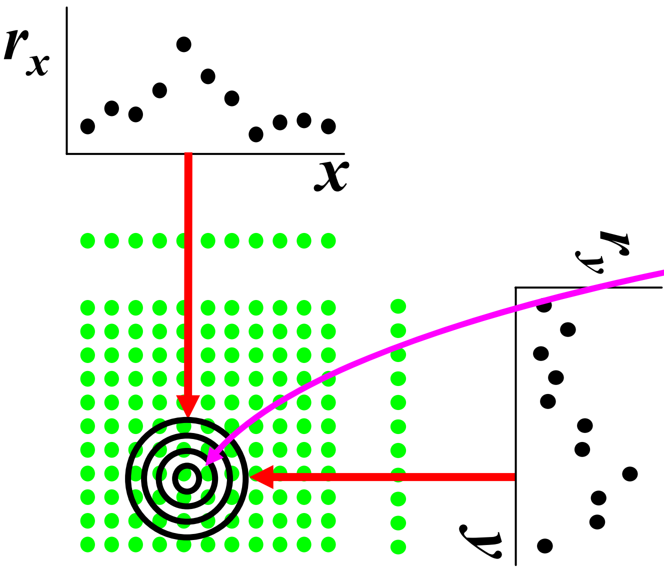
minimize

$$D_{KL}\left(p(z|\mathbf{r}_x, \mathbf{r}_y) \,||\, \exp[\mathbf{h}(z)\cdot\mathbf{r}_z]/Z\right)$$

with respect to h(z) and the parameters of the network,

$$r_{zi} = \frac{\sum_j w_{ij}^x r_{xj} + \sum_j w_{ij}^y r_{yj} + \sum_{jk} w_{ijk} r_{xj} r_{yk}}{\alpha_i + w_i \sum_j \mathbf{a}_j \mathbf{r}_{xj} + \mathbf{a}_j \mathbf{r}_{yj}}$$

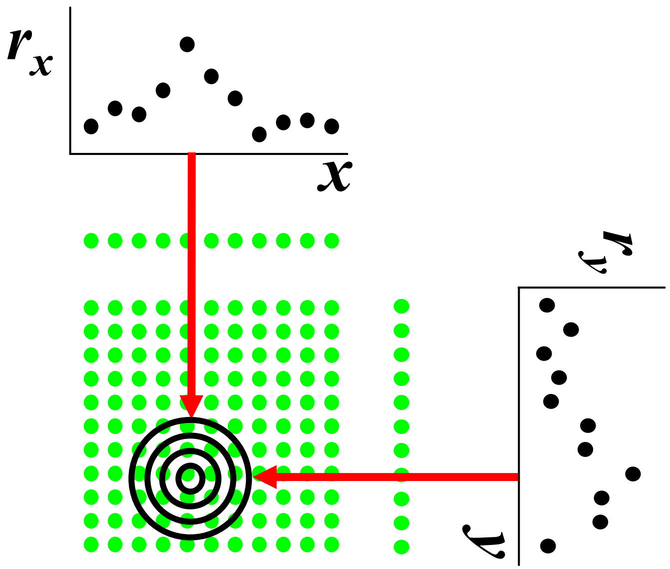
intuition



$$r_{zi} = \frac{\sum_j w^{x}_{ij}\, r_{xj} + \sum_j w^{y}_{ij}\, r_{yj} + \sum_{jk} w_{ijk}\, r_{xj}\, r_{yk}}{\alpha_i + w_i \sum_j \mathbf{a}_j\, \mathbf{r}_{xj} + \mathbf{a}_j\, \mathbf{r}_{yj}}$$

- quadratic nonlinearity produces a hill of activity
- divisive normalization corrects for it

intuition



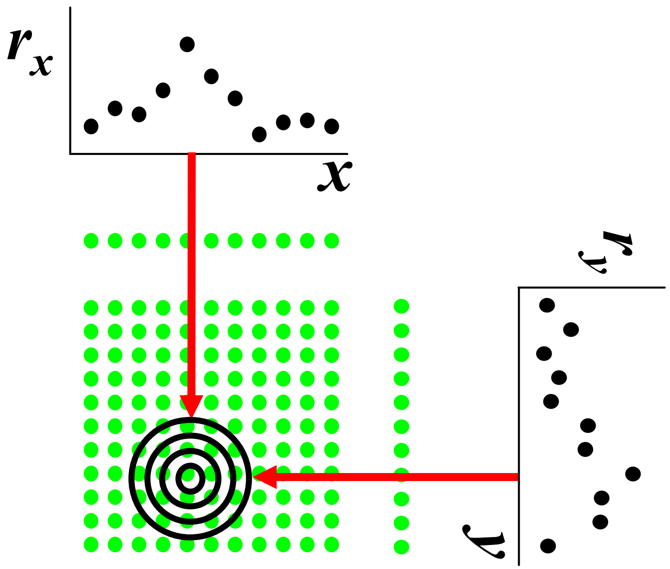
$$r_z = \frac{r_x r_y}{r_x + r_y}$$

$$1/r_z = 1/r_x + 1/r_x$$

$$\sigma_z^2 \sim 1/r_z$$

$$z = f(x, y) \Rightarrow \sigma_z^2 = \alpha\sigma_x^2 + \beta\sigma_y^2$$

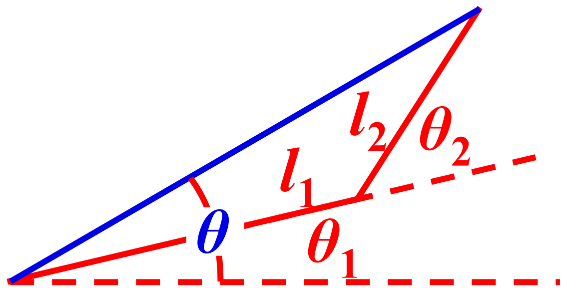
intuition



$$r_{zi} = \frac{\sum_j w^x_{ij}\, r_{xj} + \sum_j w^y_{ij}\, r_{yj} + \sum_{jk} w_{ijk}\, r_{xj}\, r_{yk}}{\alpha_i + w_i \sum_j \mathbf{a}_j\, \mathbf{r}_{xj} + \mathbf{a}_j\, \mathbf{r}_{yj}}$$

$z = f(x, y)$ => $\sigma_z^2 = \alpha\sigma_x^2 + \beta\sigma_y^2$

2-joint arm



$$\theta = \tan^{-1} \frac{l_1 \sin(\theta_1) + l_2 \sin(\theta_1 + \theta_2)}{l_1 \cos(\theta_1) + l_2 \cos(\theta_1 + \theta_2)}$$

$$p(\mathrm{r}|\theta) \sim \exp[\, \textstyle\sum_i \cos(\theta - \varphi_i)\, r_i \,]$$

minimize

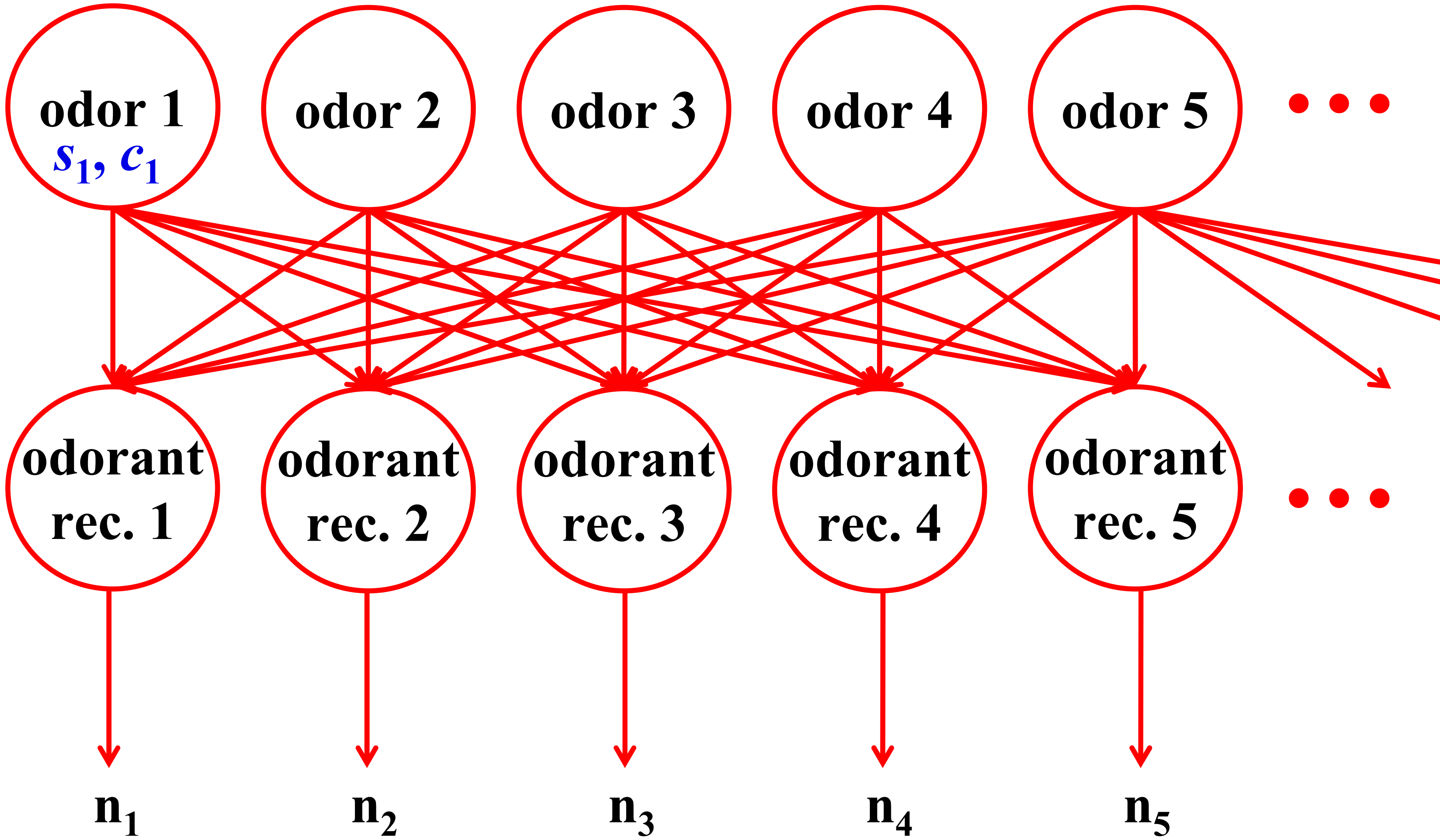
$$D_{KL}\left(p(\theta|\mathbf{r}_1, \mathbf{r}_2)\,\middle\|\,\exp[\mathbf{h}(\theta)\cdot\mathbf{r}]/Z\right)$$

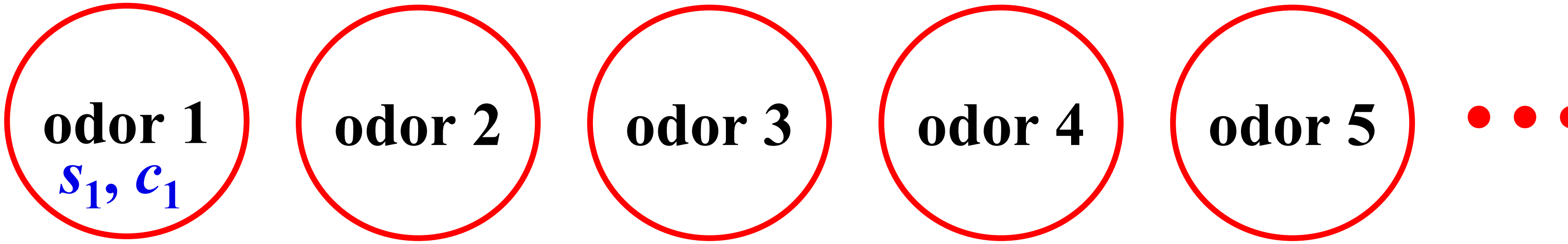
with respect to the parameters of the network

$$r_i = \frac{\sum_j w_{ij}^1 r_{1j} + \sum_j w_{ij}^2 r_{2j} + \sum_{jk} w_{ijk} r_{1j} r_{2k}}{\alpha_i + w_i \sum_j \mathbf{a}_j \mathbf{r}_{1j} + \mathbf{a}_j \mathbf{r}_{2j}}$$

$$\frac{D_{KL}\left(p(\theta|\mathbf{r}_1, \mathbf{r}_2)\,\middle\|\,p(\theta|\mathbf{r})\right)}{I(\theta; \mathbf{r}_1, \mathbf{r}_2)} \approx 0.05 \text{ (5\% information loss)}$$

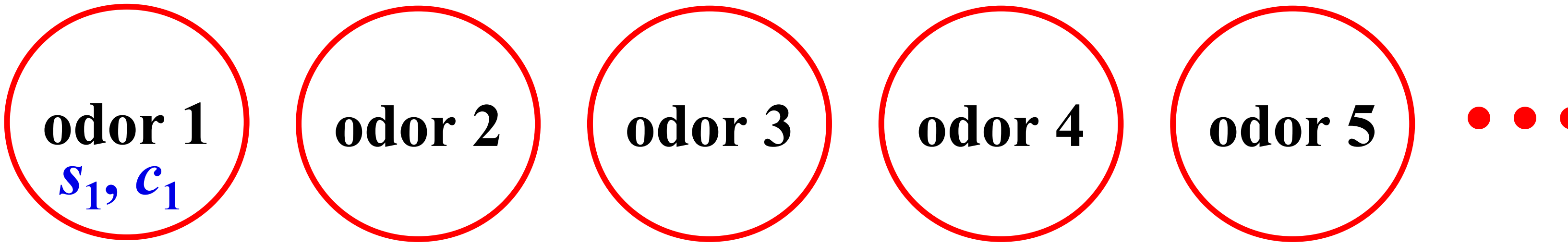
3: a hard problem (simplified olfaction)





s: 0=absent
1=present

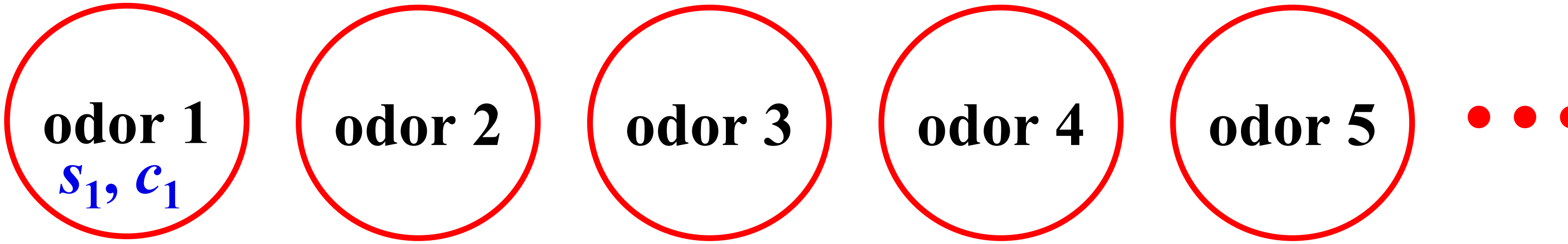
c: concentration



$$o_i = \sum_j W_{ij}\, s_j c_j$$

$$p(\mathbf{n}_i \,|\, o_i) \sim \exp[\mathbf{h}(o_i)\cdot\mathbf{n}_i]$$

goal: compute $p(s_i |\, \mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3, \ldots)$

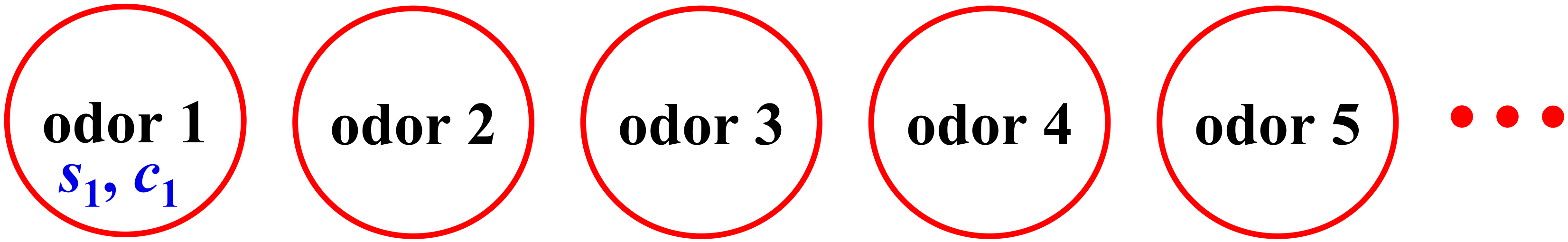


find a network

$$n_i^{\text{out}} = F(n_1, n_2, \ldots)$$

such that

$$p(s_i | n_i^{\text{out}}) = p(s_i | n_1, n_2, \ldots)$$



Intractible:

$$p(s_i|\mathrm{n}_1, \mathrm{n}_2, \ldots) = \sum_{\{s_i,\, j \neq i\}} \int \mathrm{dc}\; p(s_1, c_1, s_2, c_2, \ldots|\mathrm{n}_1, \mathrm{n}_2, \ldots)$$

***k* odors:**

2^{k-1} terms in the sum

***k*-dimensional integral**

minimize

$$\mathrm{D_{KL}}\Big(p(s_i|\mathrm{n}_1, \mathrm{n}_2, \ldots)\Big|\Big|\exp[\mathrm{h}(s_i)\cdot \mathrm{n}_i^{\mathrm{out}}]/Z\Big)$$

with respect to the parameters of the network

$$n_i^{\mathrm{out}} = \frac{\sum_{jk} w_{ij}^{k} n_{kj} + \sum_{jk} w_{ijk}^{km} n_{lj} n_{mk}}{\alpha_i + w_i \sum_{jk} \mathrm{a}_j^{k} n_{kj}}$$

$$\frac{\mathrm{D_{KL}}\Big(p(s_i|\mathrm{n}_1, \mathrm{n}_2, \mathrm{n}_3, \mathrm{n}_4)\Big|\Big|p(s_i|\mathrm{n}_i^{\mathrm{out}})\Big)}{I(s_i; \mathrm{n}_1, \mathrm{n}_2, \mathrm{n}_3, \mathrm{n}_4)} \approx 0.02 \text{ (2\% information loss)}$$

4 odors:

$$\frac{\mathrm{D_{KL}}\left(p(s_i|\mathrm{n}_1, \mathrm{n}_2, \mathrm{n}_3, \mathrm{n}_4)\,\middle\|\,p(s_i|\mathrm{n}_i^{\mathrm{out}})\right)}{I(s_i; \mathrm{n}_1, \mathrm{n}_2, \mathrm{n}_3, \mathrm{n}_4)} \approx 0.02 \text{ (2\% information loss)}$$

100s of odors:

work in progress

Summary

Our approach: minimize

$$D_{KL}\left(p(\xi|\mathbf{r}^{in})\,\|\,\exp[\mathbf{h}(\xi)\cdot\mathbf{r}^{out}]/Z\right)$$

$(x, y, z, \ldots)$

$\xi = f(x, y, z, \ldots)$

with respect to the parameters of the network

$$r_i^{\text{out}} = \frac{\sum_j w_{ij}\, r_j^{\text{in}} + \sum_{jk} w_{ijk}\, r_j^{\text{in}}\, r_k^{\text{in}}}{\alpha_i + w_i \sum_j \mathrm{a}_j\, \mathrm{r}_j^{\text{in}}}$$

Summary

Our approach: minimize

$$D_{KL}\left(p(\xi|\mathbf{r}^{in})\,\|\,\exp[\mathbf{h}(\xi)\cdot\mathbf{r}^{out}]/Z\right)$$

$q(\xi|x, y, z, \ldots)$

$(x, y, z, \ldots)$

with respect to the parameters of the network

$$r_i^{\text{out}} = \frac{\sum_j w_{ij}\, r_j^{\text{in}} + \sum_{jk} w_{ijk}\, r_j^{\text{in}}\, r_k^{\text{in}}}{\alpha_i + w_i \sum_j a_j\, r_j^{\text{in}}}$$

Summary

Our approach: minimize

$$D_{KL}\left(p(\xi|\mathrm{r}^{in})\,\middle\|\,\exp[\mathrm{h}(\xi)\cdot\mathrm{r}^{out}]/Z\right)$$

$p(\xi|\mathrm{r}^{in}) \rightarrow (x, y, z, \ldots)$

$\exp[\mathrm{h}(\xi)\cdot\mathrm{r}^{out}]/Z \rightarrow q(\xi|x, y, z, \ldots)$

Just finding $p(\xi|\mathrm{r}^{in})$ is a hard inference problem,

$$p(\xi|\mathrm{r}^{in}) = \int dx\, dy\, dz\, \ldots\, \delta(\xi - \mathrm{f}(x, y, z, \ldots))\, p(x, y, z, \ldots\, |\mathrm{r}^{in})$$

implementing it in a network is even harder.

$$p(\xi|\mathbf{r}^{in}) = \int dx\, dy\, dz\, \ldots\, \delta(\xi - f(x, y, z, \ldots))\, p(x, y, z, \ldots\, |\mathbf{r}^{in})$$

We have to do approximate inference.

- **parameterized probability distributions.**

- **approximate networks, chosen by minimizing a cost function.**

I talked about one parameterization and one class of networks. The big open questions:

what is the appropriate parameterization for the brain?
what is the appropriate class of networks?
how is all this learned?

Jeff Beck
Gatsby Computational
Neuroscience Unit. London, UK

Alex Pouget
University of Rochester
Rochester, NY

Funding: <u>Gatsby Charitable Foundation</u>

Jeff Beck
Gatsby Computational
Neuroscience Unit. London, UK

Alex Pouget
University of Rochester
Rochester, NY

Funding: Gatsby Charitable Foundation