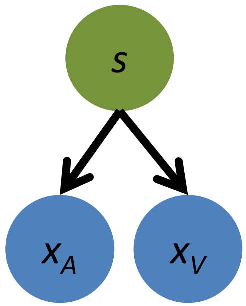
The neural basis of Bayesian inference

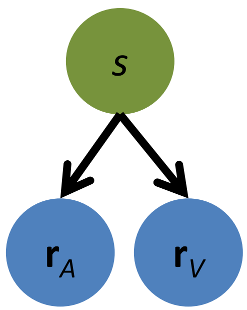
Lecture 6

Optimal cue integration (flat prior)



Behavioral model:

$$p(s \mid x_A, x_V) \propto p(x_A, x_V \mid s)$$
$$= p(x_A \mid s)\, p(x_V \mid s)$$



Neural model:

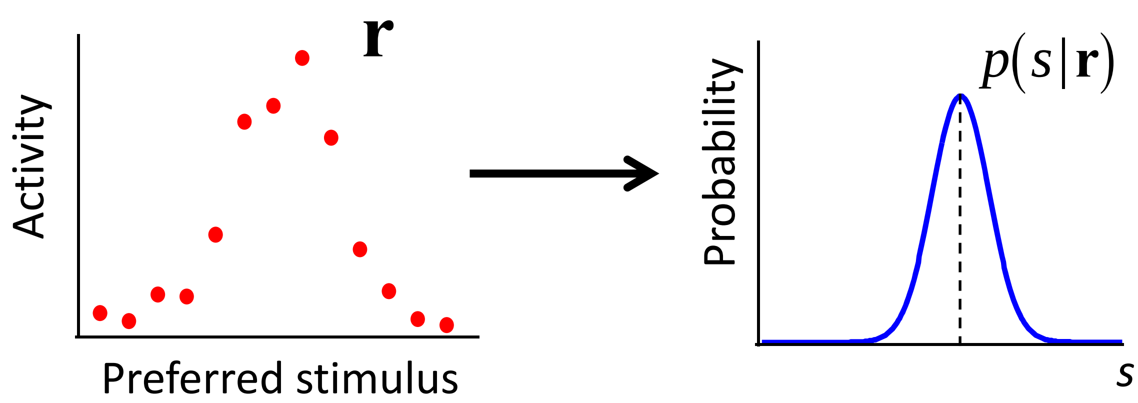
$$p(s \mid \mathbf{r}_A, \mathbf{r}_V) \propto p(\mathbf{r}_A, \mathbf{r}_V \mid s)$$
$$= p(\mathbf{r}_A \mid s)\, p(\mathbf{r}_V \mid s)$$

Response (estimate) distribution has

$$\frac{1}{\sigma_{AV}^2} = \frac{1}{\sigma_A^2} + \frac{1}{\sigma_V^2}$$

$$\frac{\langle \hat{s} \rangle}{\sigma_{AV}^2} = \frac{s_A}{\sigma_A^2} + \frac{s_V}{\sigma_V^2}$$

Uncertainty is encoded through PPC



Question: how is this PPC used to implement the computation

$$p(s \mid \mathbf{r}_A, \mathbf{r}_V) \propto p(\mathbf{r}_A \mid s)\, p(\mathbf{r}_V \mid s) \qquad ?$$

Neural variability model

Assume (only for now) independent Poisson variability.

$$p(\mathbf{r} \mid s) = \prod_{i=1}^{N} \frac{e^{-f_i(s)} f_i(s)^{r_i}}{r_i!}$$

Lecture 1 → can be rewritten as

$$p(\mathbf{r} \mid s) = \frac{\varphi(\mathbf{r})}{\eta(s)} e^{\mathbf{h}(s)\cdot\mathbf{r}}$$ with

$$\varphi(\mathbf{r}) = \prod_{i=1}^{N} \frac{1}{r_i!}$$

$$\eta(s) = \exp\left(-\sum_{i=1}^{N} f_i(s)\right)$$

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

General: *Poisson-like* variability

$$p(\mathbf{r} \mid s) = \frac{\varphi(\mathbf{r})}{\eta(s)} e^{\mathbf{h}(s)\cdot\mathbf{r}}$$

Fano can be unequal to 1; correlations allowed → more realistic

Relationship between **h**(*s*), tuning curve, and covariance

Exercise: given $p(\mathbf{r} \mid s) = \dfrac{\varphi(\mathbf{r})}{\eta(s)} e^{\mathbf{h}(s)\cdot\mathbf{r}}$

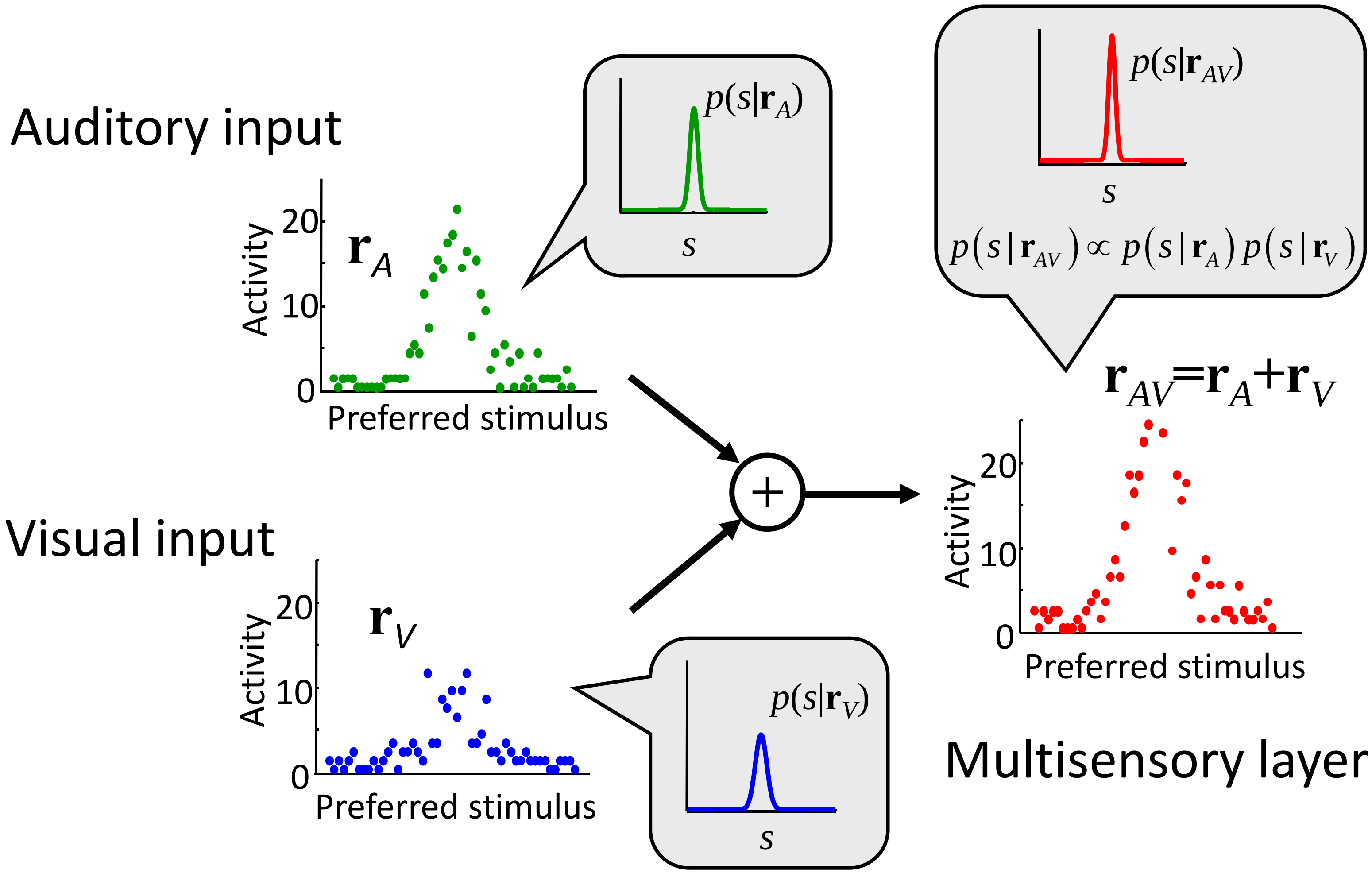
$$\mathbf{f}(s) = \langle \mathbf{r} \rangle$$

$$\mathbf{\Sigma}(s) = \langle \mathbf{r}\mathbf{r}^T \rangle - \langle \mathbf{r} \rangle \langle \mathbf{r}^T \rangle$$

Show that

$$\mathbf{f}'(s) = \mathbf{\Sigma}(s)\mathbf{h}'(s)$$

Implementation of optimal cue integration



If $\mathbf{r}_A$ and $\mathbf{r}_V$ are Poisson-like with the same $\mathbf{h}(s)$, then addition implements optimal cue integration.

Third population: $\mathbf{r}_{AV} = \mathbf{r}_A + \mathbf{r}_V$

$$
\begin{aligned}
p(\mathbf{r}_{AV} \mid s) &= \iint p(\mathbf{r}_A \mid s)\, p(\mathbf{r}_V \mid s)\, \delta(\mathbf{r}_{AV} - \mathbf{r}_A - \mathbf{r}_V)\, d\mathbf{r}_A d\mathbf{r}_V \\
&= \iint \frac{\varphi_A(\mathbf{r}_1)\varphi_V(\mathbf{r}_2)}{\eta_A(s)\eta_V(s)} \exp(\mathbf{h}(s)\cdot\mathbf{r}_A)\exp(\mathbf{h}(s)\cdot\mathbf{r}_V)\,\delta(\mathbf{r}_{AV} - \mathbf{r}_A - \mathbf{r}_V)\, d\mathbf{r}_A d\mathbf{r}_V \\
&= \int \frac{\varphi_A(\mathbf{r}_A)\varphi_V(\mathbf{r}_{AV} - \mathbf{r}_A)}{\eta_A(s)\eta_V(s)} \exp(\mathbf{h}(s)\cdot\mathbf{r}_{AV})\, d\mathbf{r}_A \\
&= \frac{\varphi_{AV}(\mathbf{r}_{AV})}{\eta_A(s)\eta_V(s)} \exp(\mathbf{h}(s)\cdot\mathbf{r}_{AV})
\end{aligned}
$$

$$p\left(\mathbf{r}_{AV} \mid s\right)=\frac{\varphi_{AV}\left(\mathbf{r}_{AV}\right)}{\eta_A(s)\eta_V(s)}\exp\left(\mathbf{h}(s)\cdot\mathbf{r}_{AV}\right)$$

$$p\left(\mathbf{r}_A \mid s\right)p\left(\mathbf{r}_V \mid s\right)=\frac{\varphi_A\left(\mathbf{r}_A\right)\varphi_V\left(\mathbf{r}_V\right)}{\eta_A(s)\eta_V(s)}\exp\left(\mathbf{h}(s)\cdot\mathbf{r}_A\right)\exp\left(\mathbf{h}(s)\cdot\mathbf{r}_V\right)$$

$$=\frac{\varphi_A\left(\mathbf{r}_A\right)\varphi_V\left(\mathbf{r}_V\right)}{\eta_A(s)\eta_V(s)}\exp\left(\mathbf{h}(s)\cdot\mathbf{r}_{AV}\right)$$

Same dependence on *s*, therefore

$$p\left(\mathbf{r}_{AV} \mid s\right)\propto p\left(\mathbf{r}_A \mid s\right)p\left(\mathbf{r}_V \mid s\right)$$

$$p\left(s \mid \mathbf{r}_{AV}\right)\propto p\left(\mathbf{r}_A \mid s\right)p\left(\mathbf{r}_V \mid s\right)$$

Relating back to behavior

$$\frac{1}{\sigma_{AV}^2} = \frac{1}{\sigma_A^2} + \frac{1}{\sigma_V^2}$$

$$\frac{\langle \hat{s} \rangle}{\sigma_{AV}^2} = \frac{s_A}{\sigma_A^2} + \frac{s_V}{\sigma_V^2}$$

Lecture 1: Fisher information

$$\frac{1}{\sigma_{\text{estimate}}^2} = I(s) = -\left\langle \frac{\partial^2}{\partial s^2} \log p(\mathbf{r} \mid s) \right\rangle$$

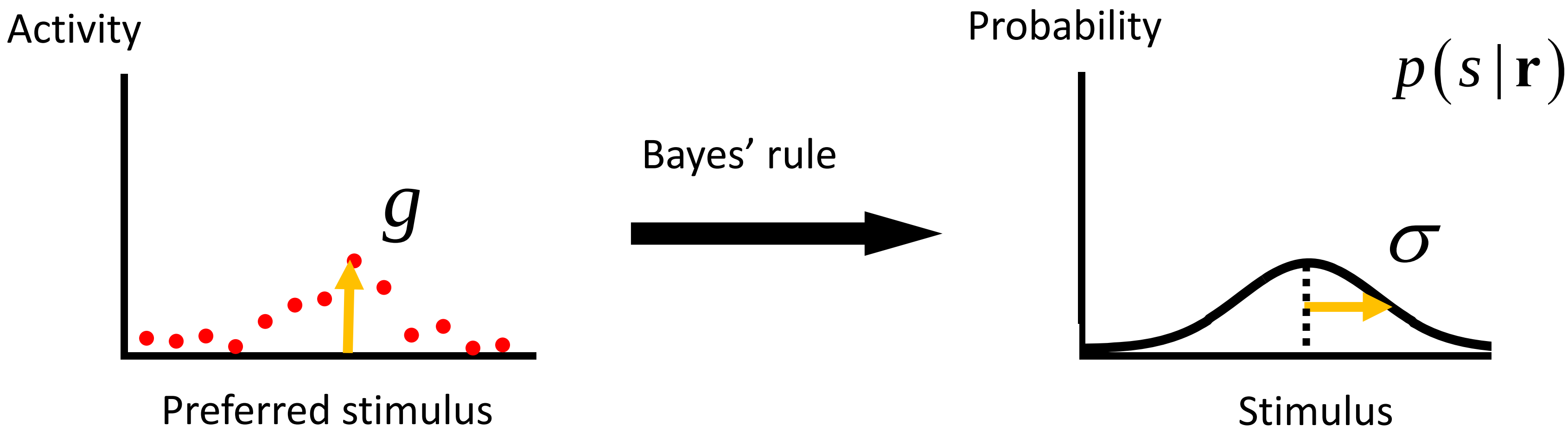
Fisher information for independent Poisson variability:

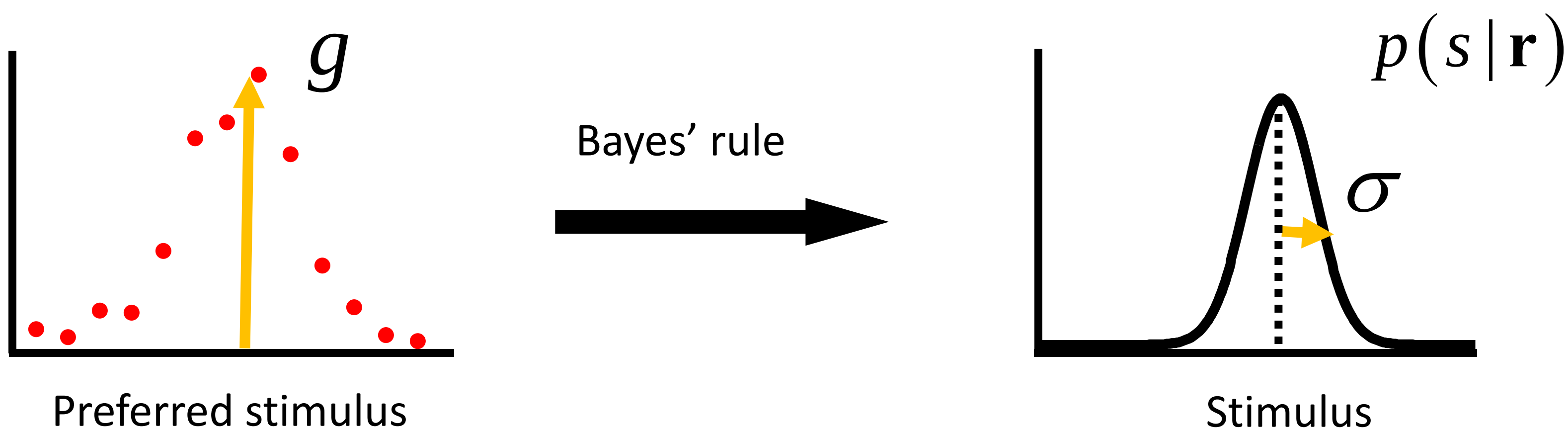
$$I(s) = \sum_{i=1}^{N} \frac{f_i'(s)^2}{f_i(s)}$$

Exercises:

- Compute Fisher information for Poisson-like variability
- Show that $\mathbf{r}_{AV} = \mathbf{r}_A + \mathbf{r}_V$ implies $\frac{1}{\sigma_{AV}^2} = \frac{1}{\sigma_A^2} + \frac{1}{\sigma_V^2}$

Special case: identical tuning curves and covariance matrices





Gain and precision

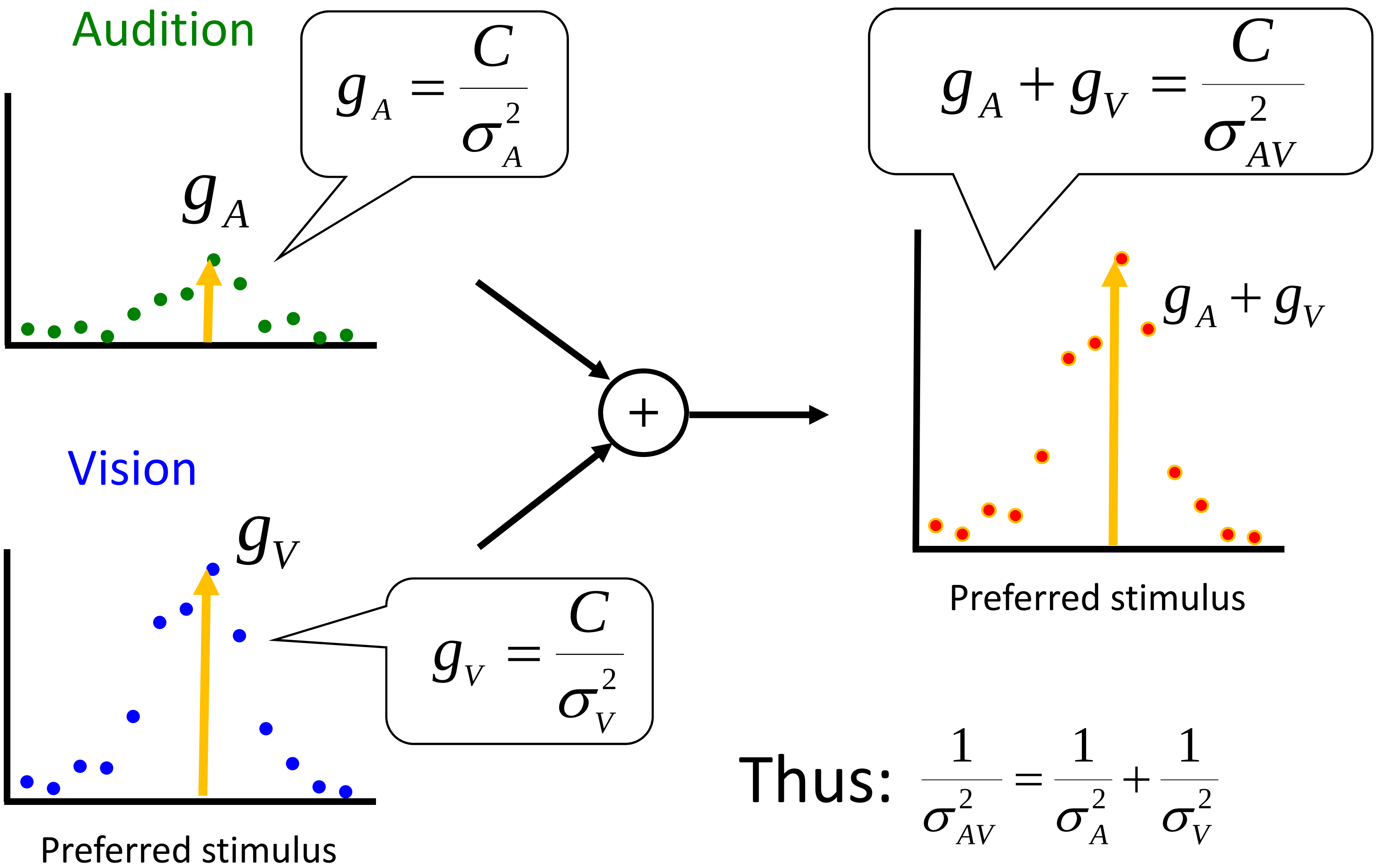
$$g \propto I(s) = \frac{1}{\sigma^2}$$

gain of population

variance of estimate distribution

High gain, high certainty: not guaranteed!

Multisensory gain and precision



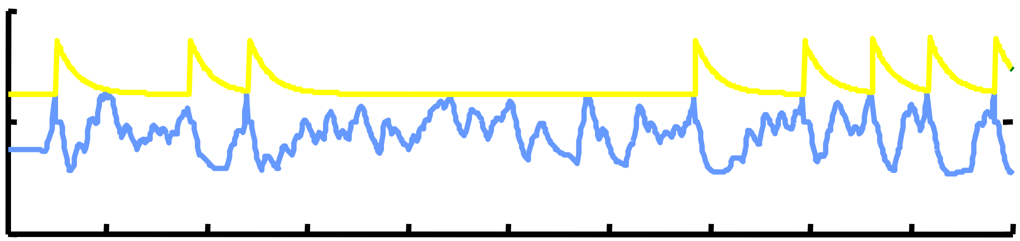
Thus: $\frac{1}{\sigma_{AV}^2} = \frac{1}{\sigma_A^2} + \frac{1}{\sigma_V^2}$

What if $\mathbf{h}(s)$ not the same in both populations?

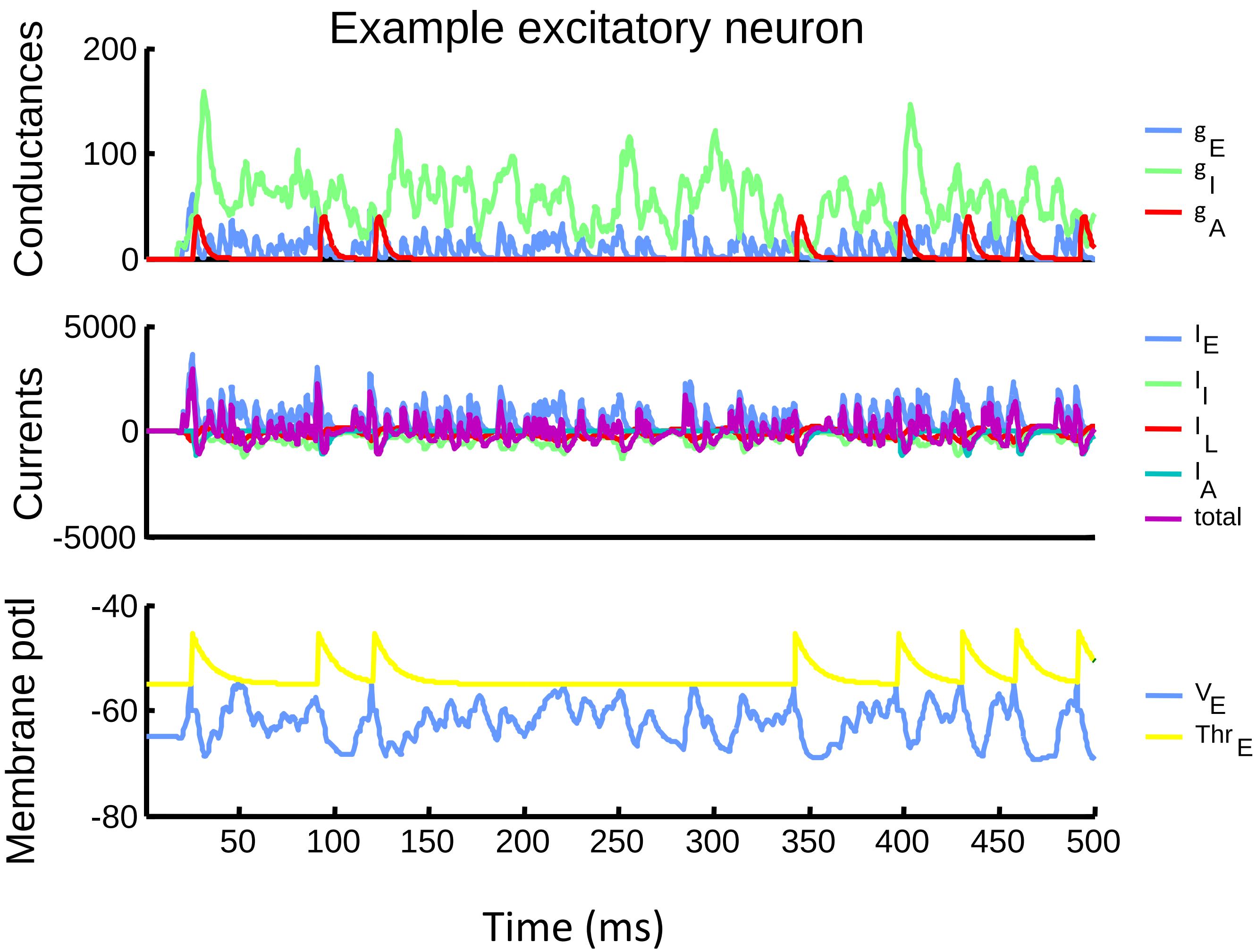
- Instead of $\mathbf{r}_{AV} = \mathbf{r}_A + \mathbf{r}_V$
- Linear combination $\mathbf{r}_{AV} = \mathbf{W}_A\mathbf{r}_A + \mathbf{W}_V\mathbf{r}_V$
- $\mathbf{W}_A$ and $\mathbf{W}_v$ are synaptic weights, fixed across trials.
- → Very general scheme for optimal cue combination.

So far: adding spike counts.

Does this work for realistic neurons?

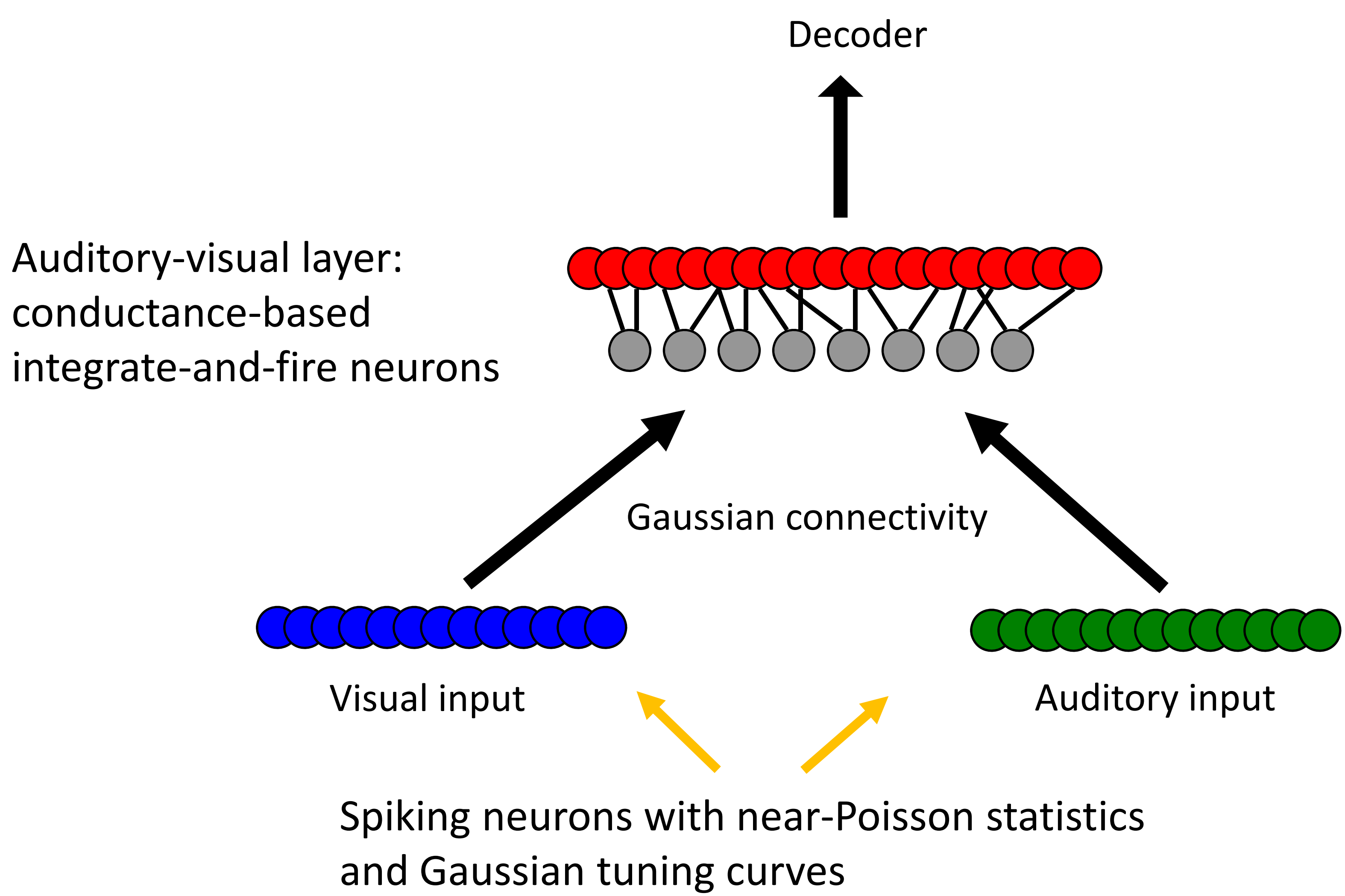


Conductance-based integrate-and-fire neuron



$$C\frac{dV_i}{dt} = -g_L\left(V_i(t) - E_L\right) - g_{iE}(t)\left(V_i(t) - E_E\right) - g_{iI}(t)\left(V(t) - E_I\right) - g_{iA}(t)\left(V_i(t) - E_A\right)$$

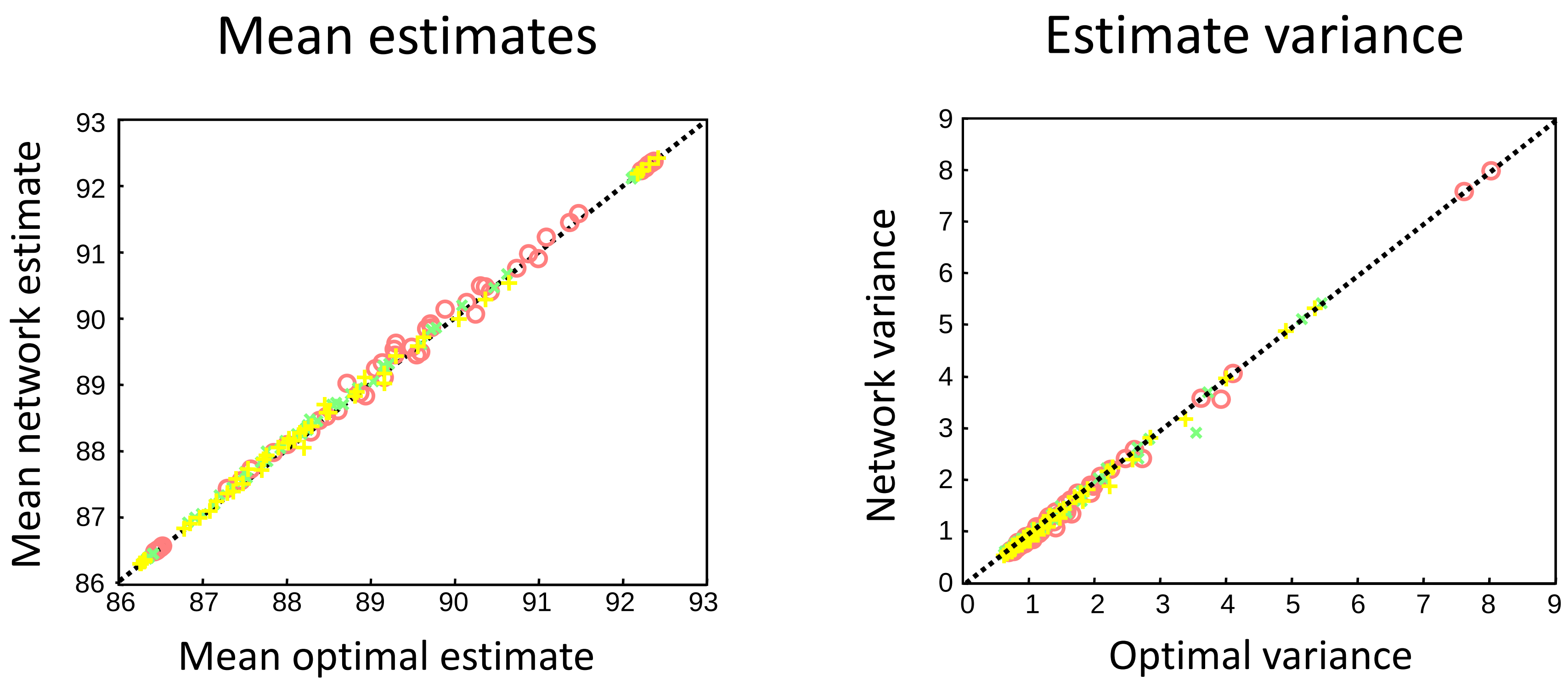
Integrate-And-Fire Network



Simulating cue combination

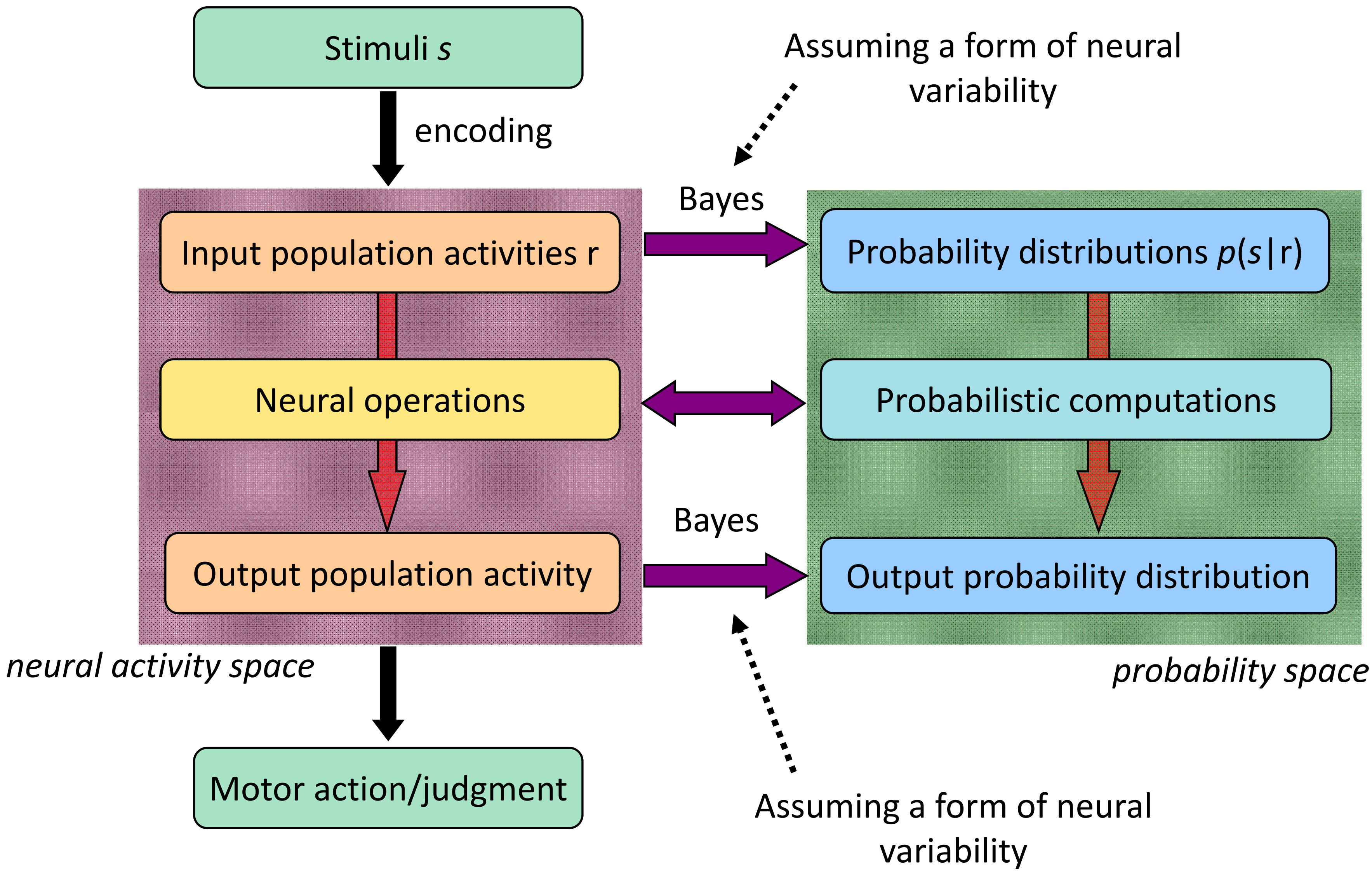
1. Visual input only. Obtain mean and variance of visual estimates from output spike counts.
2. Repeat for auditory input only.
3. Multisensory input. Obtain mean and variance of estimates.
4. Compare mean and variance from step 3 with optimal combinations based on steps 1 and 2.

Network performs near-optimally



o Same tuning curves, same covariance matrices
+ Same tuning curves, different covariance matrices
x Different tuning curves, different covariance matrices

Framework for optimal neural computation



Conclusions

- Poisson variability can be generalized to Poisson-like variability.
- For Poisson-like variability, a linear combination of population activity implements optimal cue combination.
- Mappings between Bayesian operations on probability distributions and neural operations can be found.
- For more complex computations, a wide open area → machine learning algorithms?