IJCNN 2006

Modeling Cortical Maps With Feed-Backs

Thierry Viéville and Pierre Kornprobst

ÉCOLE NORMALE SUPÉRIEURE
Département d'Informatique

Projet Odyssée
INRIA

École Nationale des Ponts et Chaussées
CERTIS
Centre d'Enseignement et de Recherche en Technologies de l'Information et Systèmes

Goals

- Implement a general variational formulation as a neural network
- Relate an optimization problem to a cortical map estimation
- Consider the coupling between several cortical maps

Variational Framework

Inputs w How? v Infered "cause"

- Minimization of an energy written in a continuous setting

$$\bar{v} = \underset{v \in H(\Omega) / c(v) = 0}{\operatorname{Argmin}} \mathcal{L}(v)$$

$$\boxed{v \in H(\Omega)}$$

General Variational Formulation

- We propose

$$\bar{v} = \operatorname*{Argmin}_{v \in H / c(v)=0} \mathcal{L}(v),$$

with

$$\mathcal{L}(v) = \boxed{\int_\Omega |\hat{w} - w|^2_{\mathbf{\Lambda}}} + \boxed{\int_\Omega \phi(|\nabla v|_L)} + \boxed{\int_\Omega \psi(v),}$$

and

$$\boxed{\hat{w} = P\,v}$$

General Variational Formulation

- The different terms

Nonlinearities

$$\bar v = \operatorname*{Argmin}_{v\in H / \boxed{c(v)=0}} \mathcal{L}(v),$$

Control of the solution

with

$$\mathcal{L}(v) = \boxed{\int_\Omega |\hat w - w|^2_{\boldsymbol\Lambda}} + \boxed{\int_\Omega \phi(|\nabla v|_L)} + \boxed{\int_\Omega \psi(v)},$$

and

$$\boxed{\hat w = P\, v}$$

Λ = Measurement information metric

Inputs from outputs Re-estimation

L = Regularization Metric

General Variational Formulation

- Λ and L are constants (now...)

$$\bar{v} = \underset{v \in H / c(v)=0}{\operatorname{Argmin}} \ \mathcal{L}(v),$$

with

$$\mathcal{L}(v) = \boxed{\int_\Omega |\hat{w} - w|^2_{\mathbf{\Lambda}}} + \boxed{\int_\Omega \phi(|\nabla v|_L)} + \int_\Omega \psi(v),$$

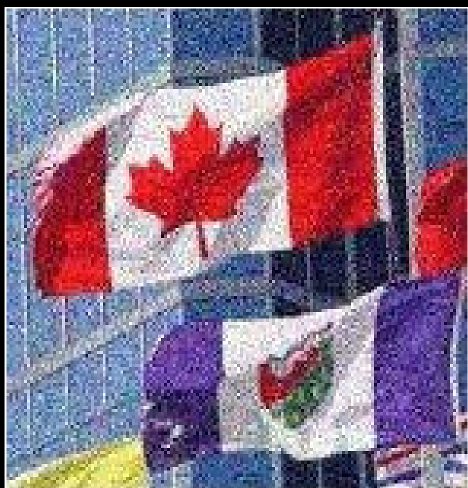
and

$$\hat{w} = P\, v$$

Λ = Measurement information metric

L = Regularization Metric

Covers Several Applications



- For example, image smoothing

$$\bar{v} = \operatorname*{Argmin}_{v \in H / \sum_i v_i = \mathrm{cte}} \mathcal{L}(v),$$

with

$$\mathcal{L}(v) = \int_\Omega |\hat{w} - w|^2_\Lambda + \boxed{\int_\Omega \phi(|\nabla v|_{\mathbf{L}})},$$

and

$$\hat{w} = v$$

Covers Several Applications

- For example, winner-take-all

$$\bar{v} = \underset{v}{\mathrm{Argmin}} \mathcal{L}(v),$$

$$\text{with}$$

$$\mathcal{L}(v) = \int_\Omega |\hat{w} - w|^2_{\mathbf{\Lambda}} + \int_\Omega |\nabla \mathbf{v}|^2 + \boxed{\int \psi(v)}$$

$$\text{and}$$

$$\hat{w} = v$$

How to Go to Discrete Domain?

$$\boxed{\inf_{v \in H(\Omega)/c(v)=0} \mathcal{L}(v)}$$

$$\frac{\partial v}{\partial t} = F\left(v, v_x, v_{xx}, \ldots\right)$$

$$\boxed{\frac{dv_i}{dt} = G(v_i, v_j)}$$

INRIA

Goal: Establish the Correspondence...

$$\inf_{v \in H(\Omega)/c(v)=0} \mathcal{L}(v)$$

$$\mathcal{L}(v) = \boxed{\int_\Omega |\hat{w} - w|^2} + \boxed{\int_\Omega \phi(|\nabla v|_L^2)} + \boxed{\int_\Omega \psi(v)}$$

Weighted Particular Methods

$$\frac{dv_i}{dt} = \boxed{+\kappa_i \, \mathbf{w}_i} + \boxed{\sum_j \sigma_{ij}(v_i) \, v_j} - \boxed{\epsilon_i(v_i)}$$

Weighted Particular Methods

- Litterature: Degond and Mas-Gallic (89), Cottet and Ayyadi (98), Edwards (96)

- General Idea

$$
\begin{array}{cccc}
L & \int_\Omega \phi(|\nabla v|_{\mathbf{L}}) & \Longrightarrow & \mathrm{div}(L\nabla v) \\
 & \Downarrow & & \Downarrow \\
\sigma_{ij} & \sum_j \sigma_{ij} v_j & \Longleftarrow & \int_\Omega \sigma(x,y) v(y) dy
\end{array}
$$

Weighted Particular Methods

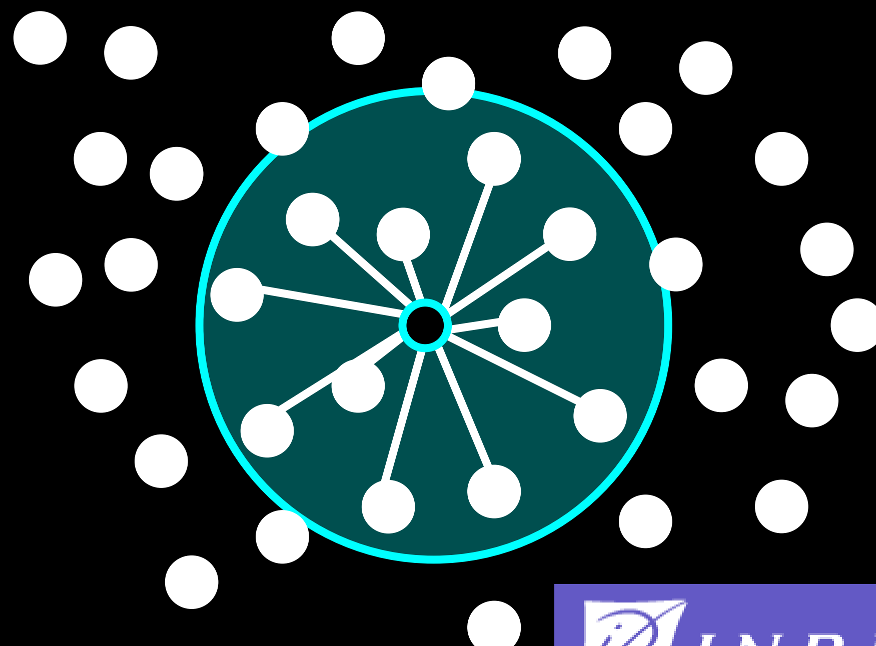
- Litterature: Degond and Mas-Gallic (89), Cottet and Ayyadi (98), Edwards (96)

- General Idea

L $\quad$ $$\int_\Omega \phi(|\nabla v|_{\mathbf{L}})$$

$$\Downarrow$$

σ_{ij} $\quad$ $$\sum_j \sigma_{ij} v_j$$



Our Contribution

- Extend result for vectorial case
- Works with nonlinear terms
  - We linearize

$$\bar{\mathbf{L}} = \phi'\left(\left|\nabla \mathbf{v}\right|_{\mathbf{L}}\right)\mathbf{L},$$

  - Linearizing, iterating and updating the smoothing term leads to the solution

Our Contribution

- We recover some compatibility conditions

$$
\begin{aligned}
\bar{\mathbf{L}}^{kl}(\mathbf{x}) &= \tfrac{1}{2}\sum_j \sigma_j \, \bar{\mu}_j^{\mathbf{e}_k+\mathbf{e}_l}(\mathbf{x}), \\
\mathbf{div}^k(\bar{\mathbf{L}}(\mathbf{x})) &= \sum_j \sigma_j \, \bar{\mu}_j^{\mathbf{e}_k}(\mathbf{x}),
\end{aligned}
$$

$$
\forall i, \quad \sum_j \sigma_{ij} \, \bar{\mu}_j^{\alpha}(\mathbf{x}) = 0 \quad 2 < |\alpha| \leq r
$$

- But also, we define an optimal solution, as close a possible to the differential operator

Our Contribution

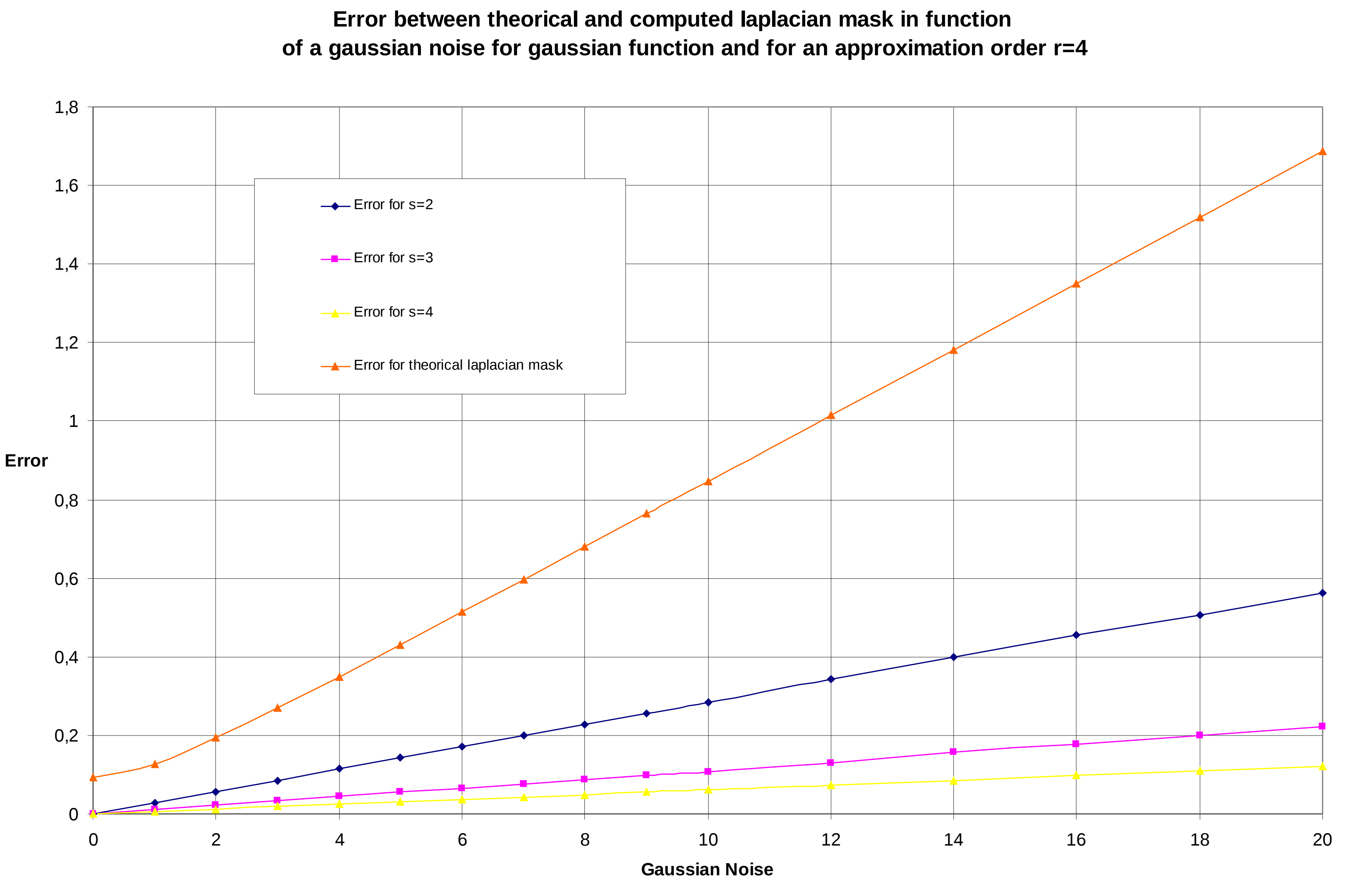
- How to prove it?

  Interest of that approach: Using a weak assumption about neuronal field geometry yields an unbiased discrete form

$$\mathbf{v}_j = \int_{\mathcal{S}_j} \mathbf{v}(\mathbf{y})\, \mu_j(\mathbf{y})\, d\mathbf{y}$$

INRIA

Approximation Error Rates



General Network Compilation Rule

- Convergence of the operator yields the convergence of solutions (L is positive but not necessarily symmetric).
- General family of solution for bounded neighbourhoods
- Definition of an optimal solution

$$\text{with } \sum_{ij} \sigma_{ij}^2 \text{ minimal}$$

(close to the differential operator).

Conversely....

- Given the network dynamic

$$\frac{\partial \mathbf{u}_i}{\partial t} = -\epsilon_i(\mathbf{u}_i) + \sum_j \sigma_{ij}(\mathbf{u}_i)\,\mathbf{v}_j + \kappa_i\,\mathbf{w}_i \text{ with } \mathbf{v}_i = Sig(\mathbf{u}_i),$$

As soon as the weights verify (C2), the NN locally minimizes the criterion

$$\mathcal{L}(v) = \int_\Omega |\hat{\mathbf{w}} - \mathbf{w}|^2 + \int_\Omega |\nabla \mathbf{v}|^2_{\mathbf{L}} + \int_\Omega \psi(\mathbf{v}),$$

With suitable weights.

Consequences

- This representation is well-defined for neural network connectivity with short-range excitatory connections

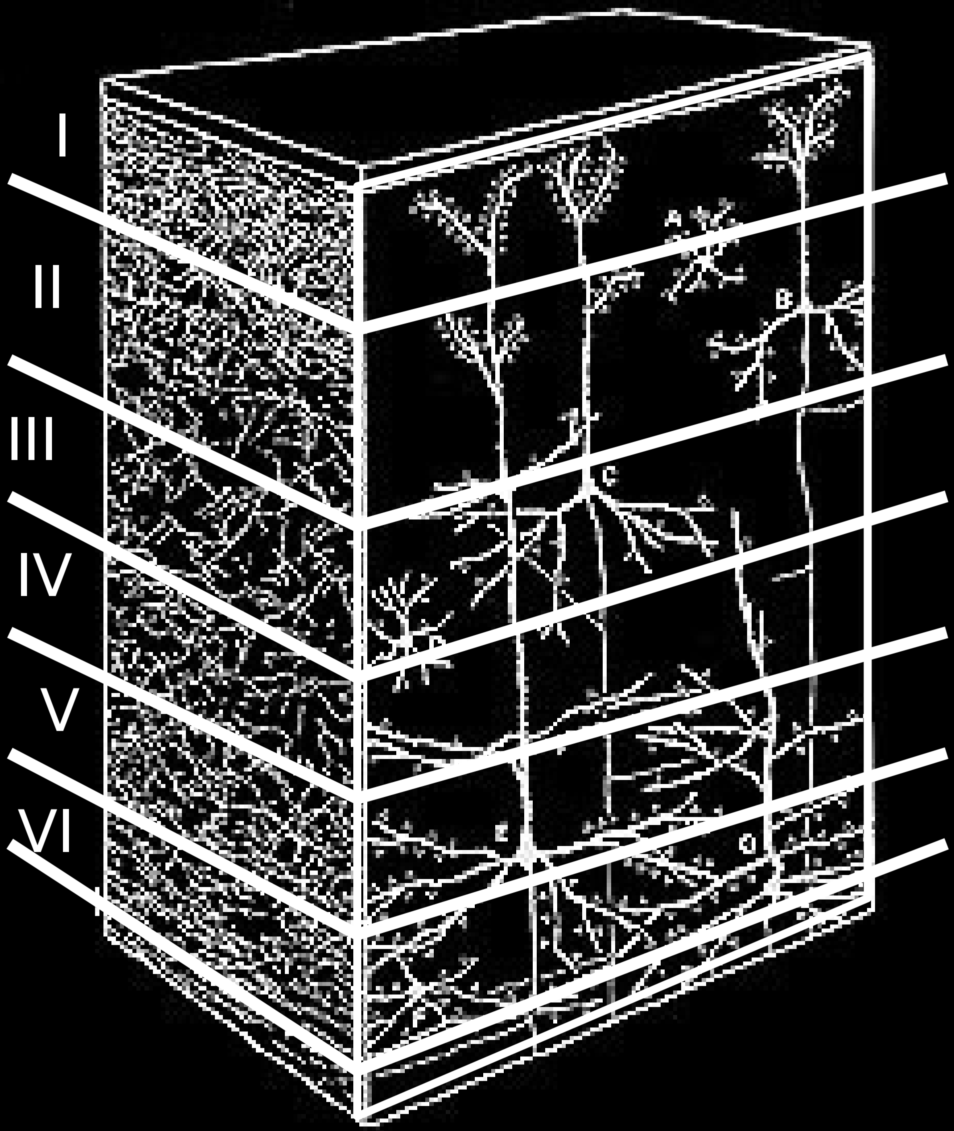
- Applicable to several classical NN, e.g., Hopfield, Cohen-Grossberg

- No more restriction on weight symmetry

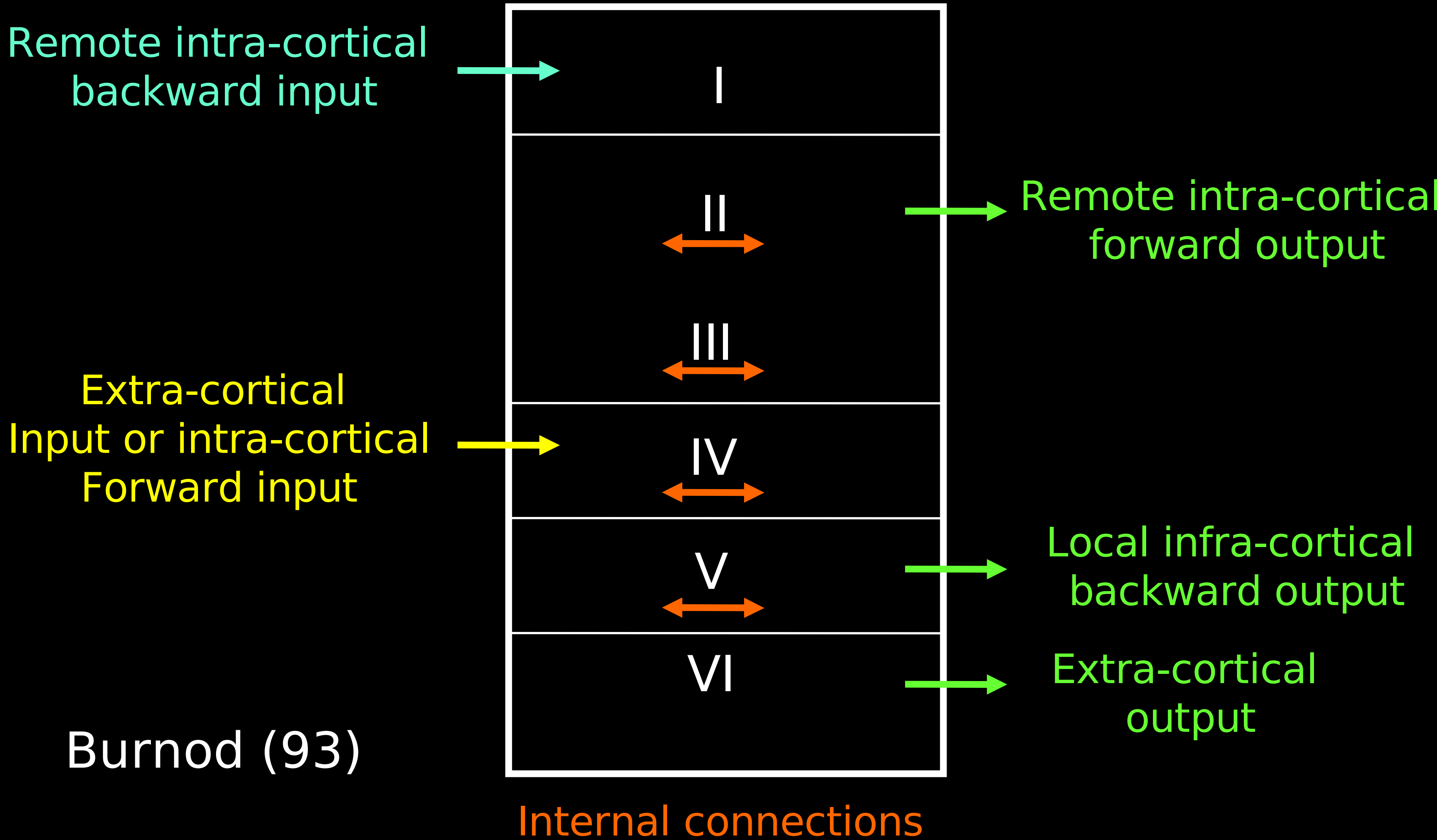
Back to Goals

- ✓ Implement a general variational formulation as a neural network
- Relate an optimization problem to a cortical map estimation
- Consider the coupling between several cortical maps

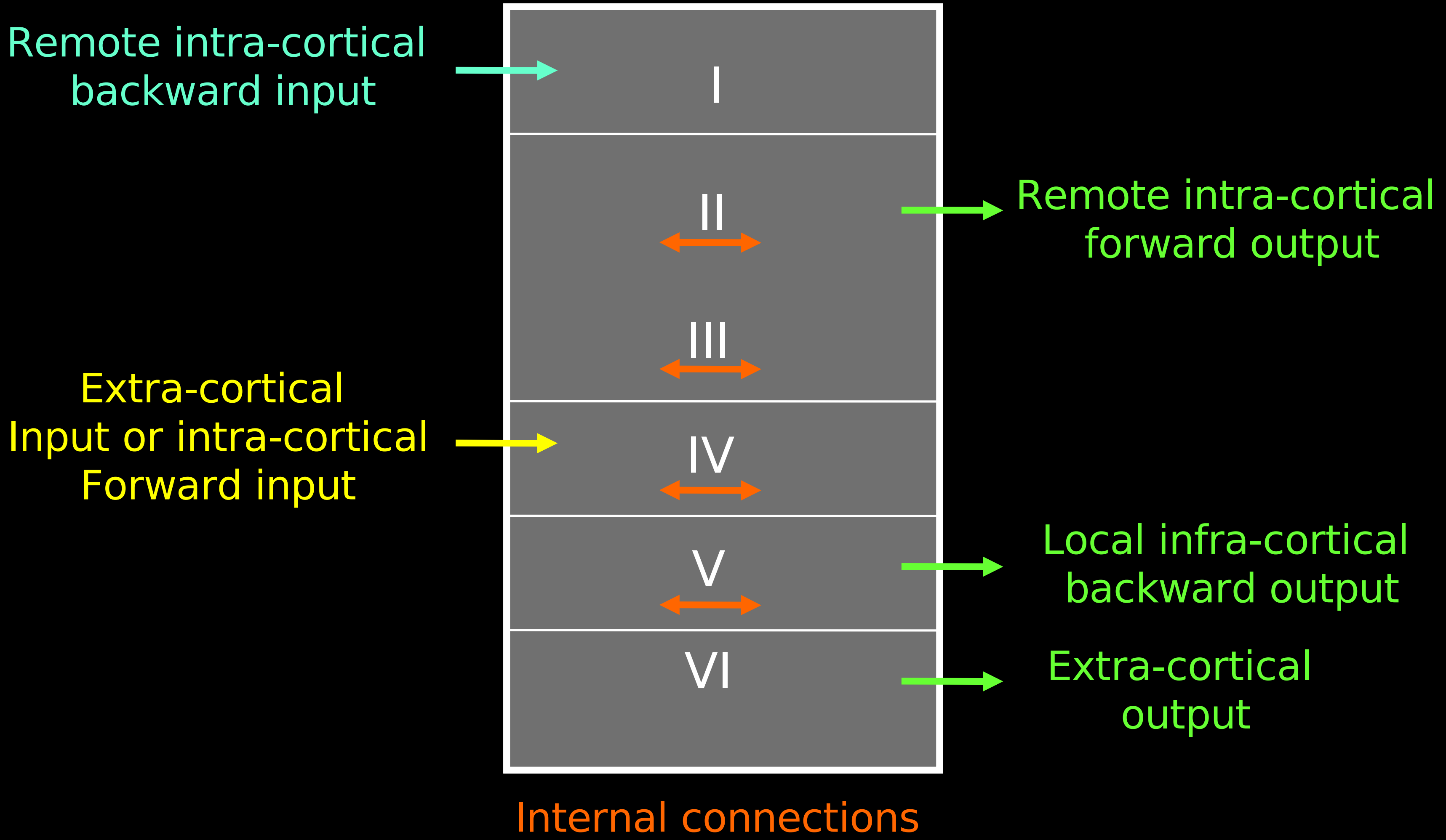
Cortical Hypercolumns



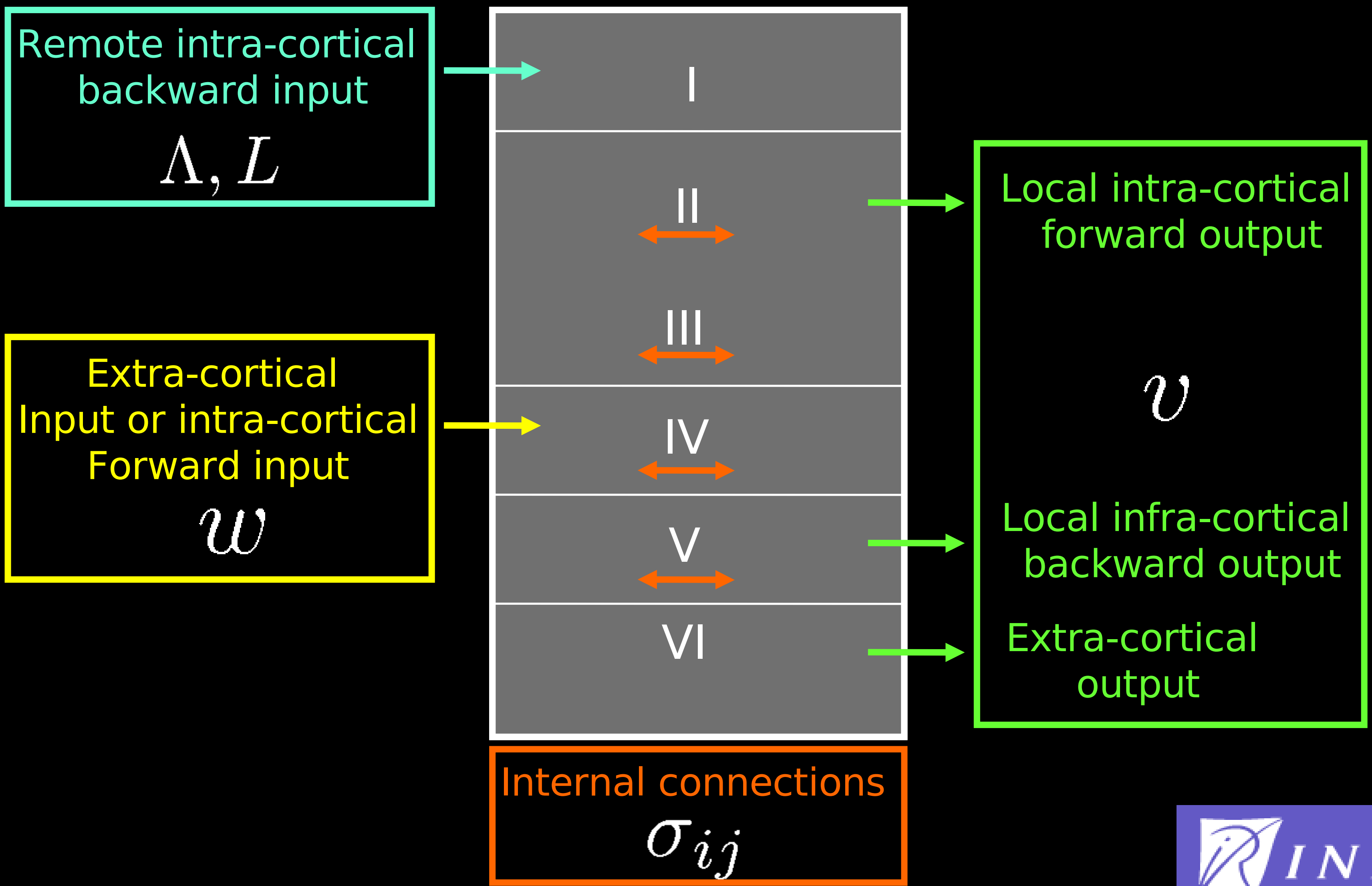
Cortical Hypercolumns



We Model an Averaged Hypercolumn



We Model an Averaged Hypercolumn



Finding Causes from Inputs...

- Related to Dayan and Abbott (01), Friston (02)

Input Prediction
(Backward connection)

P

Inputs w **Optimization Approach** v Infered "cause"

Λ, L

Cause representation
(Forward connection)

Back to Goals

- ✓ Implement a general variational formulation as a neural network
- ✓ Relate an optimization problem to a cortical map estimation
- Consider the coupling between several cortical maps

What can we say for several maps interacting together?

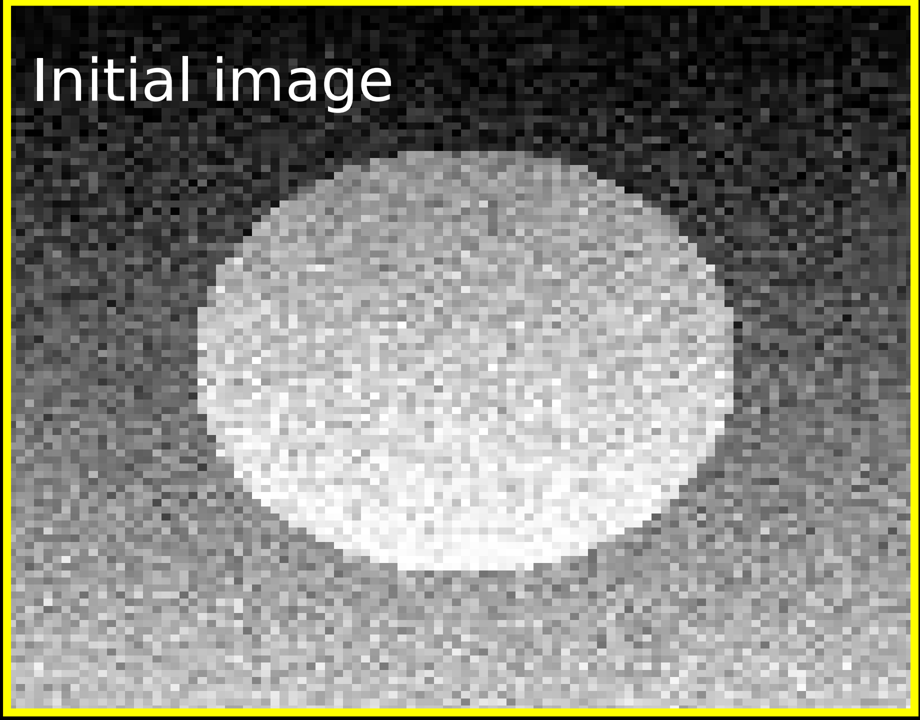
$$\begin{tikzcd} w_m \arrow[r, bend right, "\Lambda_m, L_m"'] & v_m \arrow[l, bend right, "P_m"'] \end{tikzcd}$$

$$\inf_{\mathbf{v}_m \in H_m = 0} \mathcal{L}_m(\mathbf{v}_m)$$

$$\begin{tikzcd} w_k \arrow[r, bend right, "\Lambda_k, L_k"'] & v_k \arrow[l, bend right, "P_k"'] \end{tikzcd}$$

$$\inf_{\mathbf{v}_k \in H_k = 0} \mathcal{L}_k(\mathbf{v}_k)$$

What can we say for several maps interacting together?



Edge map used for restoration

$$\frac{\partial v}{\partial t} = \operatorname{div}\left(D\nabla v\right)$$

$$D = c(|\nabla v|)$$

What can we say for several maps interacting together?

Edge map used for restoration

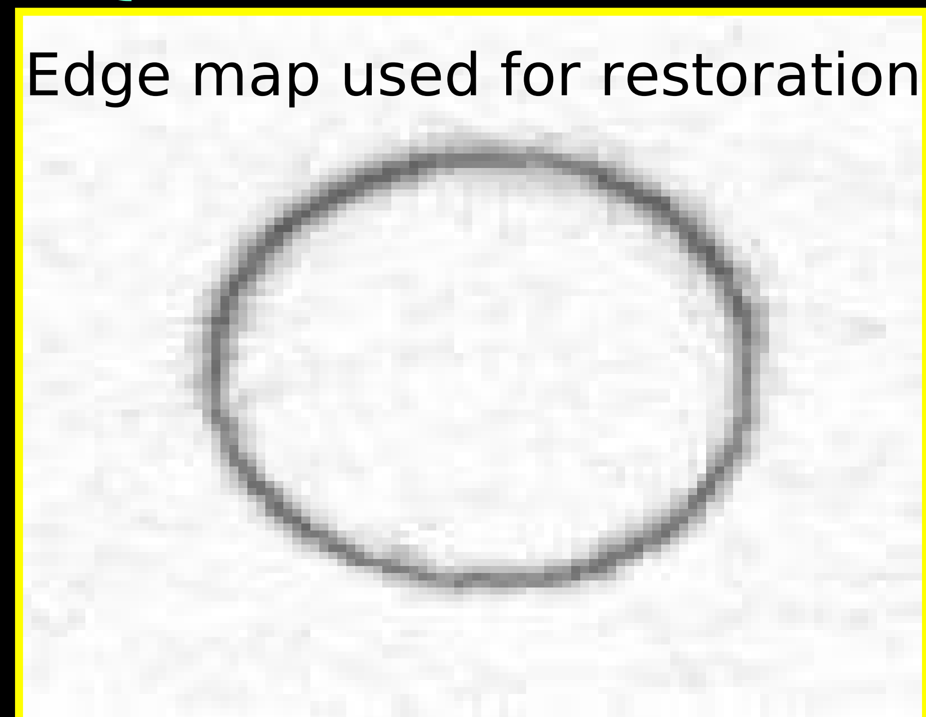
$$\frac{\partial v}{\partial t} = \mathrm{div}\left(D \nabla v\right)$$

$$D = c(|\nabla v|)$$

What can we say for several maps interacting together?

$$\frac{\partial v}{\partial t} = \operatorname{div}\left(D \nabla v\right) \qquad D = c(|\nabla v|)$$

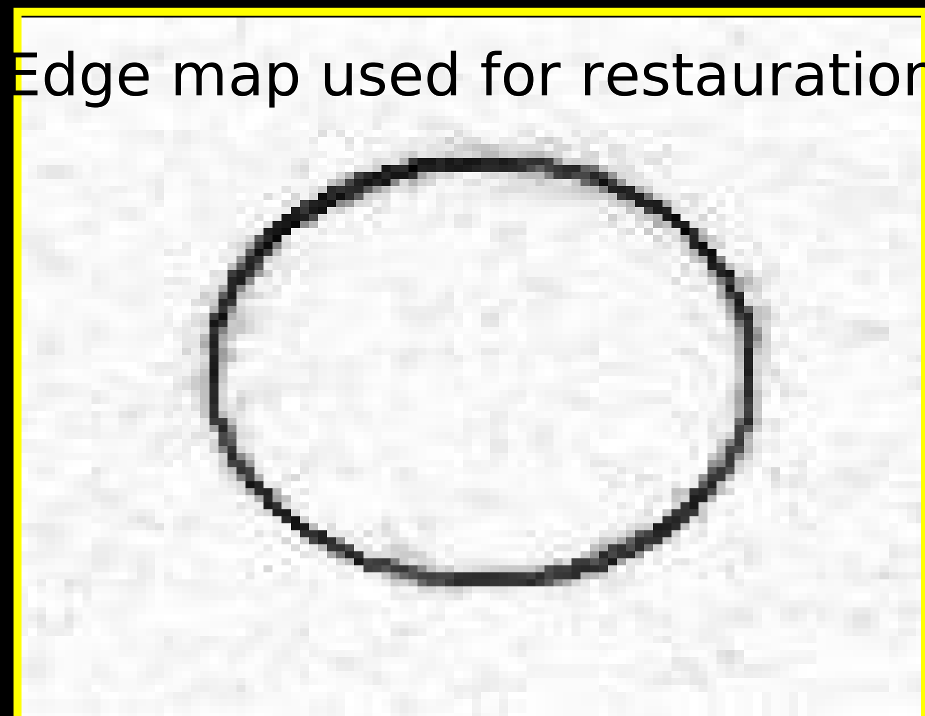
What can we say for several maps interacting together?



$$\frac{\partial v}{\partial t} = \operatorname{div}\left(D \nabla v\right)$$

$$D = c(|\nabla v|)$$

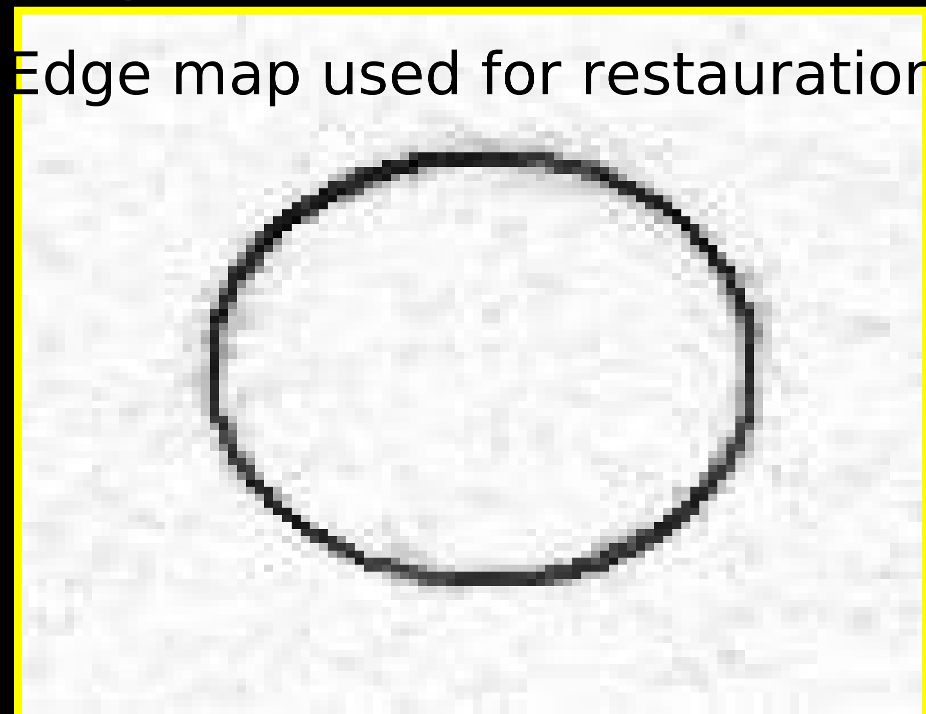
What can we say for several maps interacting together?



$$\frac{\partial v}{\partial t} = \mathrm{div}\,(D\nabla v)$$

$$D = c(|\nabla v|)$$

What can we say for several maps interacting together?



$$\frac{\partial v}{\partial t} = \mathrm{div}\left(D \nabla v\right)$$

$$D = c(|\nabla v|)$$

What Do We Learn From Biology?

Forward = Driving		Backward =
sparse axonal bifurcation	≠	frequent bif
patchy axonal terminations		al terminations
topographic projec		non-topographic projections
one-to-o divergence	≠	large spatial divergence
define a lattice	≠	transcend several levels
		slow time-constants
		more numerous

We keep two constraints

What Do We Learn From Biology?

- Forward connections define an acyclic graph

$$w_{m'} = \rho(v_m)$$

- Feedback values are smoothed in space, before influencing other maps

$$\Lambda_m(\mathcal{S} * \rho(v_\bullet)) \text{ and } L_m(\mathcal{S} * \rho(v_\bullet))$$

$v_\bullet$ ➡ Rectification ρ ➡ Smoothing kernel $\mathcal{S}$

Link with Well-posedness...

$$\inf_v \int_\Omega \phi(|\nabla v|)$$

- Perona and Malik ani diffusion model (1990) is ill-posed

$$\frac{\partial v}{\partial t} = \operatorname{div}\left(c(|\nabla v|)\nabla v\right)$$

- The regularization of the PM model, by Catte etal becomes well-posed!

$$\frac{\partial v}{\partial t} = \operatorname{div}\left(c(|G_\sigma \star \nabla v|)\nabla v\right)$$

Our Result

- Locally minimizing the criterion $\mathcal{L}_m$ with respect to $\mathbf{v}_m$ is equivalent to locally minimize with respect to $\mathbf{v}_\bullet$, in the general case:

$$\mathcal{L}_\bullet = \sum_m \lambda(|\nabla_m \mathcal{L}_m|)\, \mathcal{L}_m$$

writing $\nabla_m = \partial / \partial \mathbf{v}_m$ and $\lambda(\cdot) : \mathcal{R}^+ \rightarrow \mathcal{R}^+$ a positive strictly increasing profile with $\lambda(u) \geq 0, \lambda'(u) > 0, \lambda(0) = 0$ and $\lim_{u \rightarrow 0} \lambda'(u)/u < +\infty$., (e.g. $\lambda(u) = u^\alpha$ with $\alpha > 2$).

Common Objective for Different Cortical Maps

- Feedback links in our framework yields a well-defined process
- We need a ψ convex in this context

- Backward connections change the processing of a cortical map
- Forward connections act as a “data propagation”

INRIA

Conclusion

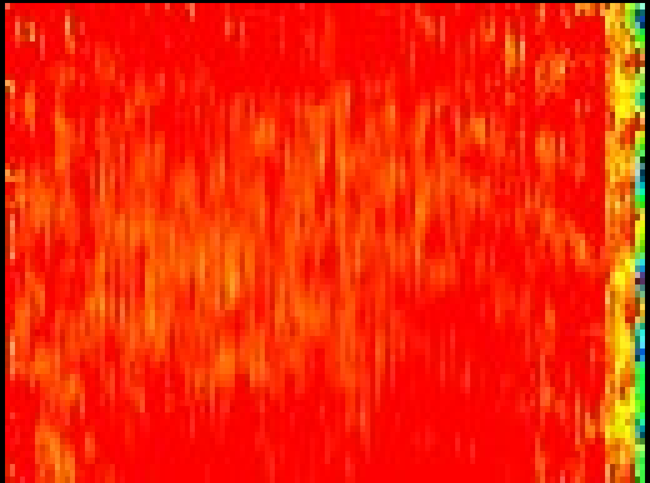
- ✓ Implement a general variational formulation as a neural network

- ✓ Relate an optimization problem to a cortical map estimation

- ✓ Consider the coupling between several cortical maps

Perspectives

- Apply this framework for visual appli (Mumford Shah)
- From an optimization approach to an **spiking** neural network
- Cf abstract ecvp!



Thank you!

This work was partially supported by the
EC IP project FP6-015879, FACETS

ÉCOLE NORMALE SUPÉRIEURE
Département d'Informatique

Projet Odyssée
INRIA

École Nationale des Ponts et Chaussées
CERTIS
Centre d'Enseignement et de Recherche en Technologies de l'Information et Systèmes