

Collective computation in nonlinear networks
and the grammar of evolvability

Jean-Jacques Slotine
M.I.T.

10¹¹

10^{11}

100 billion neurons

10⁴

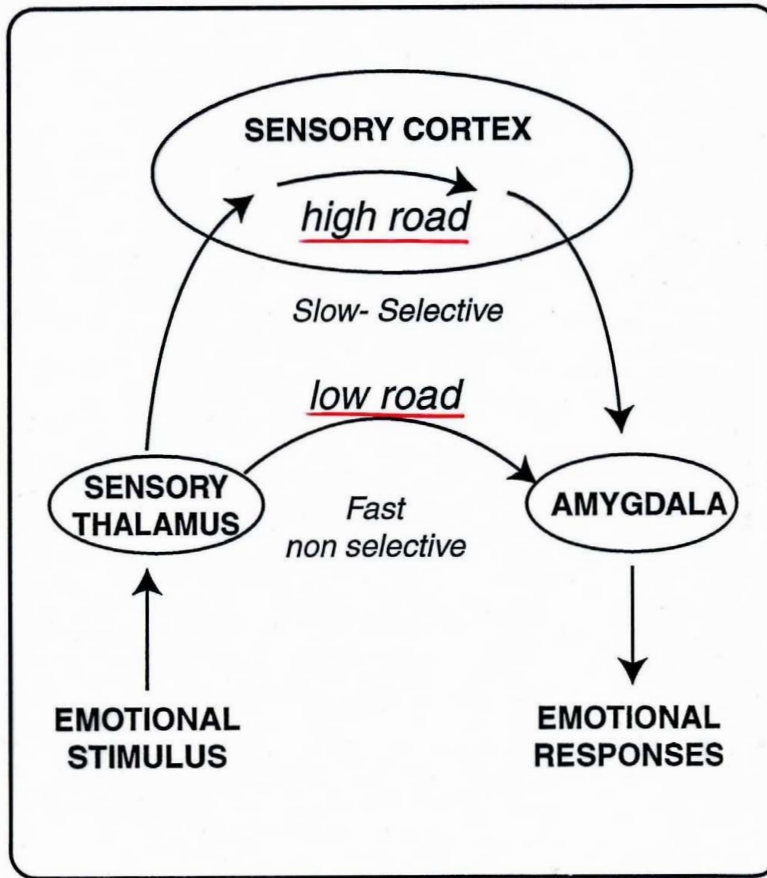
10^4

10 000 connections per neuron

10⁷

10^7

**10 million times
slower than a transistor**



Prevent the inappropriate response

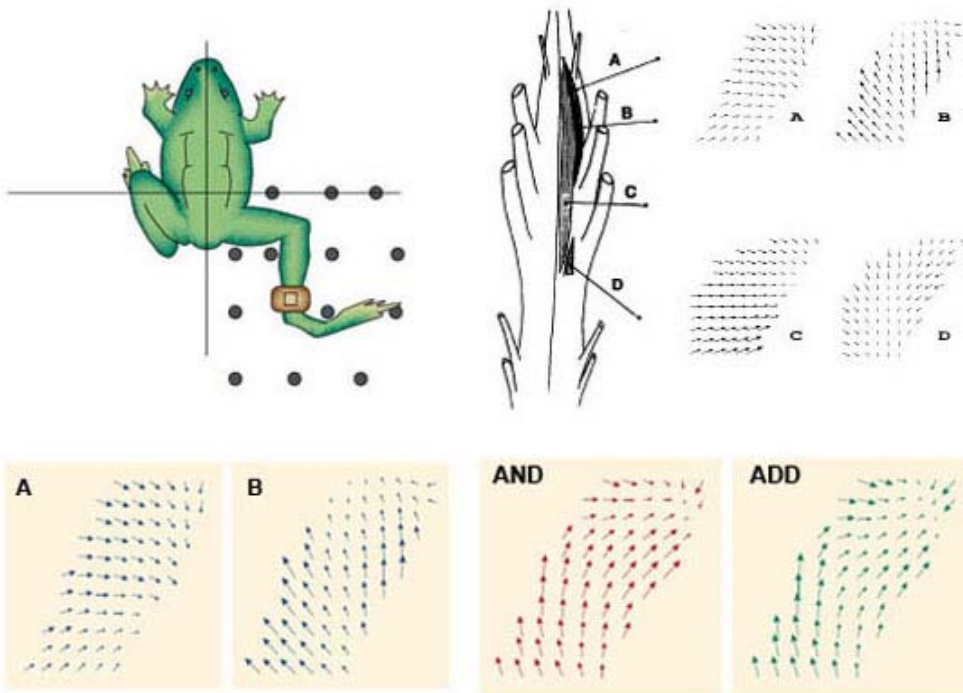
Produce the appropriate response

J. Ledoux
The Emotional Brain, 1996

In immune system

- innate immunity (ancient, fast)
- adaptive immunity (in vertebrates, slower, memory)

Experiments on Frogs



Bizzi et al., 1991, 1995

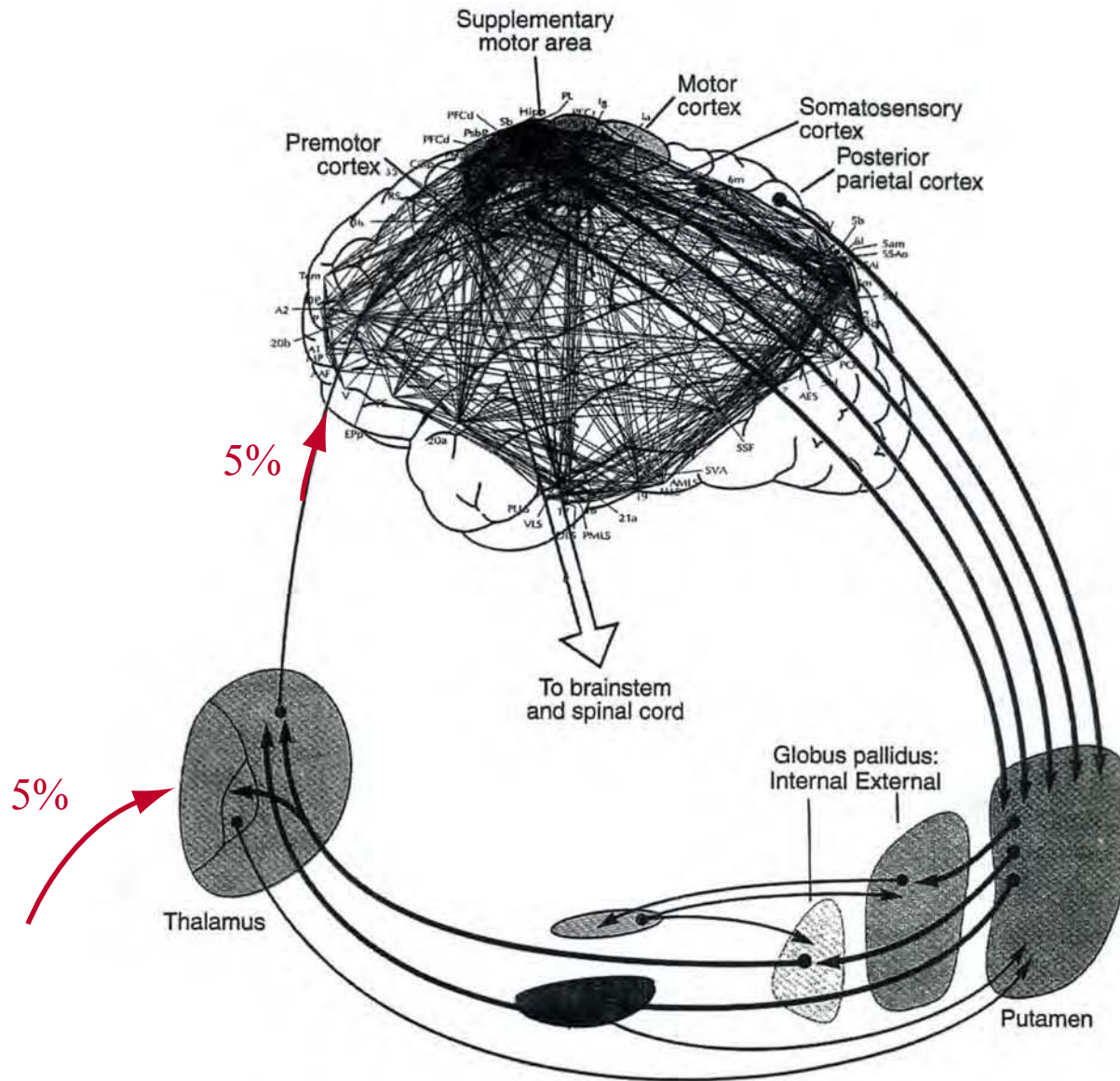
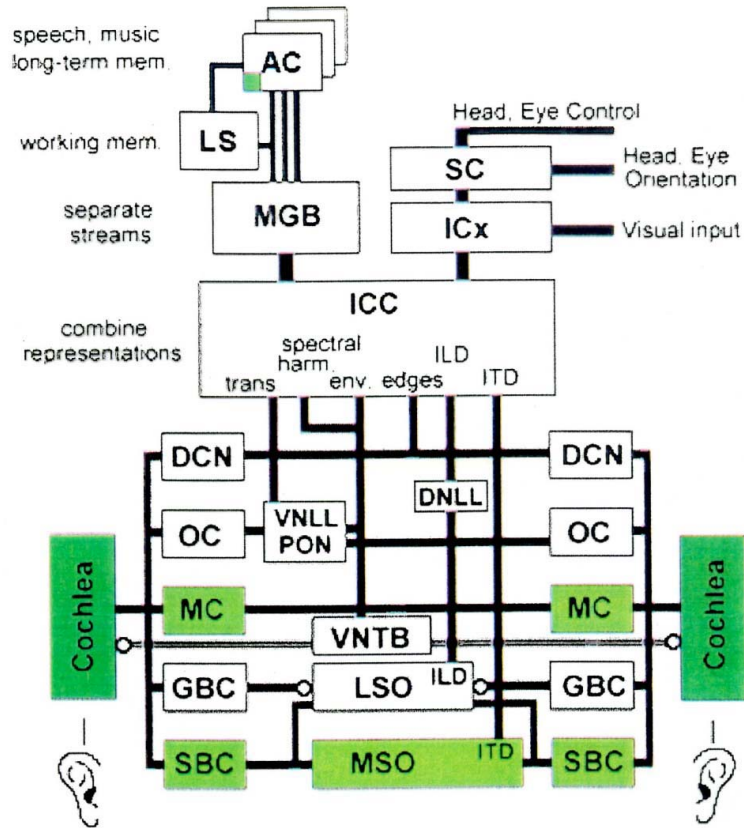


FIGURE 14.1 STRUCTURES AND CONNECTIONS MEDIATING CONSCIOUS AND UNCONSCIOUS PROCESSES.

Edelman and Tononi, 2000

In genetics

- 98% of human genome does not encode protein sequences
- Used in part for regulation and control



Lloyd Watts, 2002

Modularity/Feedback

Prediction

Prediction

Prediction

Catching

Prediction

Catching, avoiding obstacles

Prediction

Catching, avoiding obstacles

Waking up

Prediction

Catching, avoiding obstacles

Waking up

Pavlov

Prediction

Catching, avoiding obstacles

Waking up

Pavlov

Placebo effect

Prediction

Catching, avoiding obstacles

Waking up

Pavlov

Placebo effect

Wagner operas

Prediction

Catching, avoiding obstacles

Waking up

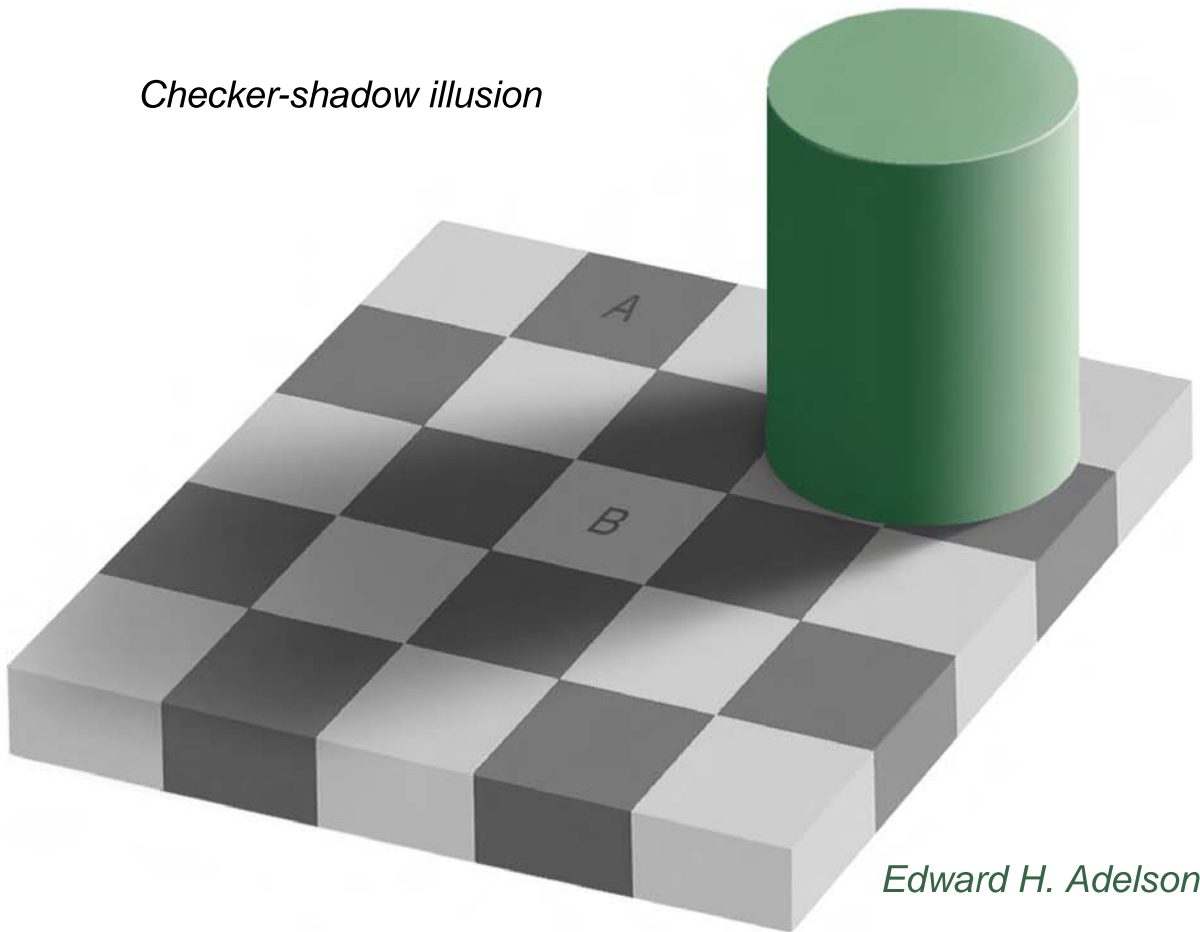
Pavlov

Placebo effect

Wagner operas

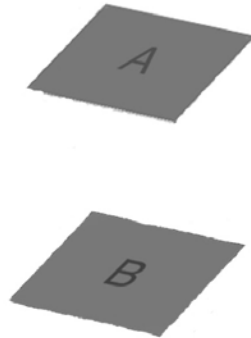
Illusions

Checker-shadow illusion



Edward H. Adelson

Checker-shadow illusion



Edward H. Adelson

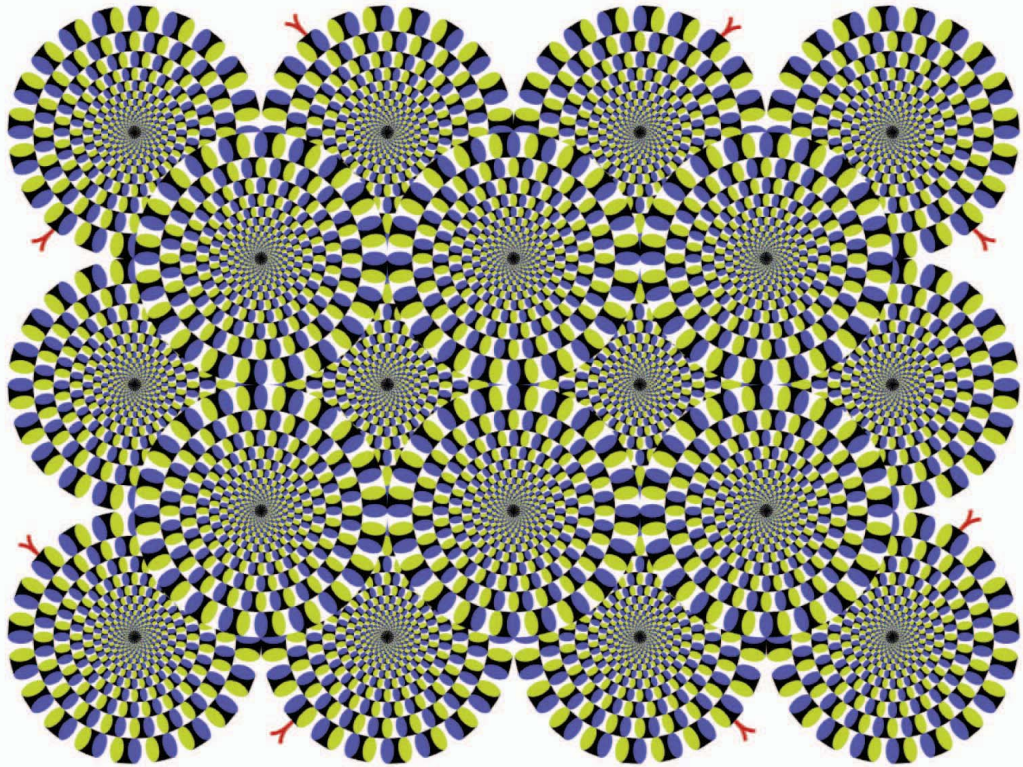
Modularity/Feedback

Prediction

Synchronization

Synchronization models neural computations at many scales

- Coincidence detection, pattern matching
- Mirror neuron response
- Signal transmission and restoration
- Temporal binding of multi-sensory data
- Attention and priming



Akiyoshi Kitaoka, 2004

Collective computation and learning in nonlinear networks
and the grammar of evolvability

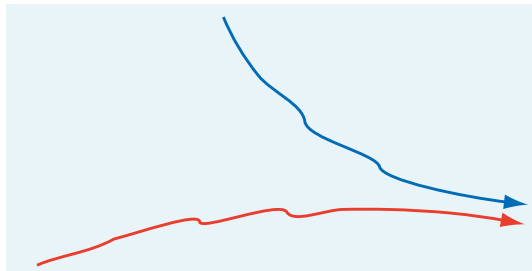
Jean-Jacques Slotine
M.I.T.

Nonlinear Contraction

Lohmiller and Slotine, Automatica 34(6), 1998

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$$

Contraction: Any two solutions converge exponentially

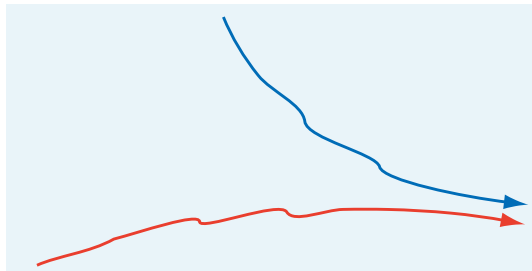


Nonlinear Contraction

Lohmiller and Slotine, Automatica 34(6), 1998

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$$

Contraction: Any two solutions converge exponentially
if Jacobian is negative definite in some metric



Contraction Theory

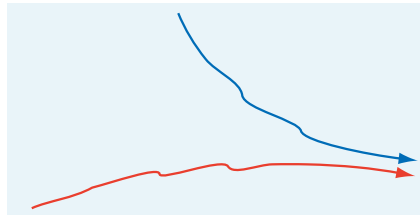
Lohmiller and Slotine, *Automatica* 34(6), 1998

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$$

If $\exists \Theta(\mathbf{x}, t)$ such that, uniformly $\forall \mathbf{x}, \forall t \geq 0$,

$$\mathbf{F} = \left(\dot{\Theta} + \Theta \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right) \Theta^{-1} < \mathbf{0} \quad \Theta(\mathbf{x}, t)^T \Theta(\mathbf{x}, t) > \mathbf{0}$$

then any two trajectories converge exponentially.



Proof: Consider virtual displacements, $\delta \mathbf{x} = \frac{\partial \mathbf{x}(t, \mathbf{x}_o)}{\partial \mathbf{x}_o} d\mathbf{x}_o$

$$\delta \mathbf{z} = \Theta \delta \mathbf{x} \quad \frac{1}{2} \frac{d}{dt} \|\delta \mathbf{z}\|^2 \leq \lambda_{\mathbf{F}} \|\delta \mathbf{z}\|^2$$

and **path integration** at fixed time.

Observer

Lorenz attractor

$$\begin{cases} \dot{x} = \sigma (y - x) \\ \dot{y} = \rho x - y - x z \\ \dot{z} = -\beta z + x y \end{cases}$$

Observer

$$\begin{cases} \dot{\hat{y}} = \rho x - \hat{y} - x \hat{z} \\ \dot{\hat{z}} = -\beta \hat{z} + x \hat{y} \end{cases}$$

$$\mathbf{J} = \begin{bmatrix} -1 & -x \\ x & -\beta \end{bmatrix} < \mathbf{0}$$

Aggregation Properties

- **positive parallel**
- **negative feedback**
- **series and cascades**
- **translation and scaling**
- **all of the above**

Aggregation Properties

- positive parallel
- negative feedback
- series and cascades
- translation and scaling
- all of the above

- **Evolvability**

Plausibly favored by evolution.
Similarly, development.

Combinations of Contracting Systems

Parallel
$$\frac{d}{dt} \delta \mathbf{z} = \sum_i \alpha_i(t) \frac{d}{dt} \delta \mathbf{z}_i$$

with $\alpha_i(t) > 0$, same metric

Hierarchies
$$\frac{d}{dt} \begin{pmatrix} \delta \mathbf{z}_1 \\ \delta \mathbf{z}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{F}_1 & \mathbf{0} \\ \mathbf{G}_2 & \mathbf{F}_2 \end{pmatrix} \begin{pmatrix} \delta \mathbf{z}_1 \\ \delta \mathbf{z}_2 \end{pmatrix}$$

with \mathbf{G}_2 bounded

Feedback
$$\frac{d}{dt} \begin{pmatrix} \delta \mathbf{z}_1 \\ \delta \mathbf{z}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{F}_1 & \mathbf{G}_1 \\ \mathbf{G}_2 & \mathbf{F}_2 \end{pmatrix} \begin{pmatrix} \delta \mathbf{z}_1 \\ \delta \mathbf{z}_2 \end{pmatrix}$$

with $\lambda(\mathbf{F}_1) \lambda(\mathbf{F}_2) > \frac{1}{4} \min_{k>0} \sigma^2(k\mathbf{G}_1 + \mathbf{G}_2^T)$ uniformly

Composite Variables

$$\dot{s} = \phi(s, t) \quad \text{contracting by choice of control law}$$

$$\dot{\tilde{x}} + \lambda \tilde{x} = s \quad \text{contracting by definition of } s$$

Qualitative dynamics. For instance, react faster to larger errors

$$\dot{\tilde{x}} + (\lambda_1 + \lambda_2 |\tilde{x}|) \tilde{x} = s \quad \frac{\partial \dot{\tilde{x}}}{\partial \tilde{x}} = -(\lambda_1 + 2\lambda_2 |\tilde{x}|)$$

More generally $x^{(n)} = f(x, \dot{x}, \dots, x^{(n-1)}, u, t)$

Target contracting system $x^{(n)} = g(x, \dot{x}, \dots, x^{(n-1)}, t)$

Choosing

$$s = x^{(n-1)} - x_r^{(n-1)}$$

$$\frac{d}{dt} x_r^{(n-1)} = -k s + g$$

yields

$$x^{(n)} - g = \dot{s} + k s$$

Parallel Combinations

Control Primitives

Dynamics \mathbf{f} and primitives ϕ_i all contracting in the same $\Theta(\mathbf{x})$

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) + \sum_i \alpha_i(t) \phi_i(\mathbf{x}, t) \quad \alpha_i(t) > 0$$

More generally

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) + \mathbf{B}(\mathbf{x}, t) \mathbf{u}$$

Assume control primitives $\mathbf{u} = \mathbf{p}_i(\mathbf{x}, t)$ make the **closed-loop** system **contracting** in common metric, $\forall i$. Then any **convex combination**

$$\mathbf{u} = \sum_i \alpha_i(t) \mathbf{p}_i(\mathbf{x}, t) \quad \alpha_i(t) \geq 0 \quad \sum_i \alpha_i(t) \stackrel{\text{MIT 2.152}}{=} 1$$

yields a contracting dynamics in the same metric.

Entrainment

Contracting systems of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}(t))$$

where the input $\mathbf{u}(t)$ is **periodic** in time, converge towards a periodic state of the same period.

Robustness

Disturbed flow field

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) + \mathbf{d}(\mathbf{x}, t)$$

Radius of metric ball

$$\dot{R} + \lambda R \leq \|\Theta \mathbf{d}\|$$

Alternate Norms

$$\frac{d}{dt} \delta \mathbf{z} = \mathbf{F}(\mathbf{x}, t) \delta \mathbf{z}$$

$$\|\delta \mathbf{z}\|_2 = \sum_i (\delta z_i)^2 \quad \rightarrow \quad \frac{d}{dt} \|\delta \mathbf{z}\|_2 \leq \lambda_{max} \left(\frac{\mathbf{F} + \mathbf{F}^T}{2} \right) \|\delta \mathbf{z}\|_2$$

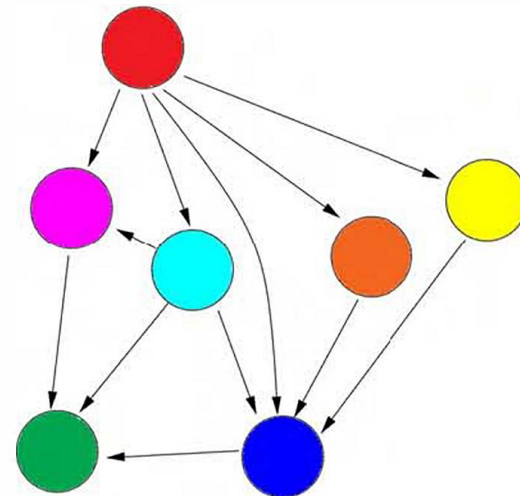
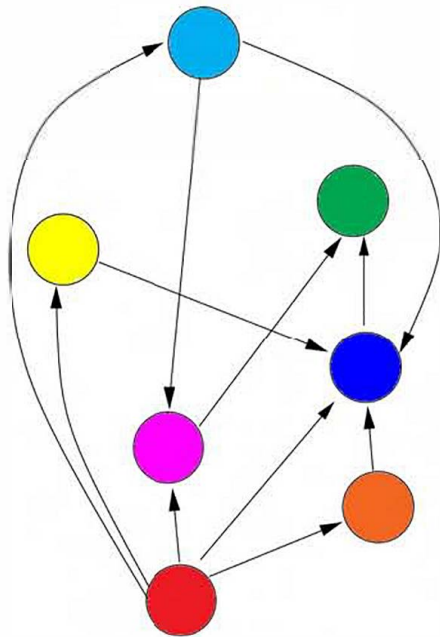
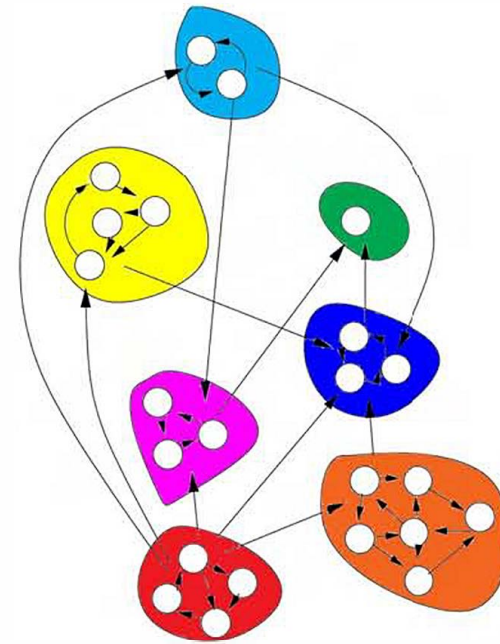
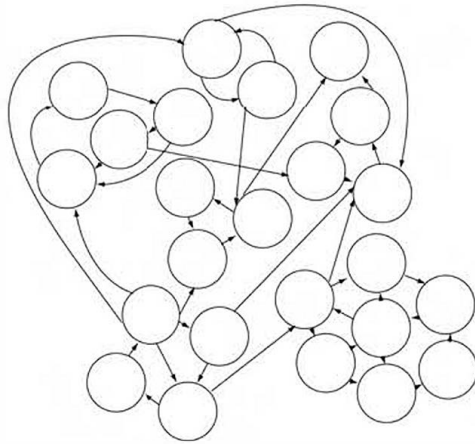
Other norms give diagonal dominance conditions

$$\|\delta \mathbf{z}\|_1 = \sum_i |\delta z_i| \quad \rightarrow \quad \frac{d}{dt} \|\delta \mathbf{z}\|_1 \leq \max_j (F_{jj} + \sum_{i \neq j} |F_{ij}|) \|\delta \mathbf{z}\|_1$$

$$\|\delta \mathbf{z}\|_\infty = \max_i |\delta z_i| \quad \rightarrow \quad \frac{d}{dt} \|\delta \mathbf{z}\|_\infty \leq \max_i (F_{ii} + \sum_{j \neq i} |F_{ij}|) \|\delta \mathbf{z}\|_\infty$$

Hierarchies of strongly connected components

Tabareau and Slotine, 2005



Facilitated Variation

Gerhart and Kirschner, P.N.A.S. 104, 2007

Three billion years ago, in early prokaryotic organisms

Components of energy metabolism, biosynthesis of the 60 building blocks, DNA replication, DNA transcription to RNA, translation of RNA to protein, lipid membrane synthesis, transmembrane transport

Two billion years ago, in early eukaryotic cells

Components of the formation of microfilament and microtubule cytoskeletons, motor proteins moving materials along the cytoskeletons, contractility processes, movement of the cell by cilia and ruffling membrane action, shuttling of materials between intracellular organelles, phagocytosis, secretion, chromosome dynamics, a complex cell cycle driven by protein kinases and protein degradation, sexual reproduction with meiosis and cell fusion

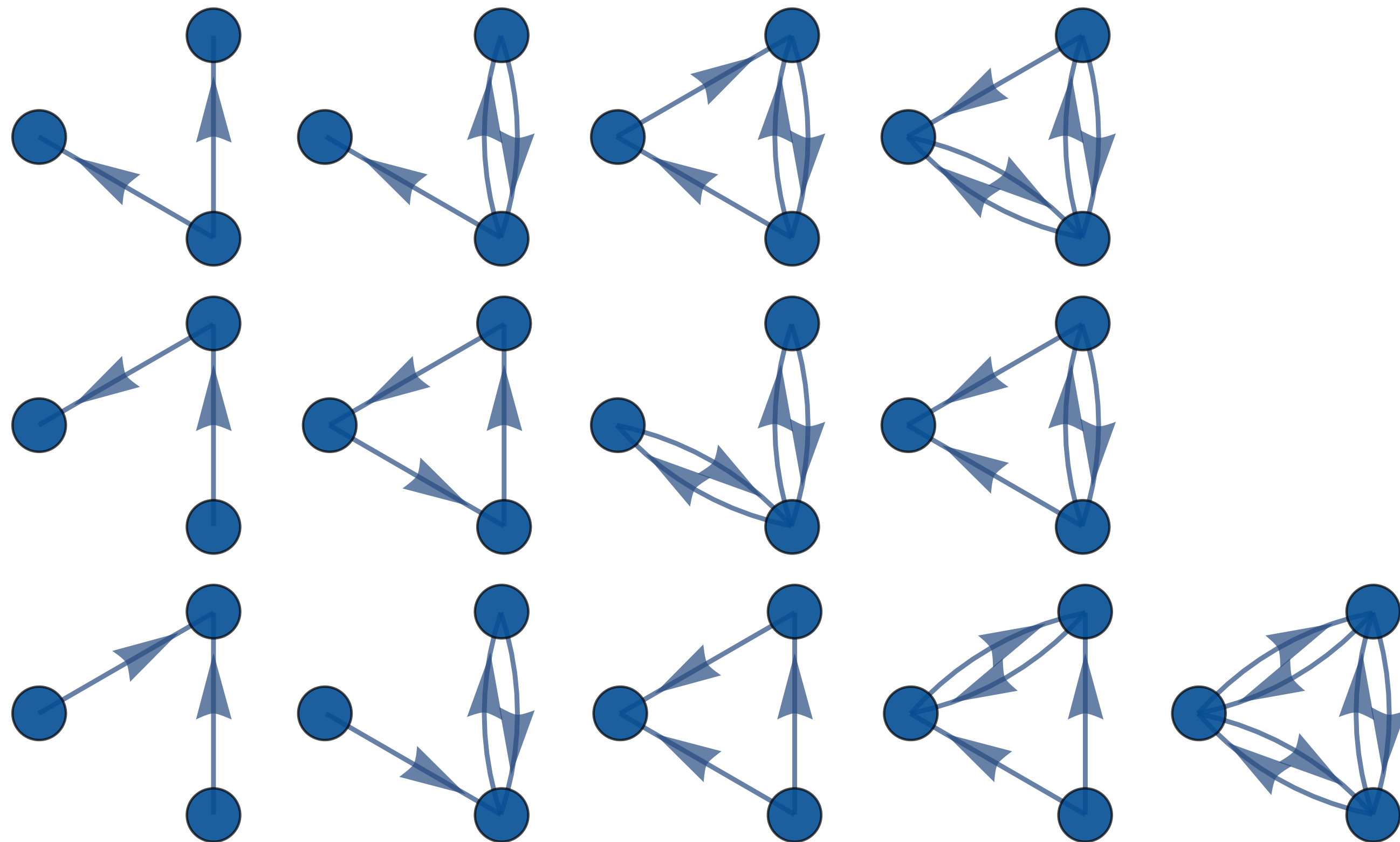
One billion years ago, in early multicellular animal life forms

Components of 15–20 cell–cell signaling pathways, cell adhesion processes, apical basal polarization of cells, junction formation, epithelium formation, specialization of cells toward physiological ends, some developmental processes of the single-celled egg to the adult

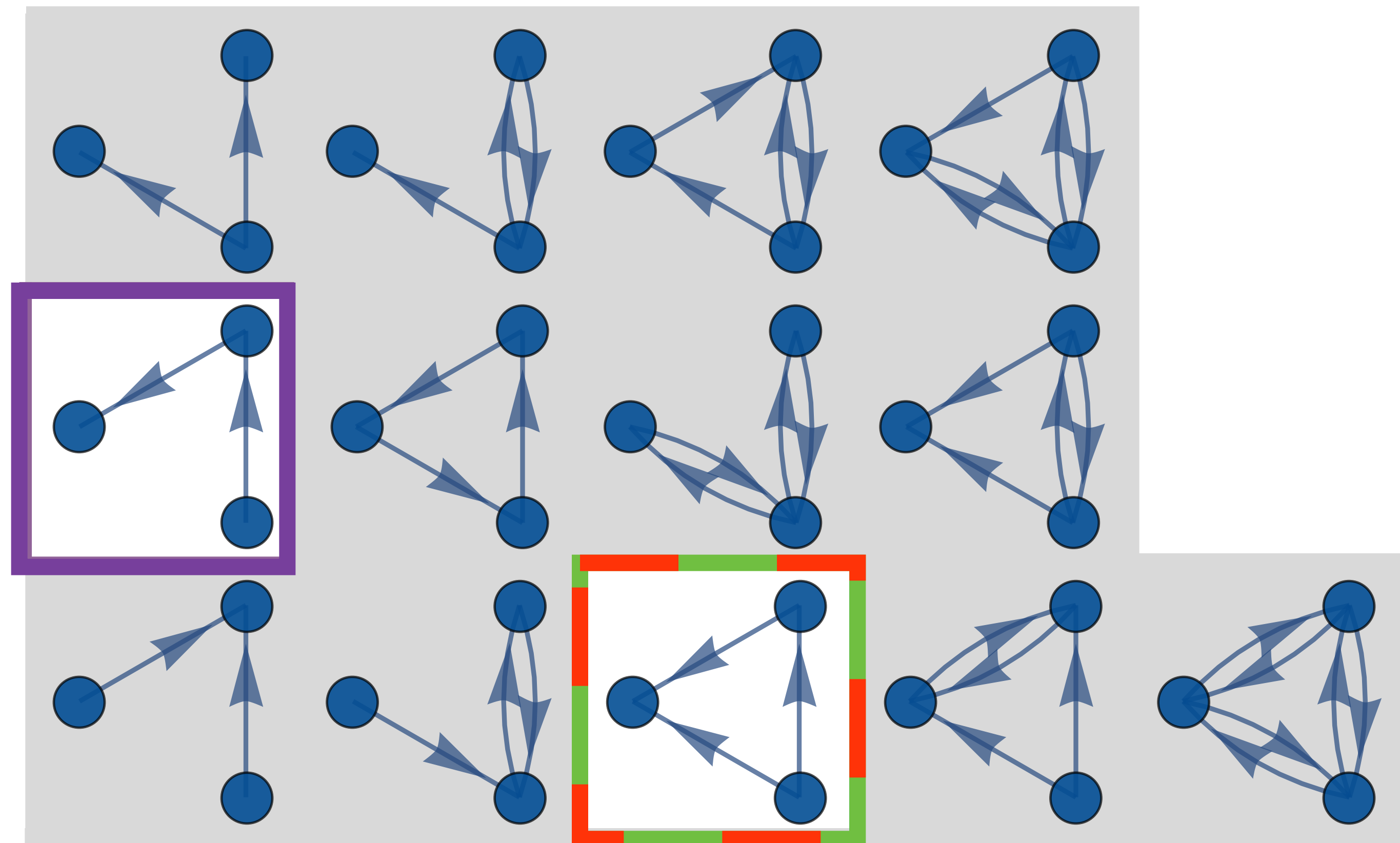
Near pre-Cambrian, in animals with early body axes

Components of complex developmental patterning, such as anteroposterior axis formation (Wnt/Wnt antagonist gradients) and dorsoventral axis formation (Bmp/antagonist gradients), inductions, complex cell competence, additional specialized cell types, formation of the body plan's map of selector gene compartments (both transcription factors and signaling proteins), various regulatory processes




The universe of 3-node interconnection patterns



The universe of 3-node interconnection patterns



NETWORK MOTIFS

-  **Gene Transcription**
(Yeast and E. Coli)
-  **Neural network**
(C. Elegans)
-  **Food webs**
(Little Rock, Ythan...)

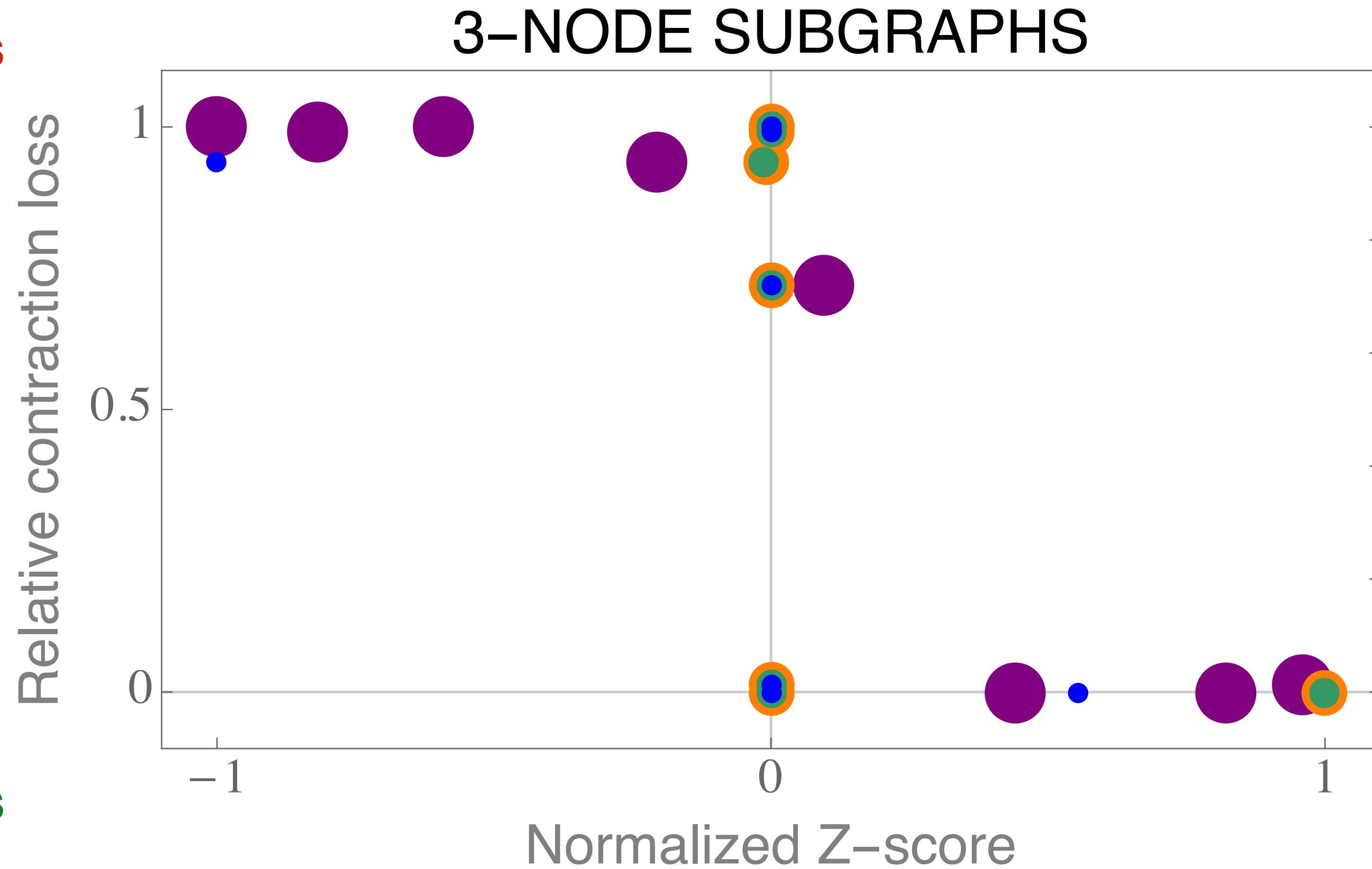
[Milo et.al. Science 2002]

Overrepresented subgraphs have low relative contraction loss

High contraction loss



Low contraction loss



- C.Elegans
- Yeast
- E.Coli
- Food-web

Anti-motif



Motif

Time delays

Two contracting systems of possibly different dimensions, with identity metric,

$$\begin{cases} \dot{\mathbf{x}}_1 = \mathbf{f}_1(\mathbf{x}_1, t) \\ \dot{\mathbf{x}}_2 = \mathbf{f}_2(\mathbf{x}_2, t) \end{cases}$$

Delayed coupling with constant \mathbf{G}_i and constants $k_i \geq 0$

$$\begin{cases} \dot{\mathbf{x}}_1 = \mathbf{f}_1(\mathbf{x}_1, t) + k_1 \mathbf{G}_2 (\mathbf{G}_1^T \mathbf{x}_2(t - T_{21}) - \mathbf{G}_2^T \mathbf{x}_1(t)) \\ \dot{\mathbf{x}}_2 = \mathbf{f}_2(\mathbf{x}_2, t) + k_2 \mathbf{G}_1 (\mathbf{G}_2^T \mathbf{x}_1(t - T_{12}) - \mathbf{G}_1^T \mathbf{x}_2(t)) \end{cases}$$

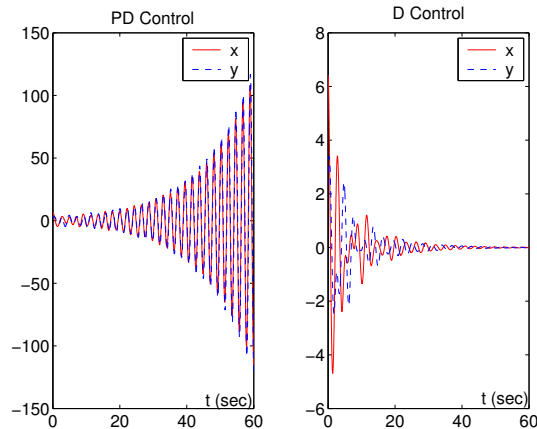
preserves asymptotic contraction.

For same dimension and $\mathbf{G}_1 = \mathbf{G}_2$, means couplings are p.s.d. in [same metric](#).

- Teleoperation between two linear mass-spring dampers

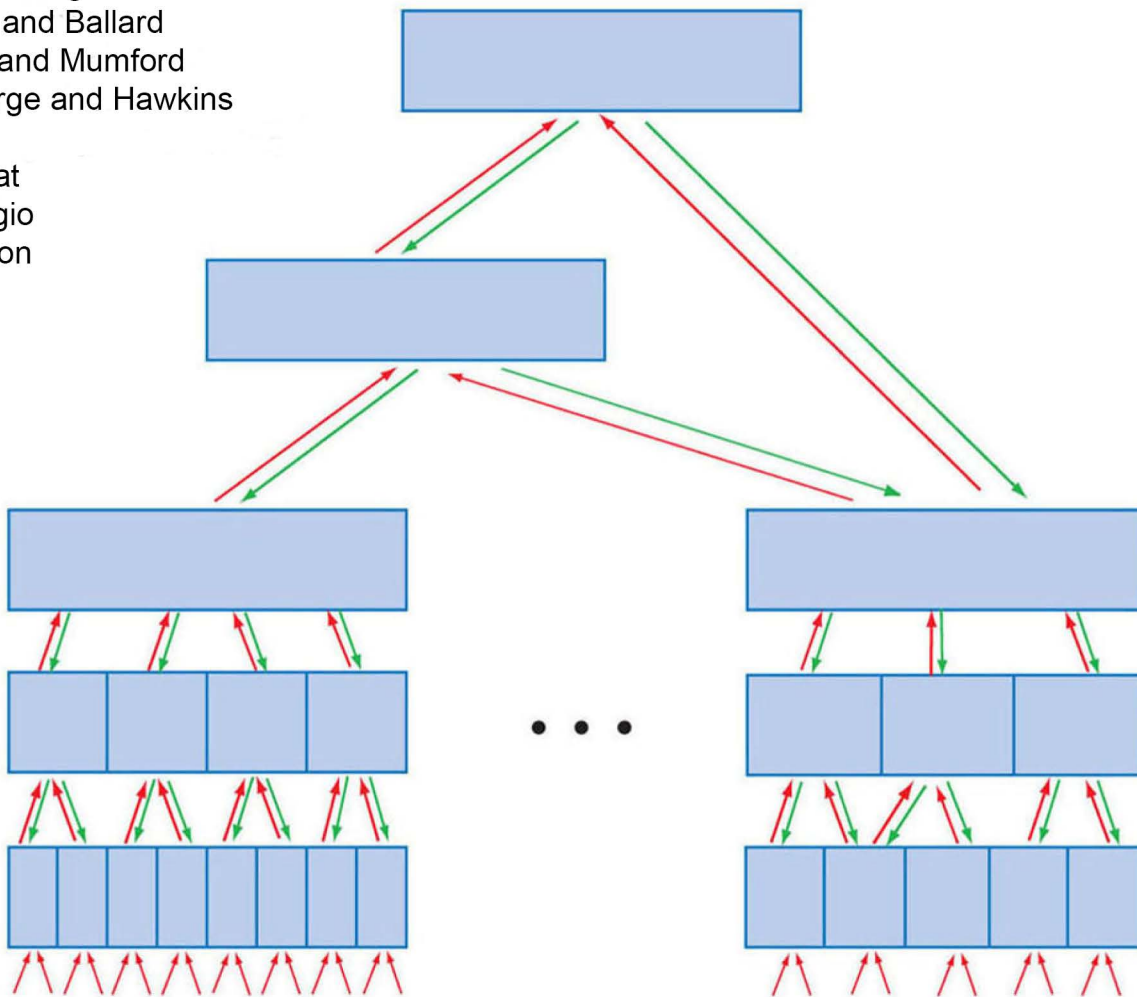
$$\begin{cases} \ddot{x}_1 + b\dot{x}_1 + \omega^2 x_1 = k_D(\dot{x}_2(t - T_{21}) - \dot{x}_1(t)) + k_P(x_2(t - T_{21}) - x_1(t)) \\ \ddot{x}_2 + b\dot{x}_2 + \omega^2 x_2 = k_D(\dot{x}_1(t - T_{12}) - \dot{x}_2(t)) + k_P(x_1(t - T_{12}) - x_2(t)) \end{cases}$$

$$\Theta = \begin{bmatrix} \omega & 0 \\ b & 1 \end{bmatrix} \quad \rightarrow \quad \mathbf{K} = \begin{bmatrix} k_D & 0 \\ \frac{k_P}{\omega} & 0 \end{bmatrix}$$



- Similarly, in Turing/Smale **diffusion-driven instability**, the diffusion gains lose positive semi-definiteness once expressed in the individual cell **metric**.

Luetzgen, et al.
Dayan, et al.
Grossberg
Rao and Ballard
Lee and Mumford
George and Hawkins
Rao
Mallat
Poggio
Friston
...



Predictive Feedback Hierarchies

Toward Contraction Analysis of Optimization

With metric $\mathbf{M} = \mathbf{\Theta}^\top \mathbf{\Theta}$ the condition

$$\left(\dot{\mathbf{\Theta}} + \mathbf{\Theta} \mathbf{A} \right) \mathbf{\Theta}^{-1} \preceq -\alpha \mathbf{I}$$

is equivalent to

$$\dot{\mathbf{M}} + \mathbf{A}^\top \mathbf{M} + \mathbf{M} \mathbf{A} \preceq -2\alpha \mathbf{M}$$

Recall that $\dot{\mathbf{M}} + \mathbf{A}^\top \mathbf{M} + \mathbf{M} \mathbf{A} \preceq -2\alpha \mathbf{M}$ implies that geodesic distances decrease exponentially,

$$d_{\mathcal{M}}(\mathbf{x}_1(t), \mathbf{x}_2(t)) \leq e^{-\alpha t} d_{\mathcal{M}}(\mathbf{x}_1(0), \mathbf{x}_2(0))$$

Contraction analysis of gradient flows

Fact: If a function $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ is α -strongly convex, then its gradient system

$$\dot{\mathbf{x}} = -\partial_{\mathbf{x}} f$$

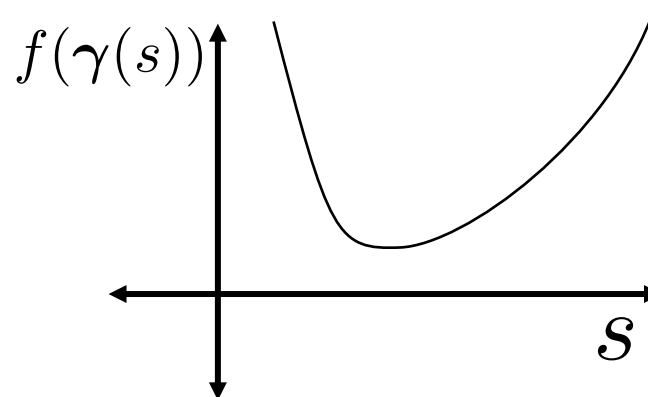
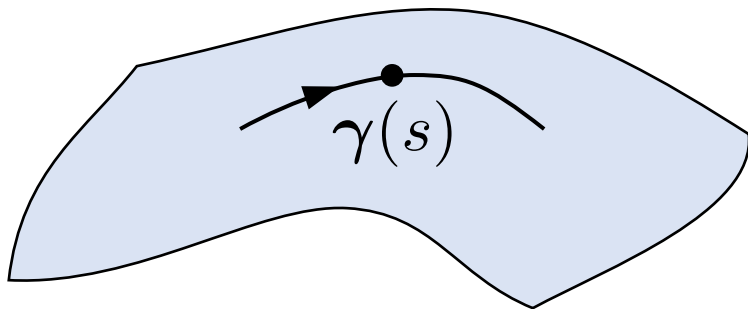
converges to the global minimum exponentially with rate α .

Proof: Consider the identity metric $\mathbf{M}(\mathbf{x}) = \mathbf{I}$.

$$\mathbf{M}\mathbf{A} + \mathbf{A}^\top \mathbf{M} + \dot{\mathbf{M}} = -2\partial_{\mathbf{xx}} f \leq -2\alpha \mathbf{I}$$

Generalization to the Riemannian Setting

- The *natural gradient* (Amari, 1998) gives the direction of steepest ascent according to distances measured on a space or manifold equipped with a Riemannian metric.
- **Natural gradient descent:** $\dot{\mathbf{x}} = -\mathbf{M}(\mathbf{x})^{-1} \partial_{\mathbf{x}} f$
- A function $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ is α -strongly *g-convex*, if it is α -strongly convex in the length parameter along any geodesic curve.
 - Equivalently, its geodesic Hessian $\mathbf{H}(\mathbf{x})$ must satisfy $\mathbf{H}(\mathbf{x}) \succeq \alpha \mathbf{M}(\mathbf{x})$.



Geodesic Hessian

$$H_{ij} = \partial_{ij} f - \Gamma_{ij}^k \partial_k f$$

Contraction Analysis of Natural Gradient Flows

Theorem: A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is α -strongly g-convex in a metric $\mathbf{M}(\mathbf{x})$ if and only if its natural gradient system is contracting at rate α in the same metric.

Basic Insight: The geodesic Hessian $\mathbf{H}(\mathbf{x})$ is given by

$$\mathbf{H}(\mathbf{x}) = -\frac{1}{2} \left(\mathbf{M}\mathbf{A} + \mathbf{A}^\top \mathbf{M} + \dot{\mathbf{M}} \right)$$

where $\mathbf{A}(\mathbf{x})$ is the Jacobian of the natural gradient dynamics.

Non-Strict Contraction Case

Consider the case with the natural gradient dynamics semi-contracting:

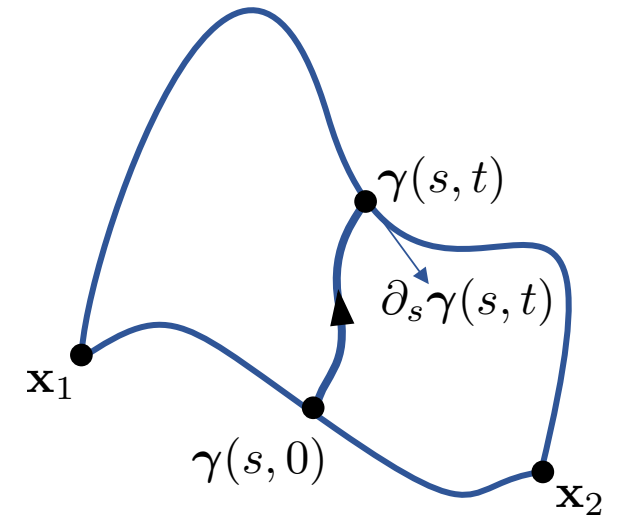
$$\frac{d}{dt} [\delta \mathbf{x}^\top \mathbf{M} \delta \mathbf{x}] = -2\delta \mathbf{x}^\top \mathbf{H}(\mathbf{x})\delta \mathbf{x} \preceq \mathbf{0}$$

Under mild assumptions, along trajectories $\mathbf{H}(\mathbf{x}(t))\delta \mathbf{x}(t) \rightarrow 0$

Theorem: If the natural gradient dynamics of f are semi-contracting in a metric $\mathbf{M}(\mathbf{x})$, then f is g-convex, every stationary point of f is a global optimum, and any geodesic between optima is composed of optima.

Sketch: Consider the family of solutions $\gamma(s, t)$ with initial conditions $\gamma(0, 0) = \mathbf{x}_1$, $\gamma(1, 0) = \mathbf{x}_2$ with \mathbf{x}_1 and \mathbf{x}_2 stationary points.

1. $\mathbf{H}(\gamma(s, t))\partial_s \gamma(s, t) \rightarrow 0$ as $t \rightarrow \infty$.
2. $\partial_{\mathbf{x}} f(\gamma(s, t)) \rightarrow 0$ as $t \rightarrow \infty$
3. $f(\mathbf{x}_2) = f(\mathbf{x}_1)$

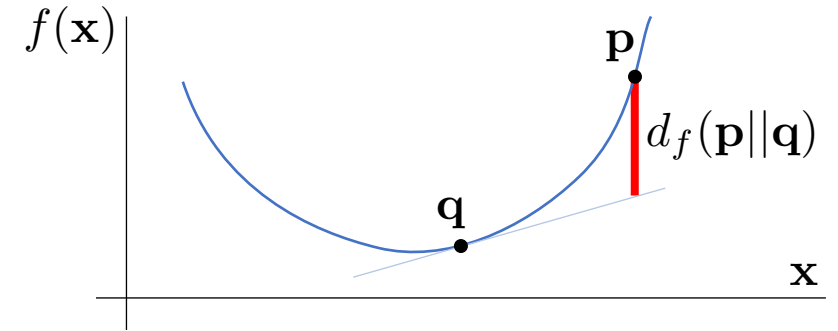


Examples: Bregman Divergence

Recall the Bregman divergence of a convex function f :

$$d_f(\mathbf{p}||\mathbf{q}) = f(\mathbf{p}) - f(\mathbf{q}) - \langle \partial_{\mathbf{x}}f(\mathbf{q}), \mathbf{p} - \mathbf{q} \rangle$$

The Bregman divergence is convex in \mathbf{p} but not necessarily in \mathbf{q} .



Example: Generalized KL-divergence $\mathbf{x} \in \mathbb{R}^n$,
 $f(\mathbf{x}) = \sum x_i \log(x_i)$.

$$d_f(\mathbf{p}||\mathbf{q}) = \sum p_i \log\left(\frac{p_i}{q_i}\right) + q_i - p_i$$

- Metric via: $ds^2 = \sum \left(\frac{\delta x_i^2}{x_i}\right)$
- Natural gradient for fixed \mathbf{p} , $\dot{\mathbf{q}} = \mathbf{p} - \mathbf{q}$.
- Differential dynamics $\frac{d}{dt} \delta \mathbf{q} = -\delta \mathbf{q} \Rightarrow$ **contracting**
- **The discrete KL-divergence is g-convex in \mathbf{q} .**

Example: For PD matrices

$$f(\mathbf{X}) = -\log \det(\mathbf{X})$$

$d_f(\mathbf{P}||\mathbf{Q})$ the KL-divergence between $\mathcal{N}(\mathbf{0}, \mathbf{P})$ and $\mathcal{N}(\mathbf{0}, \mathbf{Q})$.

- Metric via: $ds^2 = \text{tr}((\mathbf{X}^{-1} \delta \mathbf{X})^2)$
- Natural gradient for fixed \mathbf{P} , $\dot{\mathbf{Q}} = -(\mathbf{Q} - \mathbf{P})$
- Differential dynamics: $\frac{d}{dt} \delta \mathbf{Q} = -\delta \mathbf{Q} \Rightarrow$ **contracting**.
- Thus, $d_f(\mathbf{P}||\mathbf{Q})$ is g-convex in \mathbf{Q} .

Extension to Time-Varying Contexts

Theorem: A function $f(\mathbf{x}, t)$ is α -strongly g-convex in a metric $\mathbf{M}(\mathbf{x})$ for each t if and only if its natural gradient system is contracting at rate α in the same metric.

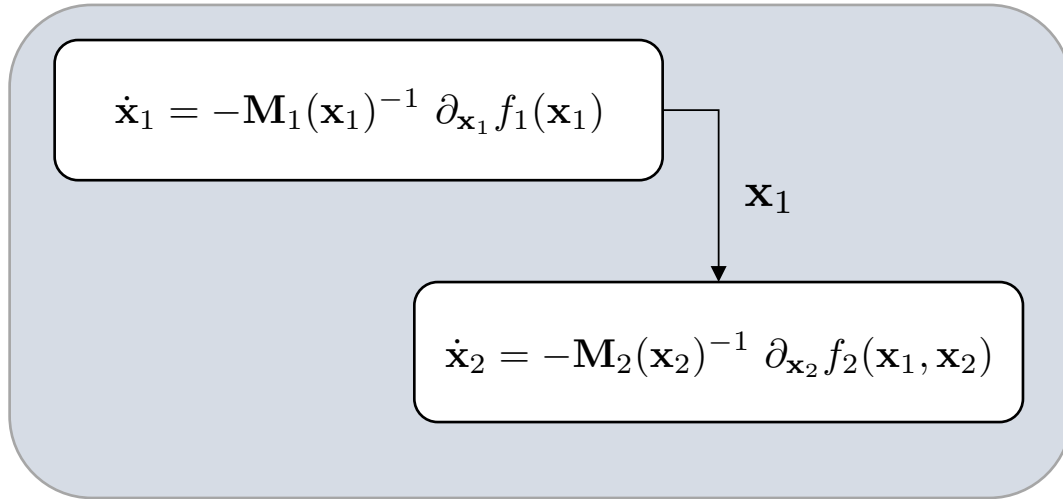
Basic Insight: The geodesic Hessian at fixed time $\mathbf{H}(\mathbf{x}, t)$ is given by

$$\mathbf{H}(\mathbf{x}, t) = -\frac{1}{2} \left(\mathbf{M}(\mathbf{x}) \mathbf{A}(\mathbf{x}, t) + \mathbf{A}(\mathbf{x}, t)^\top \mathbf{M}(\mathbf{x}) + \dot{\mathbf{M}}(\mathbf{x}, t) \right)$$

where $\mathbf{A}(\mathbf{x}, t)$ is the Jacobian of the natural gradient dynamics at time t .

Combination Properties

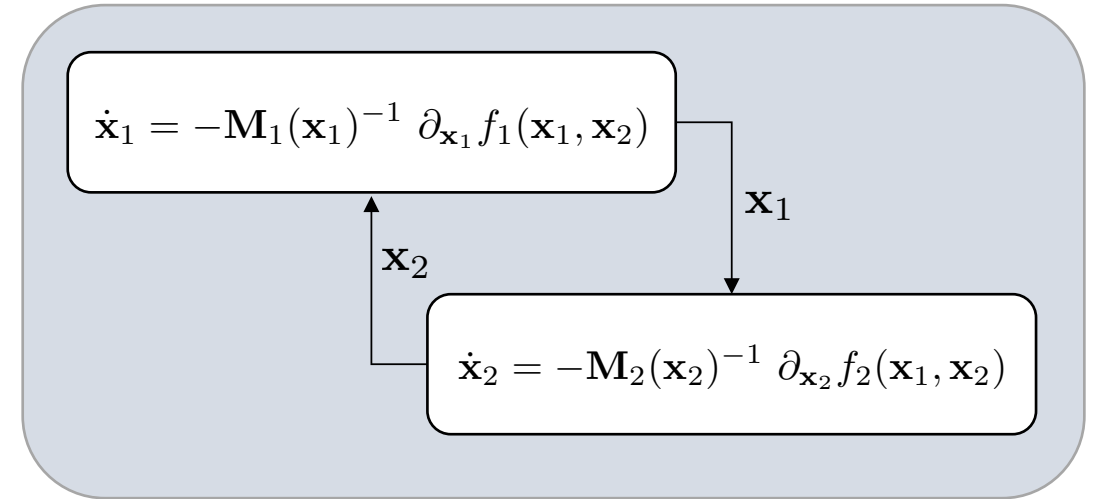
Hierarchical Natural Gradient



Contracting when:

- f_1 strongly g-convex
- f_2 strongly g-convex for each \mathbf{x}_1

Multi-player games



Contracting to a unique Nash when:

- f_1 strongly g-convex for each \mathbf{x}_2
- f_2 strongly g-convex for each \mathbf{x}_1
- $\partial_{\mathbf{x}_1 \mathbf{x}_2} f_1 = -k (\partial_{\mathbf{x}_2 \mathbf{x}_1} f_2)^\top$

Extension to the Primal Dual Setting

Consider the Lagrangian $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, t)$ with metric $\mathbf{M}_{\mathbf{x}}$ and $\mathbf{M}_{\boldsymbol{\lambda}}$. The associated **geodesic primal-dual dynamics**

$$\begin{aligned}\mathbf{M}_{\mathbf{x}}(\mathbf{x}) \dot{\mathbf{x}} &= -\partial_{\mathbf{x}}\mathcal{L} \\ \mathbf{M}_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) \dot{\boldsymbol{\lambda}} &= \partial_{\boldsymbol{\lambda}}\mathcal{L}\end{aligned}$$

are contracting if \mathcal{L} g-convex over \mathbf{x} (in $\mathbf{M}_{\mathbf{x}}$) and g-concave over $\boldsymbol{\lambda}$ (in $\mathbf{M}_{\boldsymbol{\lambda}}$).

Contraction in metric $\mathbf{M} = \text{diag}(\mathbf{M}_{\mathbf{x}}, \mathbf{M}_{\boldsymbol{\lambda}})$:

$$\mathbf{M}^{\top} \mathbf{A} + \mathbf{A}^{\top} \mathbf{M} + \dot{\mathbf{M}} = -2 \begin{bmatrix} \mathbf{H}_{\mathbf{x}} & \mathbf{0} \\ \mathbf{0} & -\mathbf{H}_{\boldsymbol{\lambda}} \end{bmatrix} < \mathbf{0}$$

Potential applications:

- Min-max problems in adversarial training
- Distributed constrained g-convex optimization

Partial Contraction

Wang and Slotine, 2002

Two nonlinear systems synchronize if their trajectories are both particular solutions of a **virtual** contracting system.

Non-Autonomous Setting: Virtual Systems

Example: Time-varying learning rates in natural gradient descent

$$\dot{\mathbf{x}} = -p(\mathbf{x}, t) \mathbf{M}(\mathbf{x})^{-1} \partial_{\mathbf{x}} f(\mathbf{x})$$

If $p(\mathbf{x}, t) > p_{min}$, and f is α -strongly g -convex in $\mathbf{M}(\mathbf{x})$, then the above converges exponentially to the minimum of f with rate αp_{min} .

Virtual System:

$$\dot{\mathbf{y}} = -p(\mathbf{x}, t) \mathbf{M}^{-1}(\mathbf{y}) \partial_{\mathbf{y}} f(\mathbf{y})$$

Quorum sensing

All-to-all coupling

$$\dot{\mathbf{x}}_i = \mathbf{f}(\mathbf{x}_i, t) + k \sum_j (\mathbf{x}_j - \mathbf{x}_i) \quad i = 1, \dots, N$$

equivalent to **local damping** + **quorum sensing**

$$\dot{\mathbf{x}}_i = \mathbf{f}(\mathbf{x}_i, t) - kN\mathbf{x}_i + k \sum_j \mathbf{x}_j \quad i = 1, \dots, N$$

Use virtual system

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, t) - kN\mathbf{y} + k \sum_j \mathbf{x}_j$$

contracting for upper-bounded Jacobian and kN large enough.

Note $N^2 \rightarrow 2N$ connections.

Some non-intuitive properties

- **Global from local**

- **Leader following**

Global synchronization to single, locally-connected element.

- **Fast inhibition**

A single inhibitory link can turn off an entire network.

- **Robustness**

Spike synchrony preserved under large parameter variations.

- **Adaptation leader**

A single element can control entire qualitative behavior.

- **Long range synchronization through different dynamics**

- **Synchronization protects from noise**

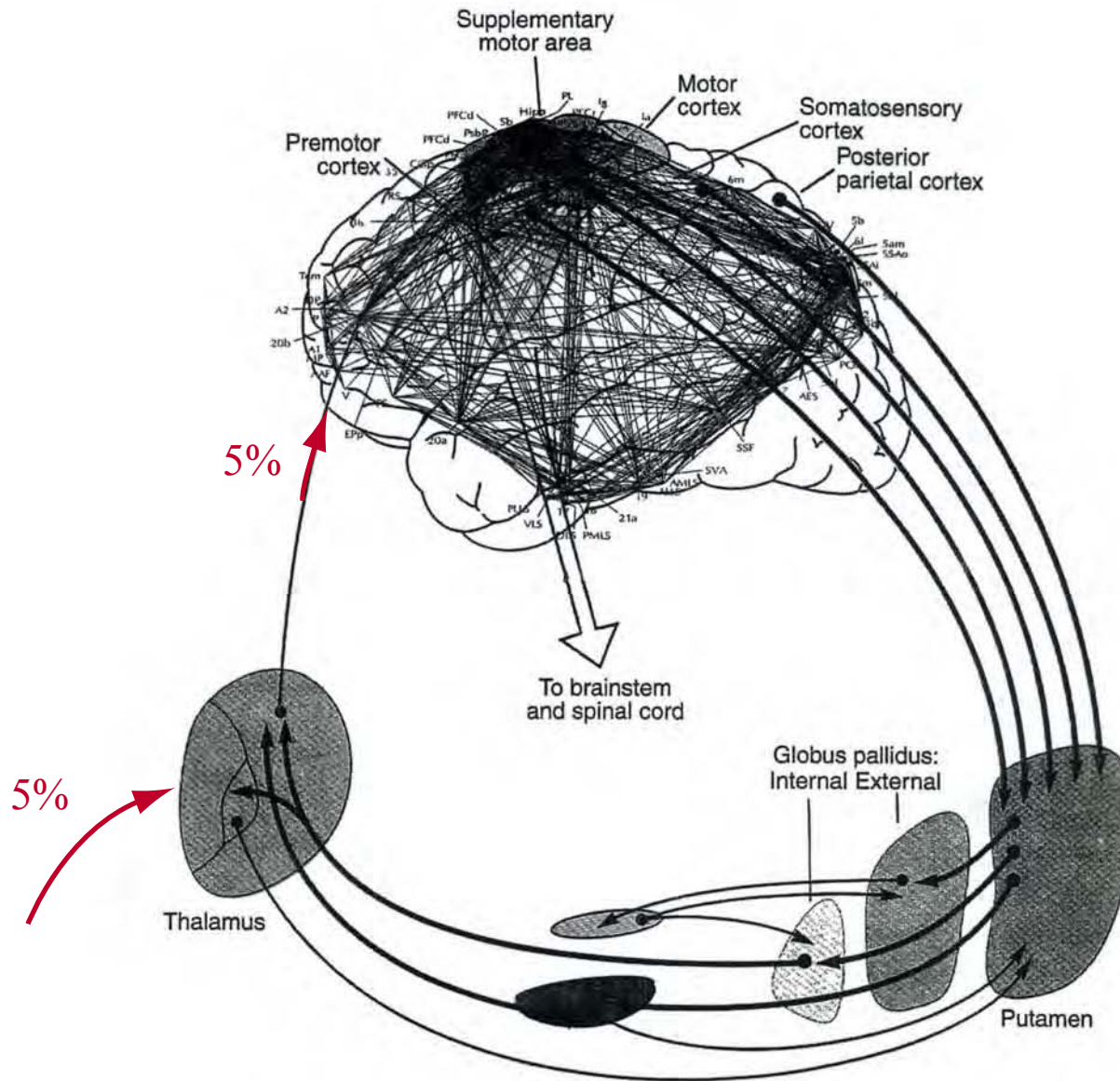


FIGURE 14.1 STRUCTURES AND CONNECTIONS MEDIATING CONSCIOUS AND UNCONSCIOUS PROCESSES.

Edelman and Tononi, 2000

Fast Inhibition

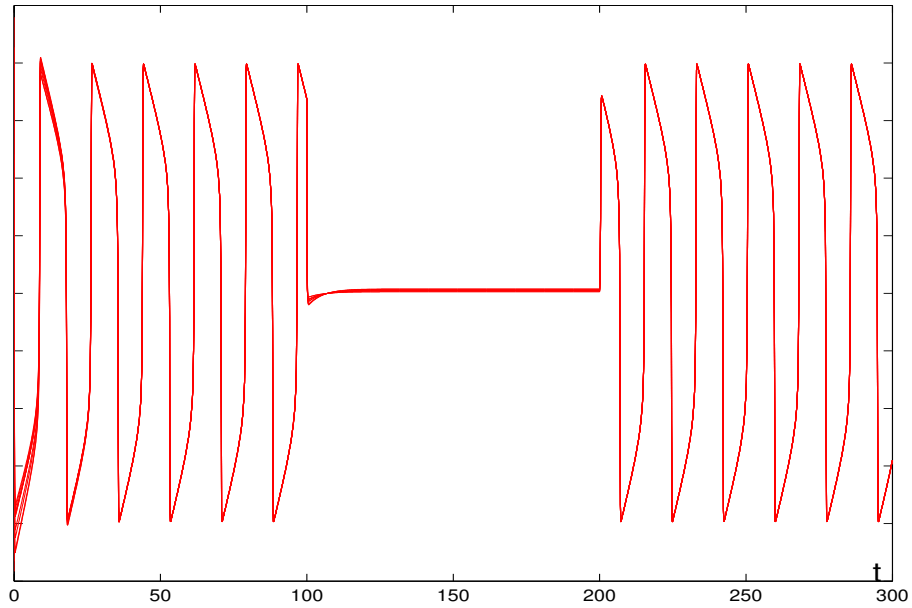
$$\dot{\mathbf{x}}_i = \mathbf{f}(\mathbf{x}_i, t) + \sum_{j \in \mathcal{N}_i} \mathbf{K}_{ji} (\mathbf{x}_j - \mathbf{x}_i) \quad i = 1, \dots, n$$

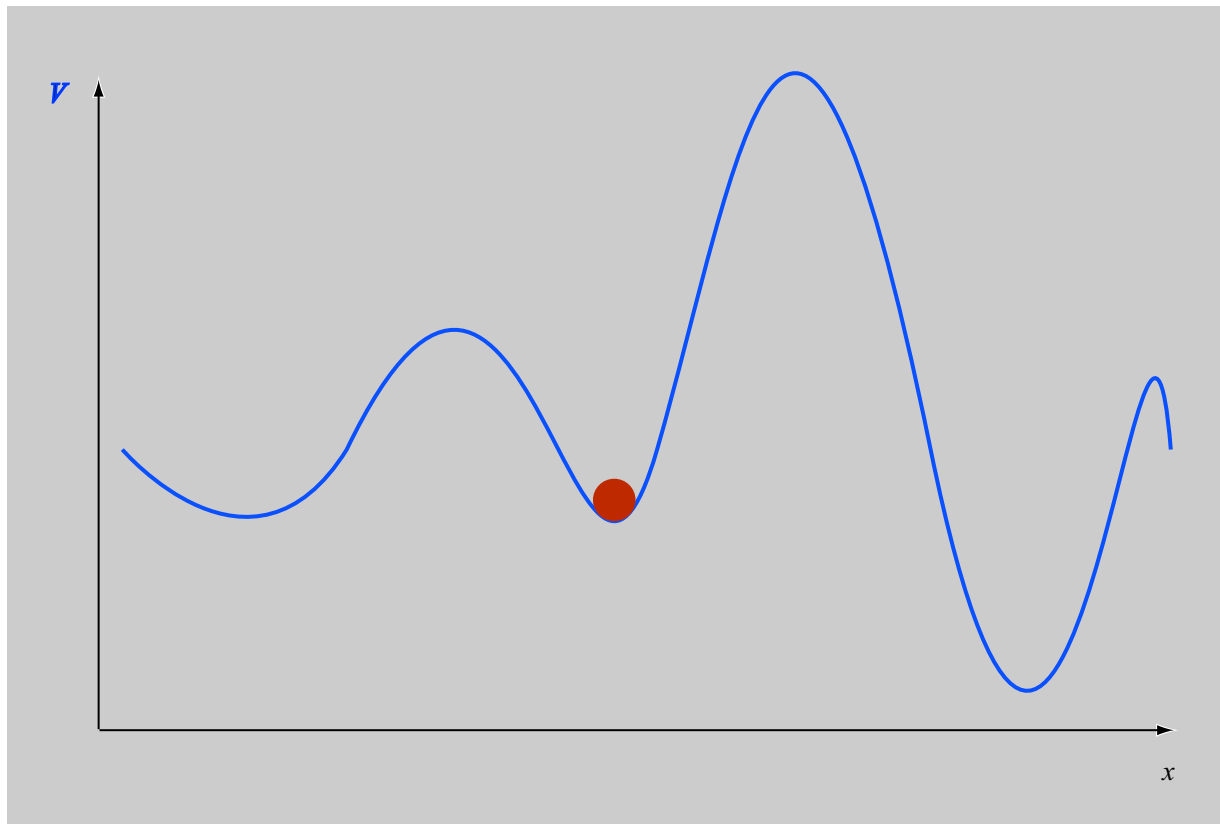
A **single inhibitory link** between two arbitrary elements has the ability to turn off the entire network.

$$\dot{\mathbf{x}}_a = \mathbf{f}(\mathbf{x}_a, t) + \sum_{j \in \mathcal{N}_a} \mathbf{K}_{ja} (\mathbf{x}_j - \mathbf{x}_a) + \mathbf{K} (-\mathbf{x}_b - \mathbf{x}_a)$$

$$\dot{\mathbf{x}}_b = \mathbf{f}(\mathbf{x}_b, t) + \sum_{j \in \mathcal{N}_b} \mathbf{K}_{jb} (\mathbf{x}_j - \mathbf{x}_b) + \mathbf{K} (-\mathbf{x}_a - \mathbf{x}_b)$$

Fast Inhibition





$$\dot{\mathbf{x}} = -\text{grad } V(\mathbf{x})$$

Polyrhythms

71 *Largo*

Fl. picc.
1
2
3

Fl. gr.
3

Fl. alto

Ob. 1, 2
3, 4

C. ing.

Cl. picc.
in Re

Cl. in Sib 1
2

Cl. bas.
in Sib 1
2

Fag.
1
3

C. Fag.
1
2
3

Cor. in Fa
1, 3
2, 4
3
6

Tr. picc.
in Re
1
2

Tr. in Do 3
4

Trbn.
1
2
3

Ten. in Sib 1
Tbn.
bas. 1
2
3

Timp.
G. C.
Tam-t.

Guero

Vl. I div.
1
2

Vl. II

Vla.

Vc.

Cb.

Largo

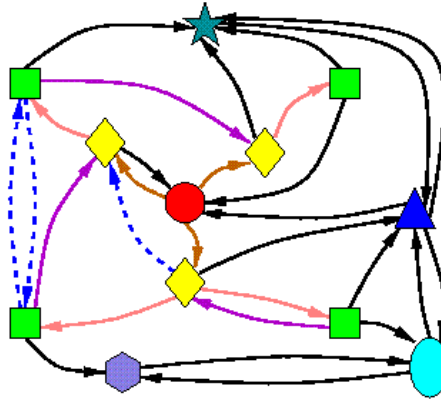
Tutti

Thb. 2 male
in Cor. 5

71 *Largo*

Concurrent Synchronization

Pham and Slotine, 2005



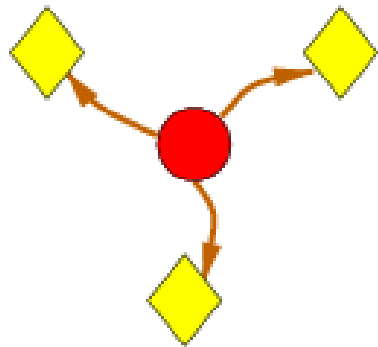
Under simple conditions on the coupling strengths, the ■ group globally exponentially synchronizes, thus providing synchronized inputs to the outer elements. So does the ◆ group.

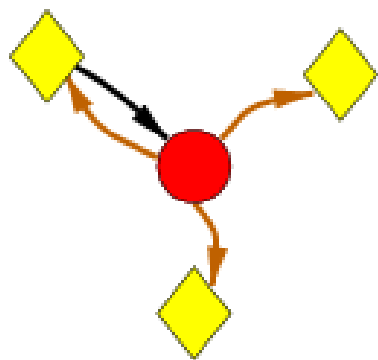
Regardless of the dynamics, connections, or inputs of the other systems.

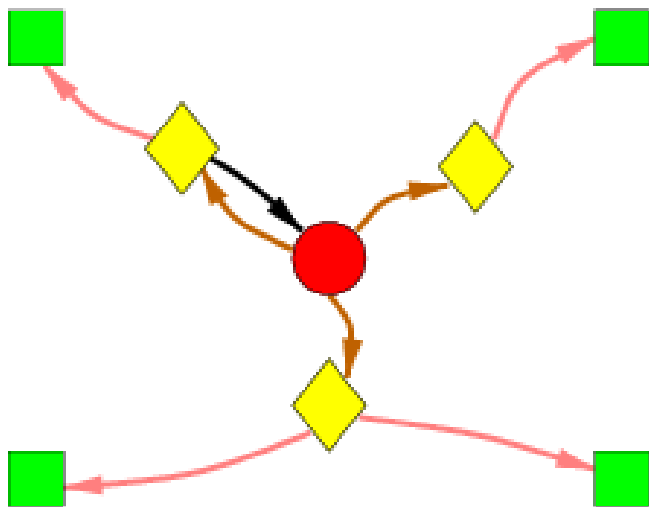
"As stable" as **global exponential convergence** to an equilibrium. But now to a possibly very **complex coordinated behavior**.

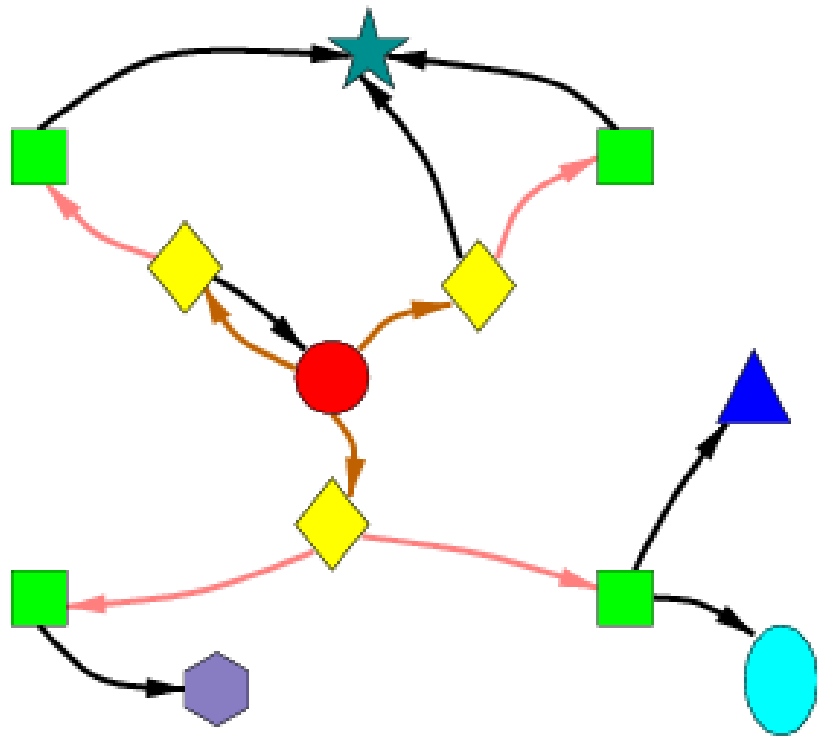
The **invariance** itself (but not the convergence) is closely related to the notion of **input-equivalence** (Golubitsky, et al.).

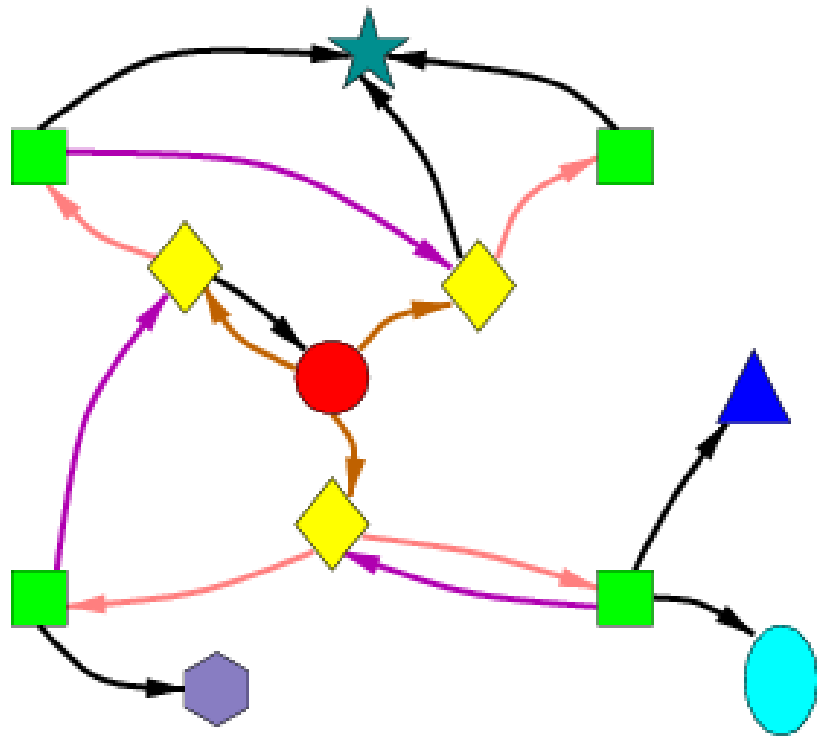
Evolution/Development friendly.

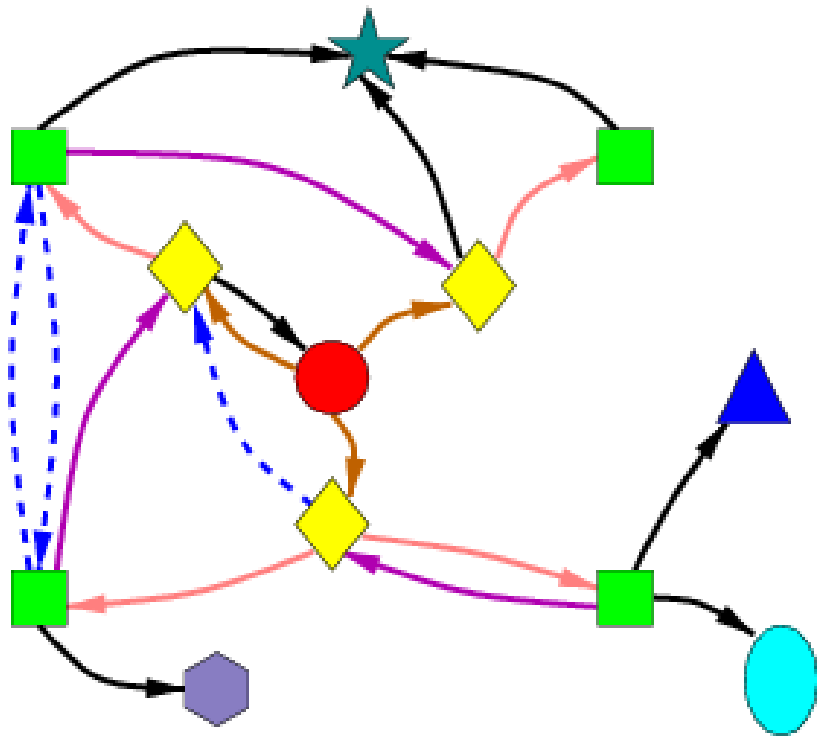


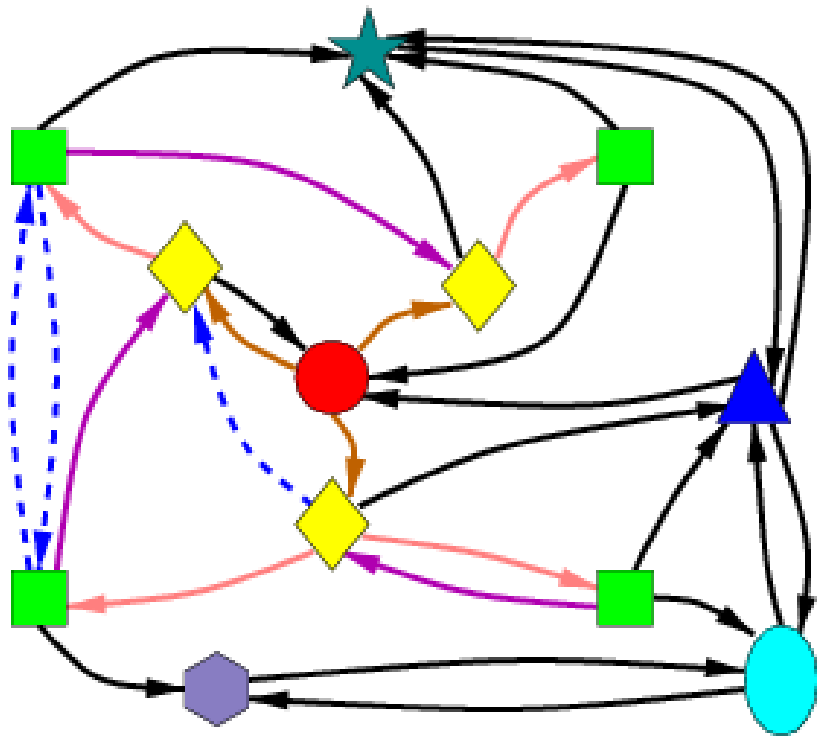


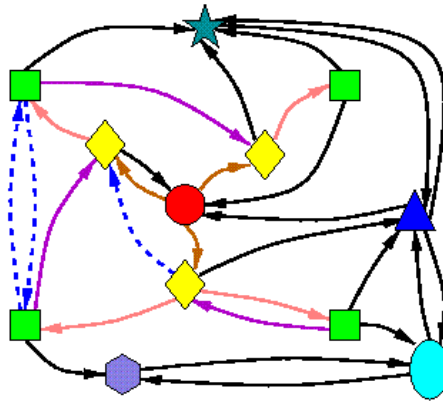








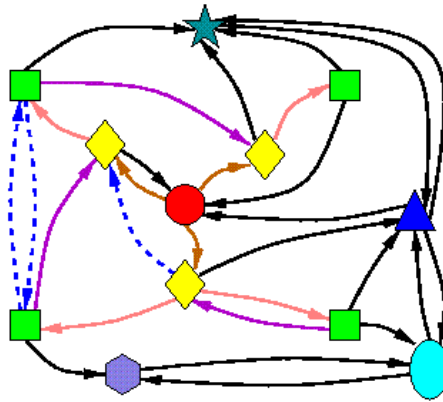




Global exponential concurrent synchronization

=

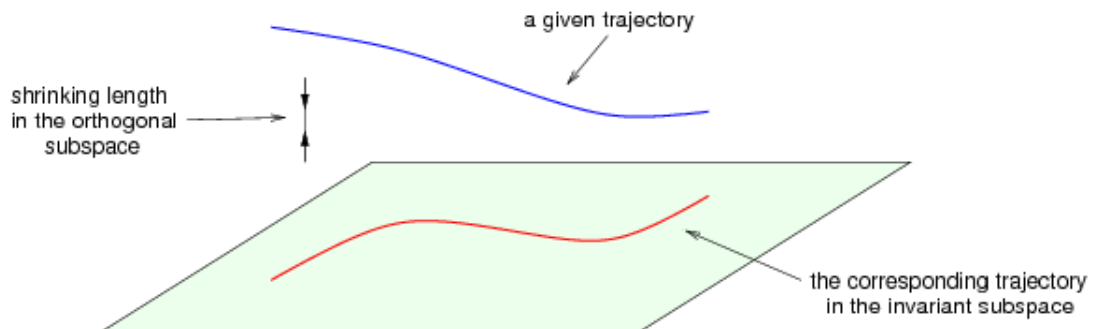
Contraction to a flow-invariant linear subspace

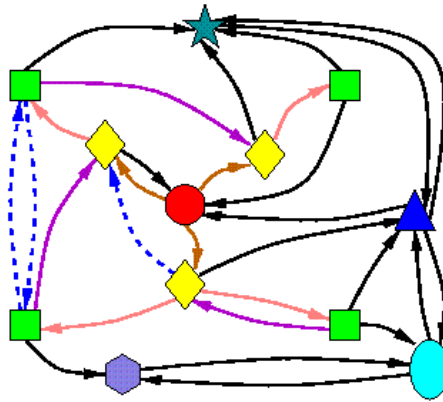


Global exponential concurrent synchronization

=

Contraction to a flow-invariant linear subspace





Global exponential concurrent synchronization
 =
 Contraction to a flow-invariant linear subspace

- Simple conditions based on Jacobians
- Combination properties

Contraction to a Linear Subspace

Theorem Consider a linear subspace \mathcal{M} invariant for

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$$

Let \mathbf{V} be the orthonormal projection on \mathcal{M}^\perp . All solutions converge exponentially to \mathcal{M} if dynamics

$$\dot{\mathbf{y}} = \mathbf{V}\mathbf{f}(\mathbf{V}^\top \mathbf{y}, t)$$

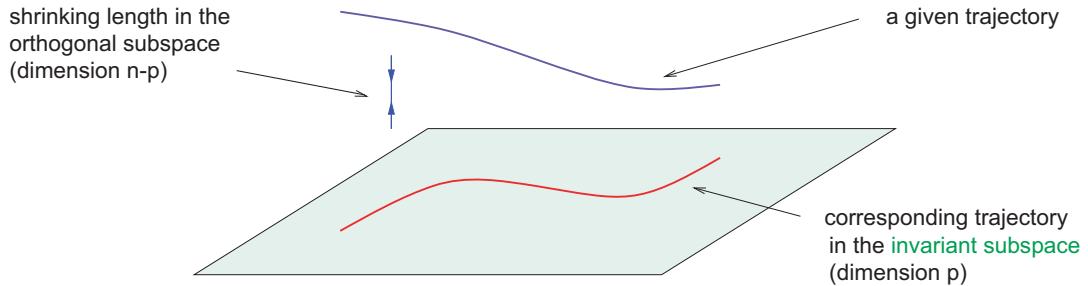
is contracting (in a constant metric).

Note

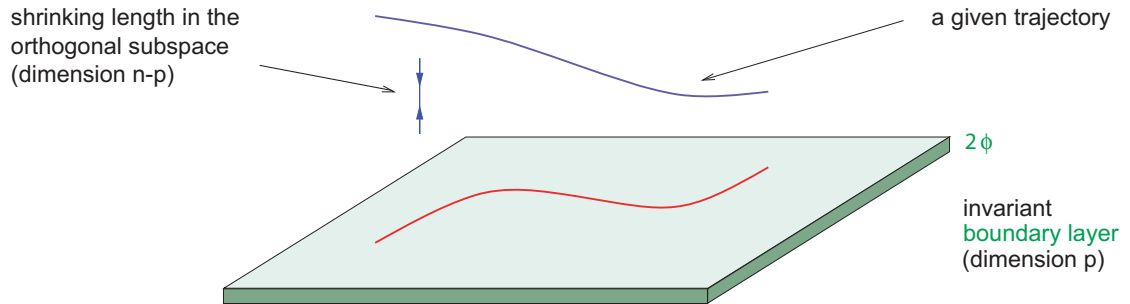
$$\mathbf{x} \in \mathcal{M} \iff \mathbf{V}\mathbf{x} = \mathbf{0}$$

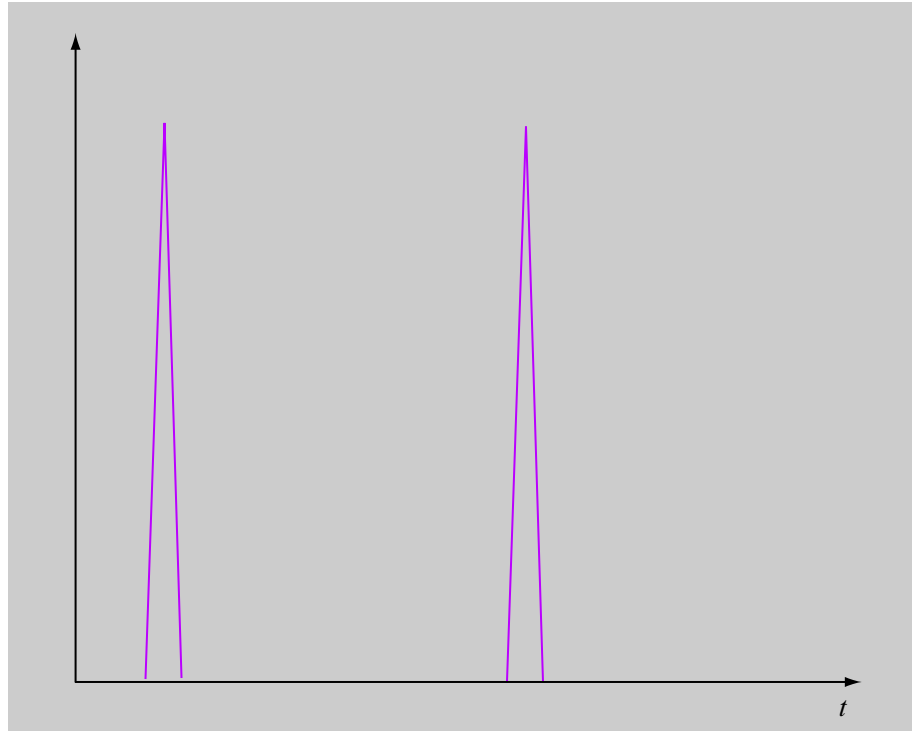
Synchronization rate = Contraction rate of \mathbf{y}

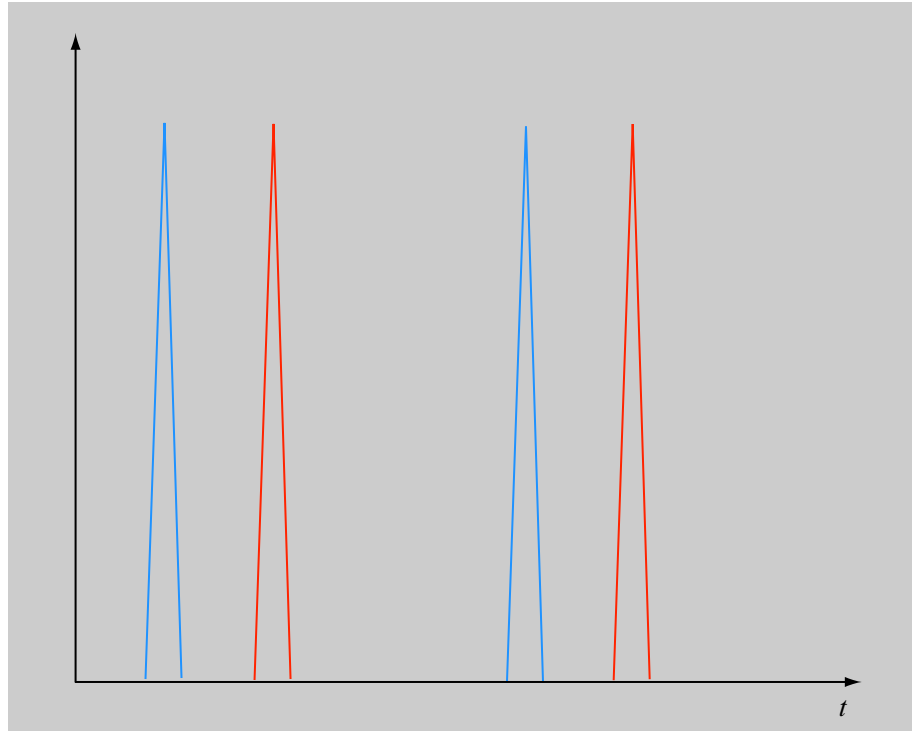
$$\dot{\mathbf{x}}_i = \mathbf{f}(\mathbf{x}_i, t) + \sum_{j \neq i} \mathbf{K}_{ij}(\mathbf{x}_j - \mathbf{x}_i) \quad i = 1, \dots, n$$



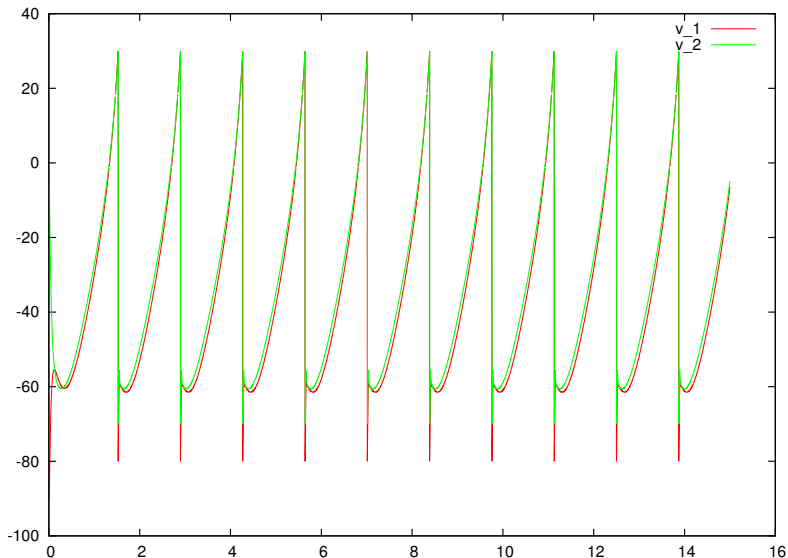
$$\dot{\mathbf{x}}_i = \mathbf{f}_i(\mathbf{x}_i, t) + \sum_{j \neq i} \mathbf{K}_{ij}(\mathbf{x}_j - \mathbf{x}_i) \quad i = 1, \dots, n$$



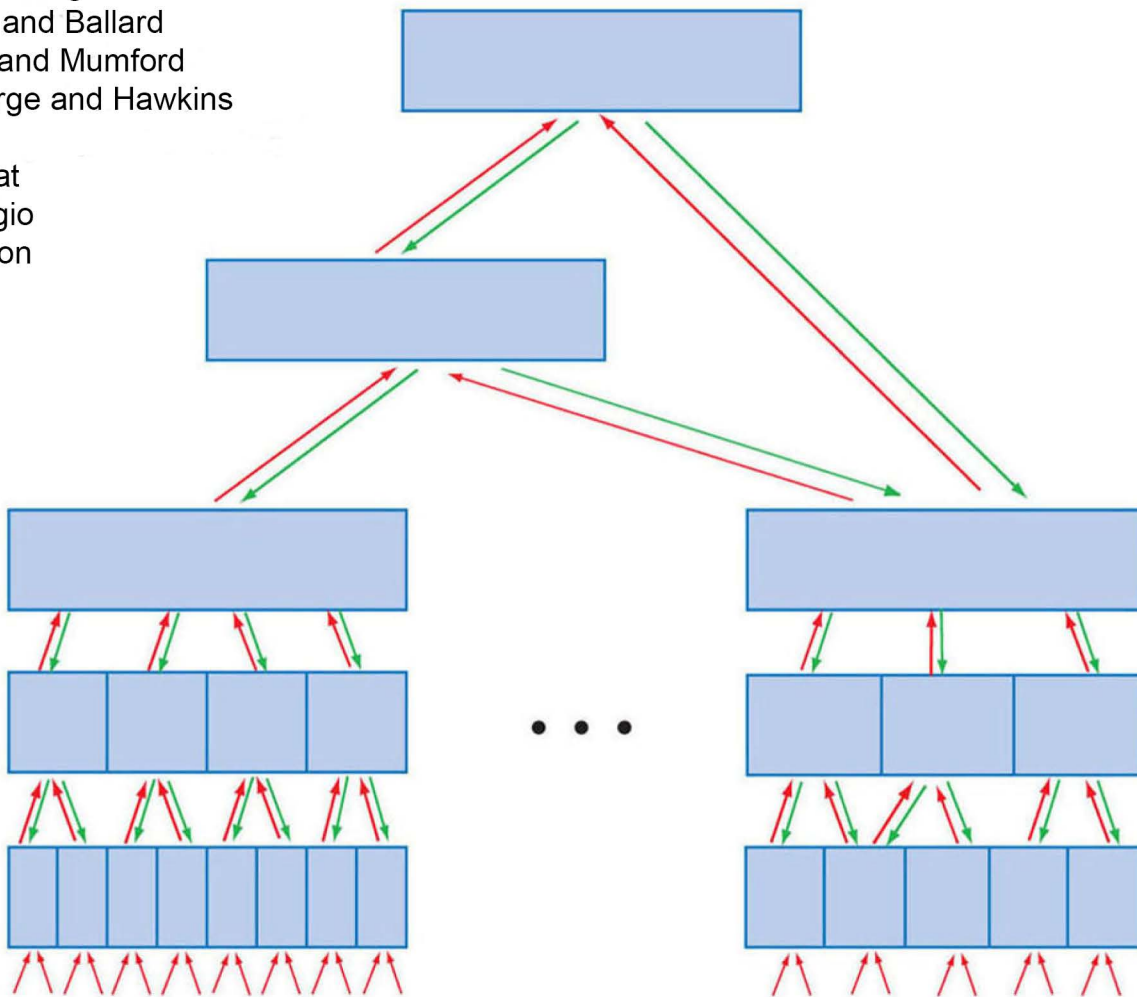




Two way coupling ($k_1 = k_2 = k$)



Luetzgen, et al.
Dayan, et al.
Grossberg
Rao and Ballard
Lee and Mumford
George and Hawkins
Rao
Mallat
Poggio
Friston
...



Predictive Feedback Hierarchies

Generalized diffusive connections

$$\begin{cases} \dot{\mathbf{x}}_1 = \mathbf{f}_1(\mathbf{x}_1, t) + k\mathbf{A}^\top(\mathbf{B}\mathbf{x}_2 - \mathbf{A}\mathbf{x}_1) \\ \dot{\mathbf{x}}_2 = \mathbf{f}_2(\mathbf{x}_2, t) + k\mathbf{B}^\top(\mathbf{A}\mathbf{x}_1 - \mathbf{B}\mathbf{x}_2) \end{cases}$$

where \mathbf{x}_1 and \mathbf{x}_2 can be of different dimensions, and \mathbf{A} and \mathbf{B} are constant matrices of appropriate dimensions.

$$\mathbf{J} = \begin{pmatrix} \frac{\partial \mathbf{f}_1}{\partial \mathbf{x}_1} & \\ & \frac{\partial \mathbf{f}_2}{\partial \mathbf{x}_2} \end{pmatrix} - k\mathbf{L}, \quad \text{where } \mathbf{L} = \begin{pmatrix} \mathbf{A}^\top\mathbf{A} & -\mathbf{A}^\top\mathbf{B} \\ -\mathbf{B}^\top\mathbf{A} & \mathbf{B}^\top\mathbf{B} \end{pmatrix}$$

$$\mathbf{L} = \mathbf{L}^\top \geq \mathbf{0} \text{ since } \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{pmatrix} \mathbf{L} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} = \|\mathbf{A}\mathbf{x}_1 - \mathbf{B}\mathbf{x}_2\|^2$$

Assume that the subspace $\mathcal{M} : \mathbf{A}\mathbf{x}_1 - \mathbf{B}\mathbf{x}_2 = \mathbf{0}$ is flow-invariant. Using the projection \mathbf{V} on \mathcal{M}^\perp (so $\mathbf{V}\mathbf{L}\mathbf{V}^\top > \mathbf{0}$), for upper bounded individual Jacobians, large enough k ensures exponential convergence to \mathcal{M} .

By recursion for larger systems.

How Synchronization Protects from Noise

Tabareau, et al., PLoS Comp. Bio., January 2010

Consider the synchronized systems

$$dx_i = \left[f(x_i) + k \sum_{j \neq i} (x_j - x_i) \right] dt + \sigma_i(x_i) dW_i$$

The average, $M = \frac{1}{N} \sum_i x_i$, satisfies

$$dM = \frac{1}{N} \left[\sum_i f(x_i) dt + \sum_i \sigma_i(x_i) dW_i \right]$$

This can be written

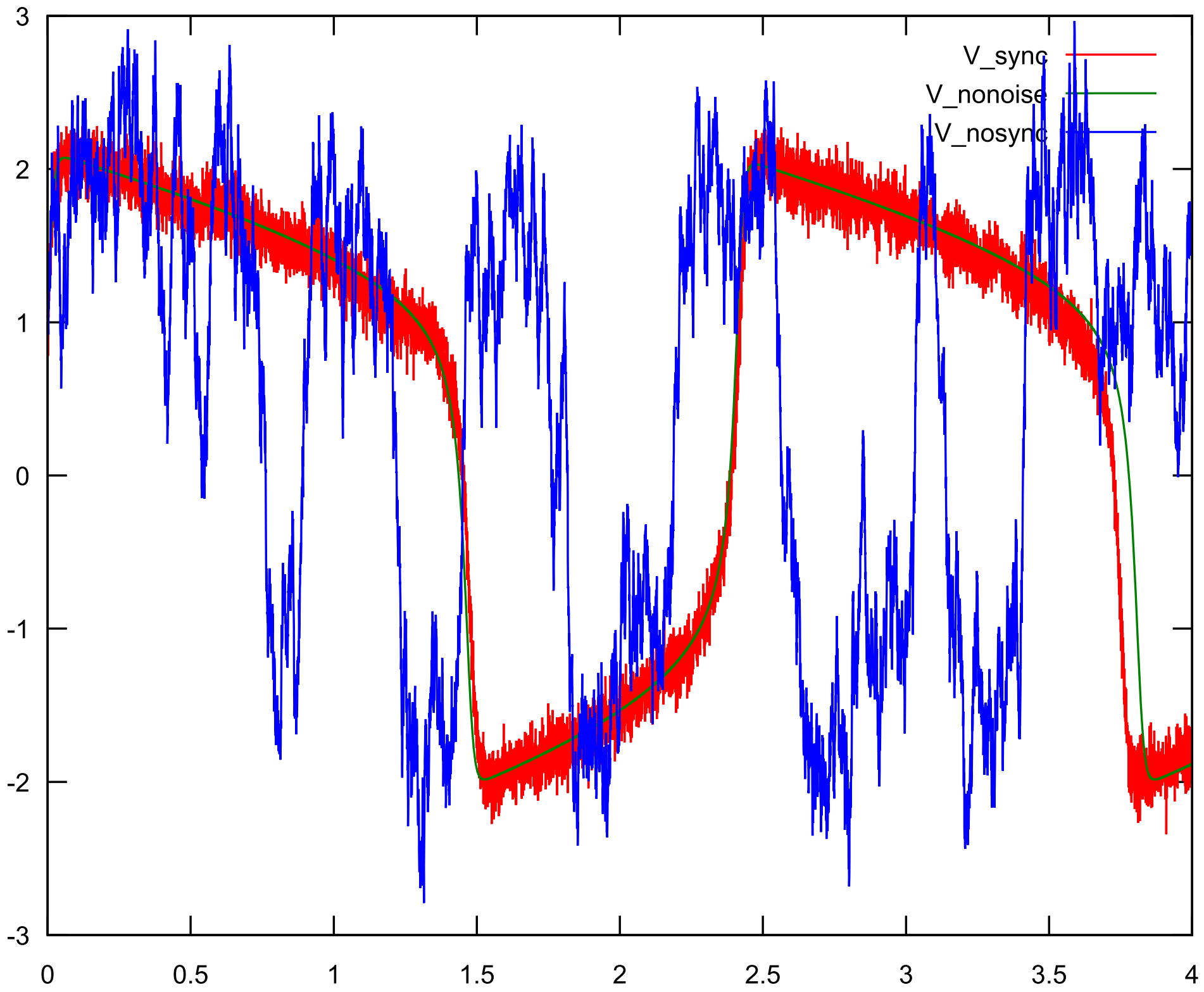
$$dM = \left(f(M) + \varepsilon \right) dt + \frac{1}{N} \sum_i \sigma_i(x_i) dW_i$$

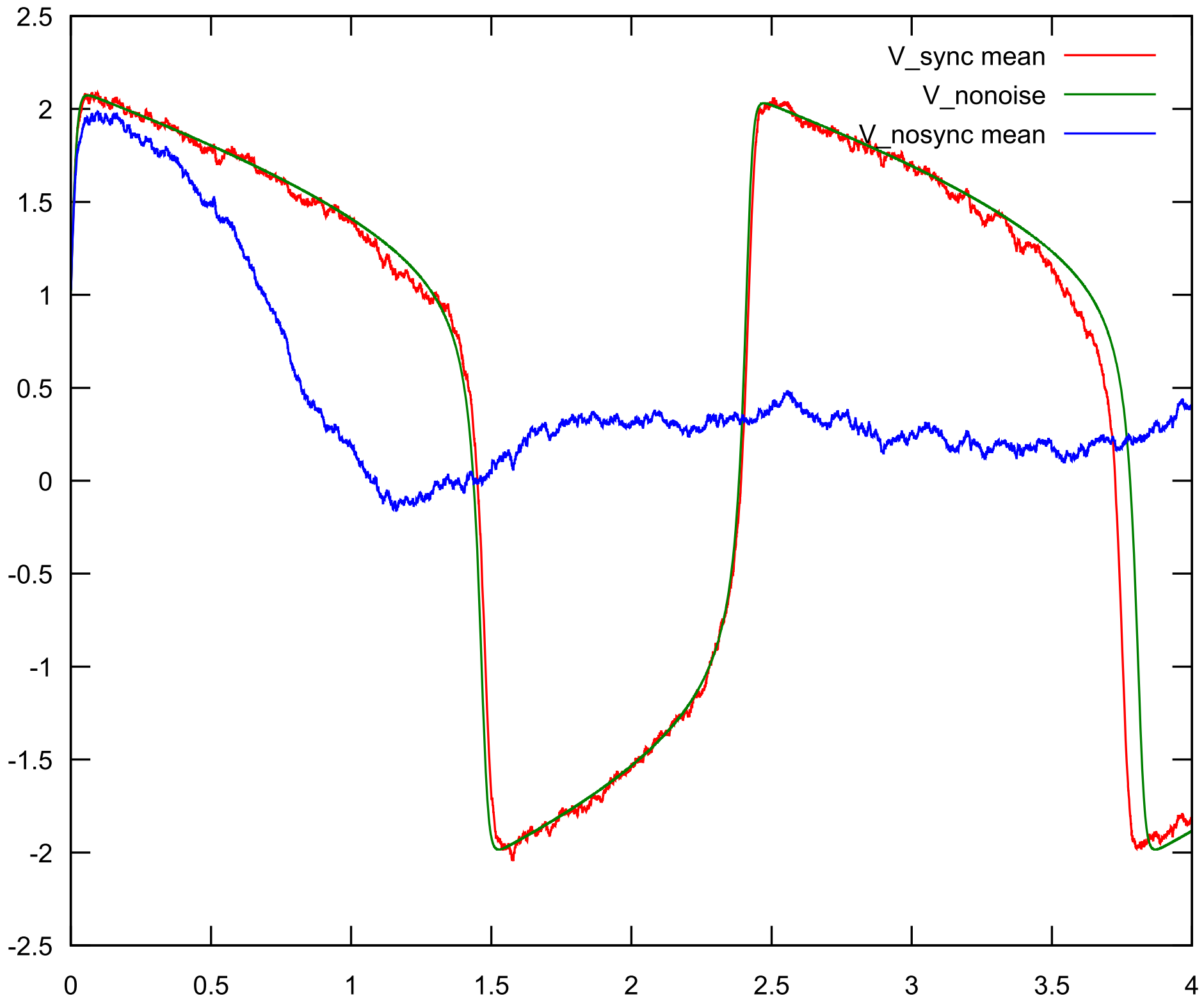
with

$$\mathbb{E} \|\varepsilon\| \leq \rho_{\max} \left(\frac{\partial^2 f}{\partial x^2} \right) \kappa(kN)$$

where ρ denotes spectral radius and

$$\kappa(kN) \rightarrow 0 \text{ as } kN \rightarrow +\infty$$





Multiple Time Scales

N layers, each partially contracting in state z_k at rate β_k ,

$$\dot{z}_k = f_k(z_{k-1}, z_k, z_{k+1}, t)$$

and the reduced system approximation

$$\dot{z}_k = \hat{f}_k(z_{k-1}, z_k, z_{k+1}^*(z_k, t), t)$$

Theorem 1 Assume that f_k is η_k - and ξ_k -Lipschitz with respect to z_k and z_{k+1} , and that each $z_{k+1}^*(z_k, t)$ is ρ_k -Lipschitz. Let $\gamma_k = 1 - \eta_k \tau_{k+1} \rho_{k+1}$. For small enough constants,

$$\gamma_k > 0 \quad \gamma_k \beta_k > \xi_k$$

the reduced system is contracting and its deviation from the exact system can be explicitly bounded.

Can also apply to virtual contracting systems, e.g. to sync of multi-time scale oscillators or quorum sensing.

(Nguyen et al., 2018)

Transverse Contraction

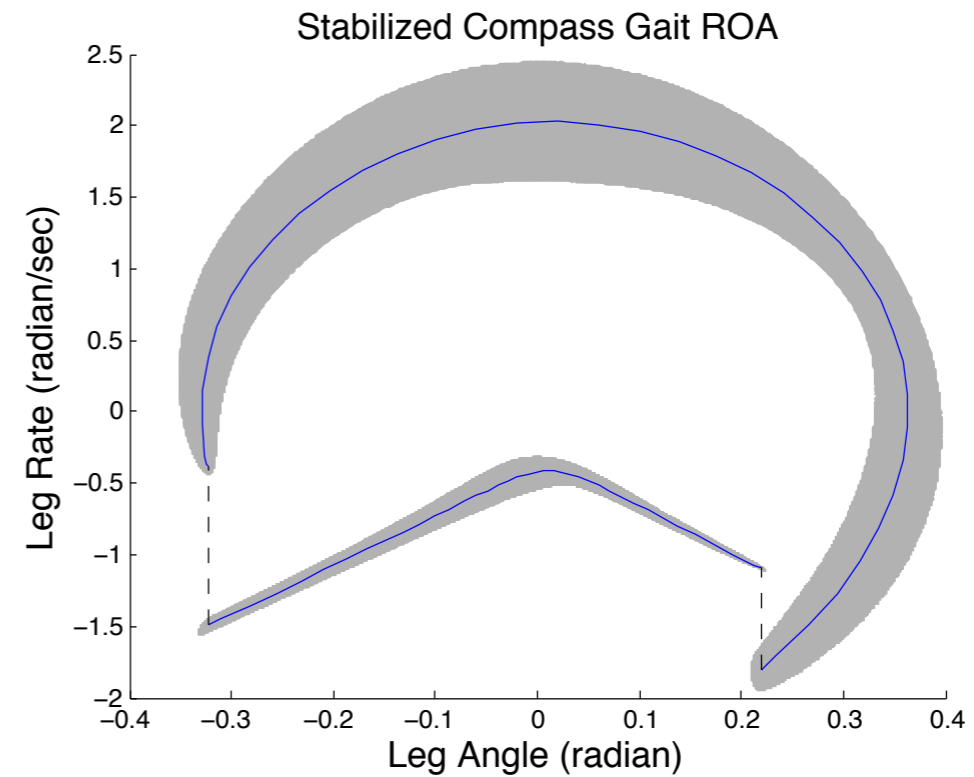
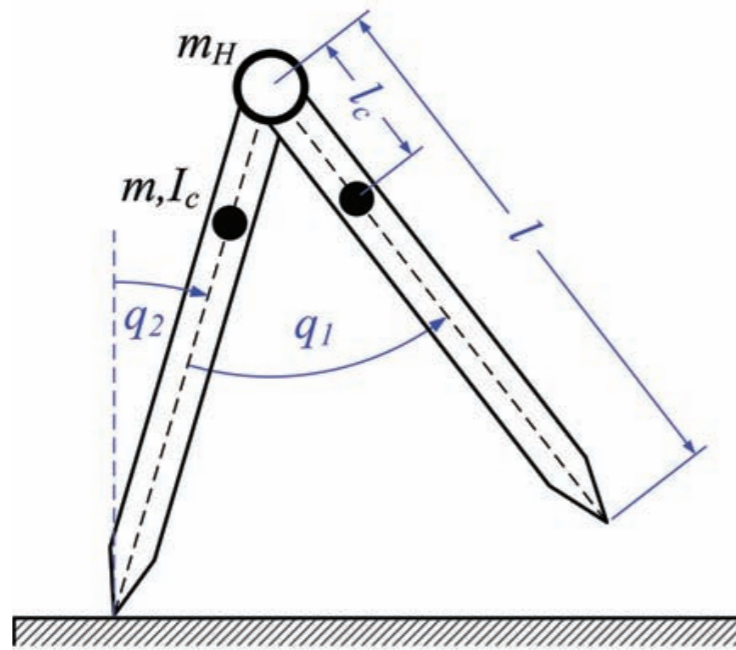
If the contraction condition holds on a compact manifold:

$$\delta'(\dot{M}(x) + J(x)'M(x) + M(x)J(x) + \lambda M(x))\delta \leq 0$$

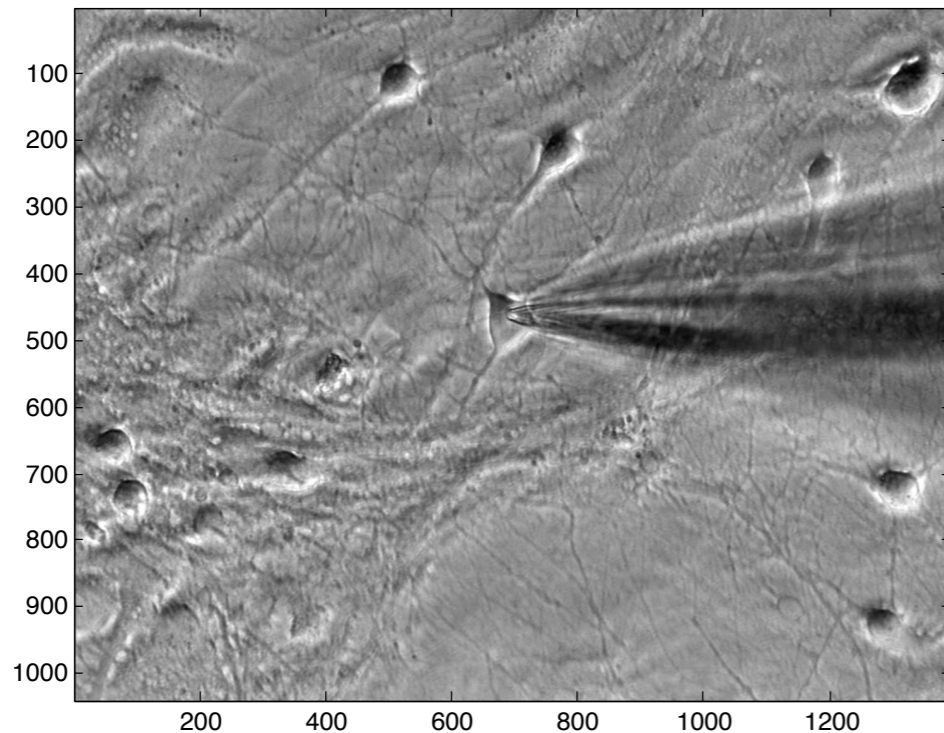
For all δ satisfying $\delta' M(x) f(x) = 0$
i.e. *transversal* w.r.t. $M(x)$

Then there exists a unique, globally orbitally stable limit cycle.

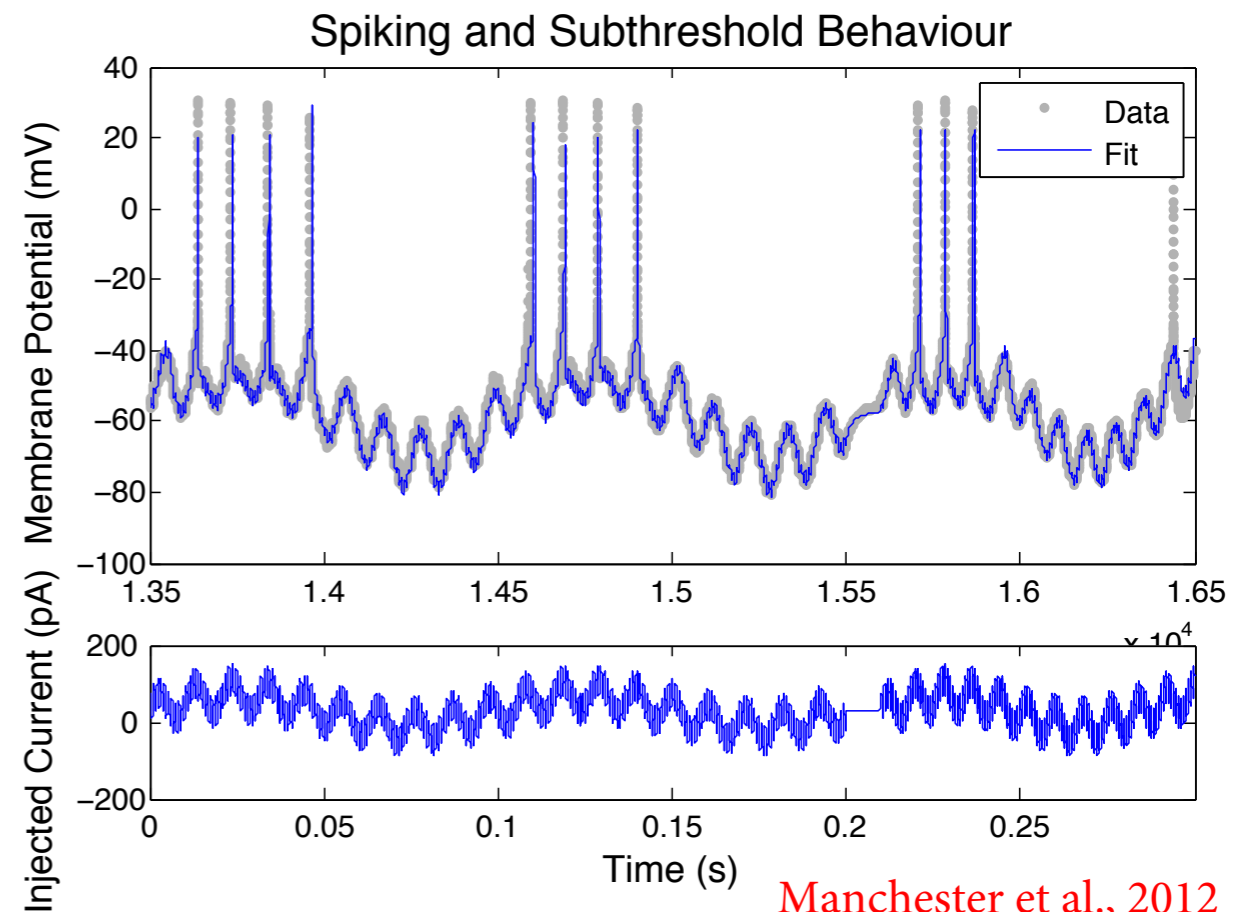
- Robust regions of stability for walking robots



- Identification of nonlinear systems with limit cycles

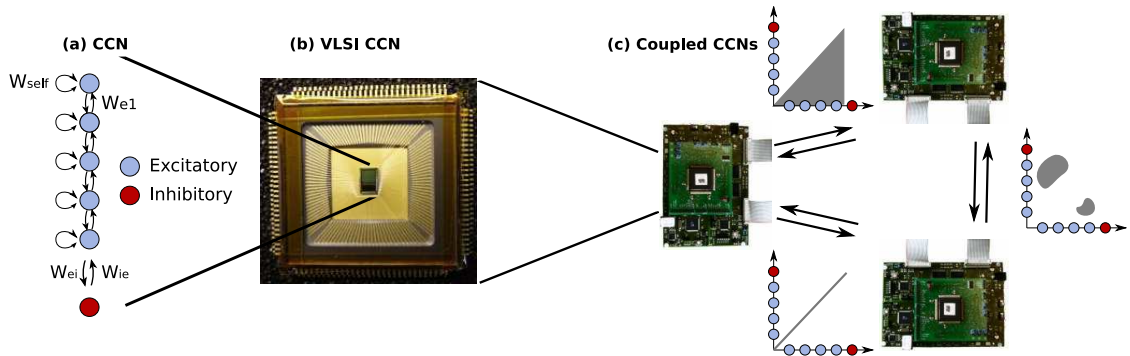


live rat neurons in culture



Manchester et al., 2012

Some Other Applications



- (Left) A network of 128 silicon Integrate-and-Fire neurons in a Cooperative and Competitive Network (CCN) topology (mediated respectively by 124 locally coupled excitatory neurons and 4 all-to-all coupled inhibitory neurons)
- (Right) The input and output from the network are events (also called spikes or action potentials). They can be mapped from one chip to another.

(G. Indiveri, E. Chicca, R. Douglas 2007)



Triantafyllou, et al., 2005

Seo, et al., 2007

Miller, et al., 2006

Chung, et al., 2008; Ramirez, et al., 2009



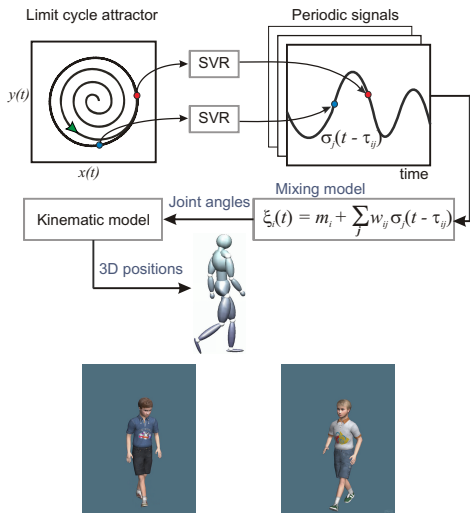
System Architecture for Real-Time Animation

- ▶ Movement primitives are extracted from MOCAP data by unsupervised learning.

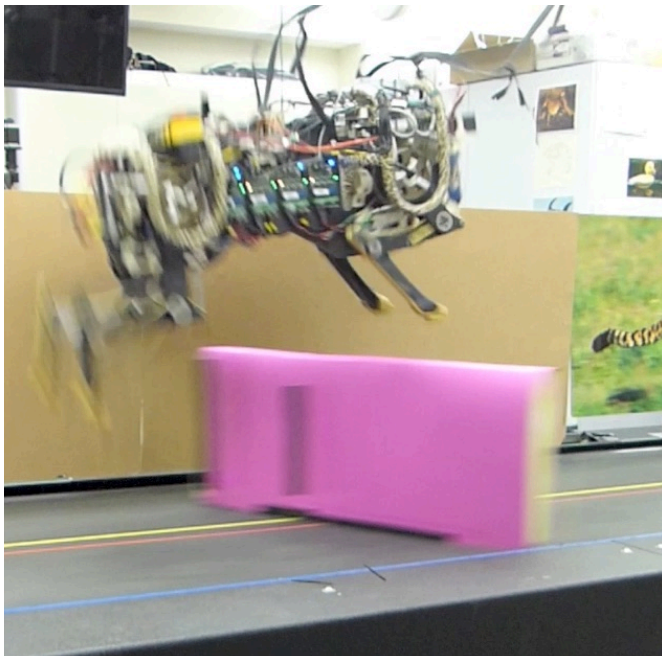
(Omlor & Giese, 2006)

- ▶ Primitives modelled by dynamical systems.
 - ▶ The phase space mapped onto the 'source signals' using Support Vector Regression.
- (Giese et al., 2008)

- ▶ Joint angles reconstructed by combining the (time-shifted) source signals linearly according to a learned mixture model.

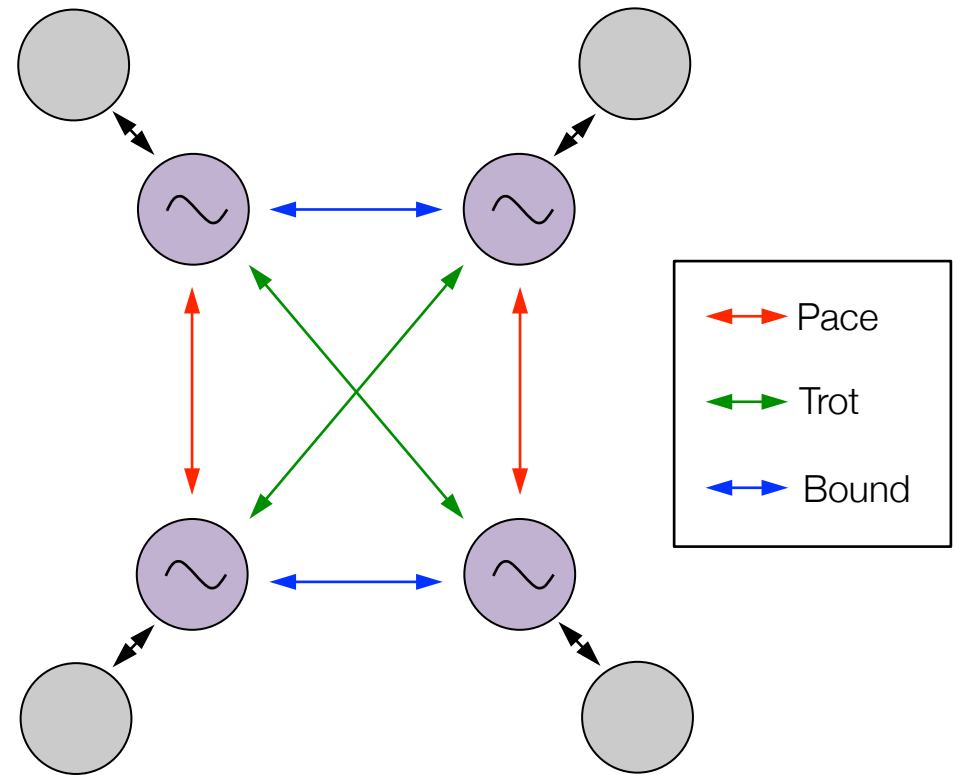


Movement Primitives and Sparse Synchronization for Gaited Locomotion



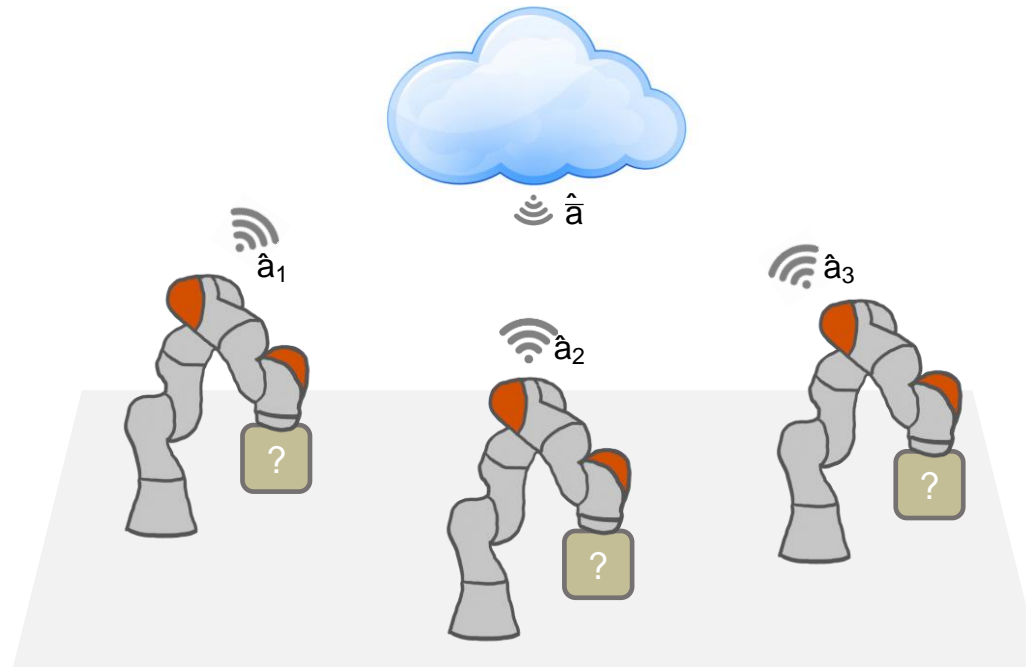
MIT Cheetah

[Park, Wensing, Kim, *RSS*, 2015]



[Wensing and Slotine, 2016]

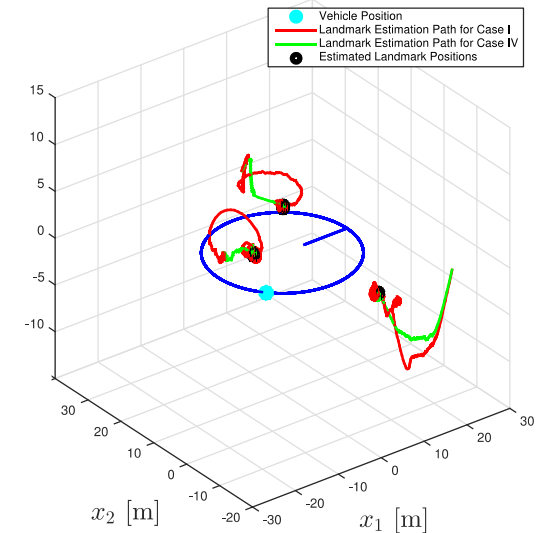
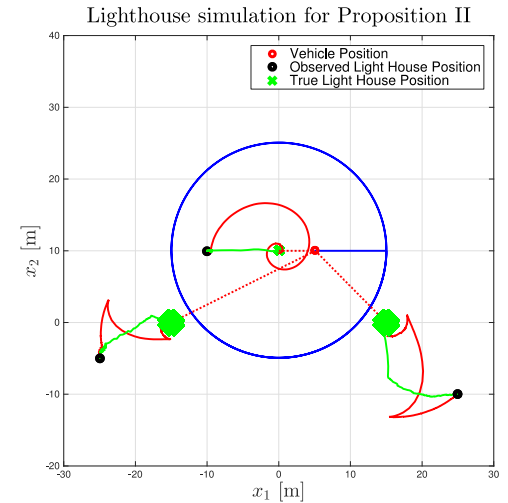
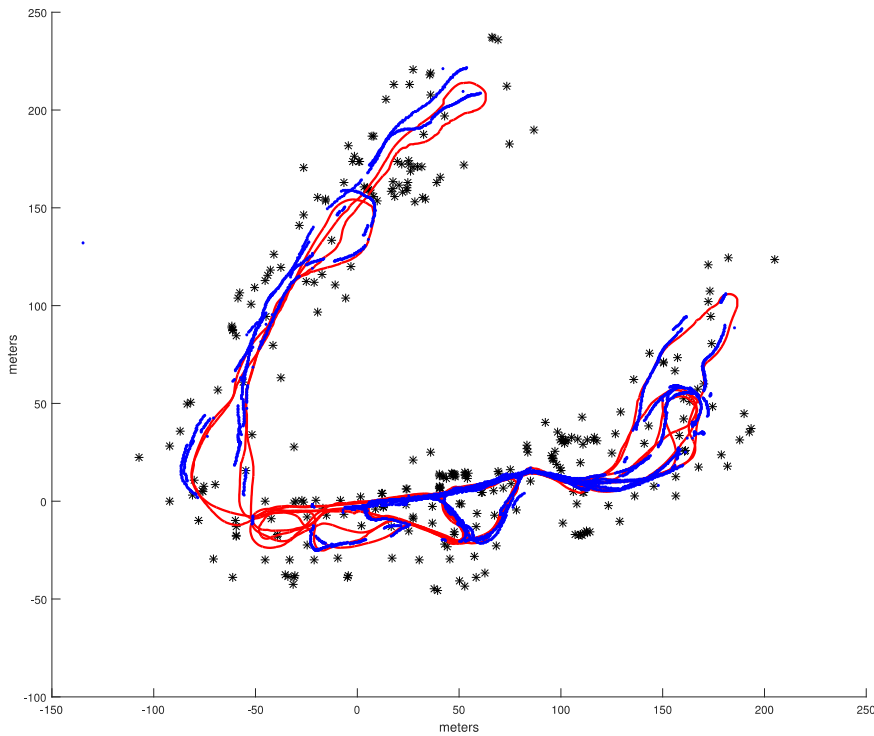
Decentralized Cloud-Based Teaming With Delay



Cloud-Based Update Law for Teaming

$$\dot{\hat{\mathbf{a}}}_i = -\mathbf{P}_i^{-1} [\mathbf{Y}_i^T \mathbf{s}_i - \frac{\mathbf{K}}{n} \sum_{j=1}^N (\hat{\mathbf{a}}_j(t - T) - \hat{\mathbf{a}}_i)]$$

Analytical SLAM Without Linearization



BACKGROUND

EKF-SLAM

- Take Measurements

$$y_1 = \theta = \arctan\left(\frac{x_1}{x_2}\right) \quad y_2 = r = \sqrt{x_1^2 + x_2^2}$$

- Standard extended Kalman Filter

$$\dot{\hat{\mathbf{x}}} = \mathbf{u} + \mathbf{K}(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}})$$

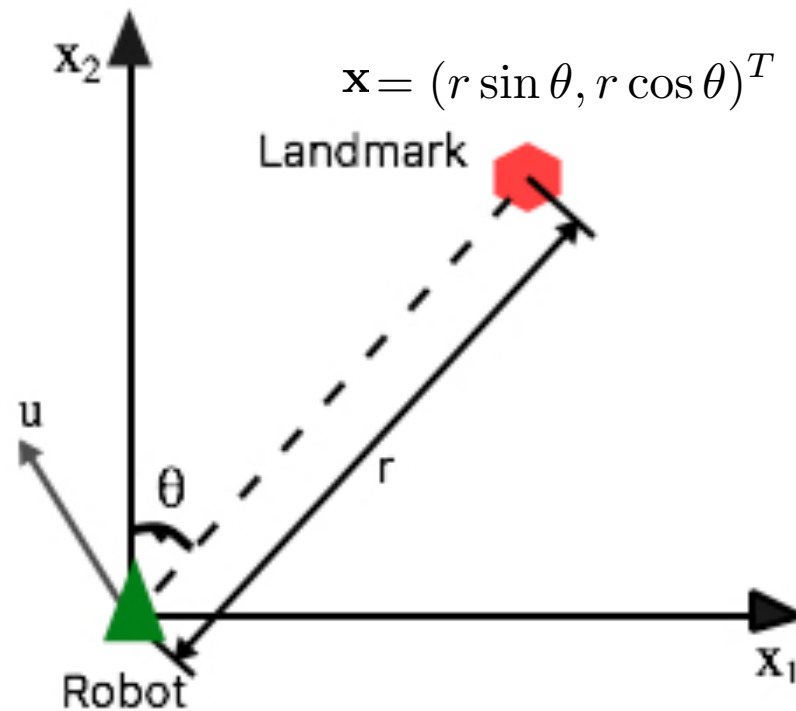
$$\dot{\mathbf{P}} = \mathbf{Q} - \mathbf{P}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{P}$$

- Linearize measurements with estimated Jacobian $\mathbf{H}_i(\hat{\mathbf{x}})$

$$\delta = \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} = \begin{bmatrix} \hat{x}_{i1} - \hat{x}_{v1} \\ \hat{x}_{i2} - \hat{x}_{v2} \end{bmatrix} \quad q = \delta^T \delta$$

$$\mathbf{H}_i(\hat{\mathbf{x}}) = \frac{1}{q} \begin{bmatrix} -\sqrt{q}\delta_1 & -\sqrt{q}\delta_2 & 0 & -\sqrt{q}\delta_1 & -\sqrt{q}\delta_2 \\ \delta_2 & -\delta_1 & q & -\delta_2 & \delta_1 \end{bmatrix}$$

$$\mathbf{y} = \mathbf{H} \mathbf{x} + \mathbf{v}(t)$$



LOCAL LTV KALMAN FILTER SLAM

Our Algorithm

- The fictive measurement vectors

$$\mathbf{h} = (\cos\theta, -\sin\theta)$$

$$\mathbf{h}^* = (\sin\theta, \cos\theta)$$

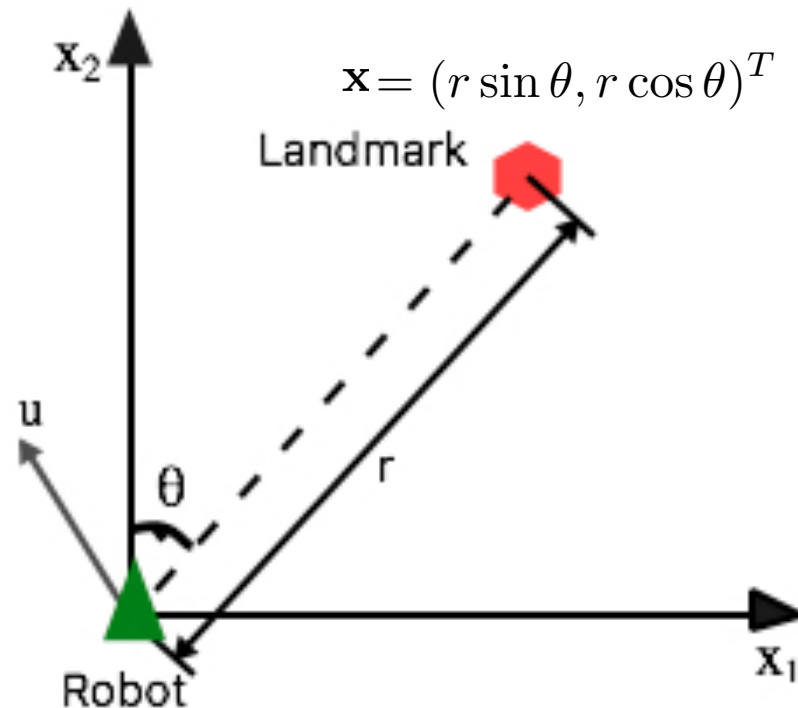
- In azimuth model

$$\mathbf{x} = (x_1, x_2)^T = (r \sin \theta, r \cos \theta)^T$$

- Rewrite the measurements into linear virtual constrains

$$\theta = \arctan\left(\frac{x_1}{x_2}\right) \Leftrightarrow \mathbf{h}\mathbf{x} = 0$$

$$r = \sqrt{x_1^2 + x_2^2} \Leftrightarrow \mathbf{h}^*\mathbf{x} = r$$



EXTENSION TO PINHOLE CAMERA MODEL

Pinhole Camera Model

- The pinhole camera model

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = -\frac{f}{x_3} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

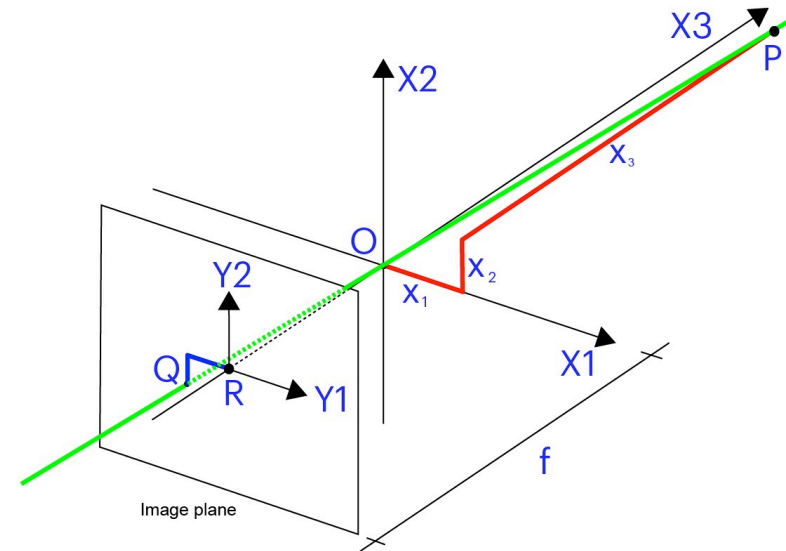
- which can be rewritten as LTV constraints on the states (x_1, x_2, x_3) as

$$\begin{bmatrix} f & 0 & y_1 \\ 0 & f & y_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{H}\mathbf{x} = \mathbf{0}$$

- LTV Kalman Filter

$$\dot{\hat{\mathbf{x}}} = -\mathbf{u} - \Omega\hat{\mathbf{x}} + \mathbf{K}(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}})$$

$$\dot{\mathbf{P}} = \mathbf{Q} - \mathbf{P}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{P} - \Omega\mathbf{P} - \mathbf{P}\Omega^T$$



EXTENSION TO STRUCTURE FROM MOTION

Structure from Motion

- Remember that for the pinhole camera model

$$\begin{bmatrix} f & 0 & y_1 \\ 0 & f & y_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{H}\mathbf{x} = \mathbf{0}$$

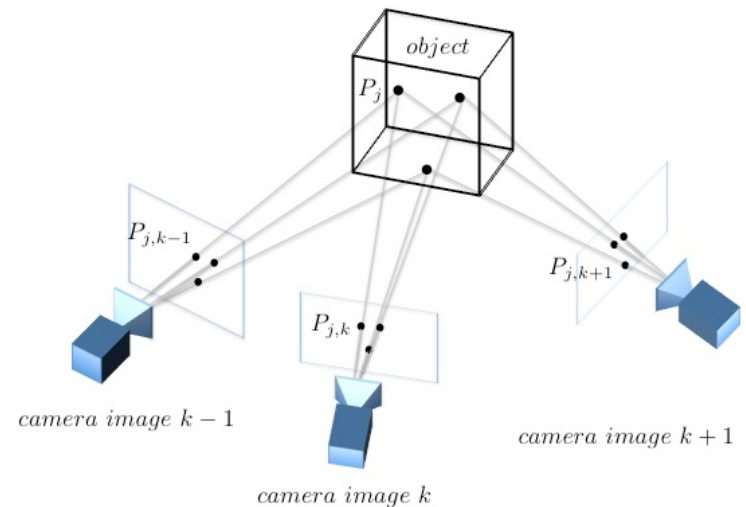
- Rewrite in global coordinates

$$\begin{bmatrix} f & 0 & y_{i1} \\ 0 & f & y_{i2} \end{bmatrix} \mathbf{T}(\beta)(\mathbf{x}_i - \mathbf{x}_c) = \mathbf{0}$$

- LTV Kalman Filter

$$\dot{\hat{\mathbf{x}}} = -\mathbf{u} - \Omega\hat{\mathbf{x}} + \mathbf{K}(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}})$$

$$\dot{\mathbf{P}} = \mathbf{Q} - \mathbf{P}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{P} - \Omega\mathbf{P} - \mathbf{P}\Omega^T$$



DECOUPLED UNLINEARIZED NETWORKED KALMAN-FILTER

Complete Algorithm for SLAM-DUNK

- Separate LTV Kalman filter for each single pair of landmark and virtual vehicle, including both the measurement and the consensus following for the virtual vehicle

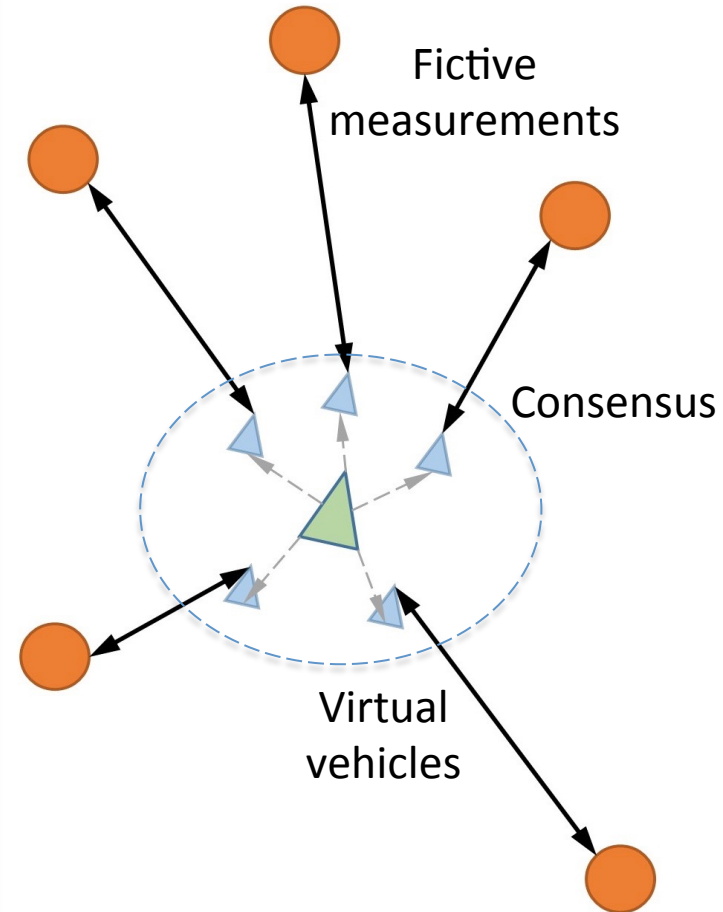
$$\mathbf{H}_{iv} = \begin{bmatrix} \mathbf{H}_i \mathbf{T}(\beta) \\ \mathbf{H}_{i'} \end{bmatrix}$$


$$\begin{bmatrix} \dot{\hat{\mathbf{x}}}_i \\ \dot{\hat{\mathbf{x}}}_{vi} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{u} \end{bmatrix} + \mathbf{P}_i \mathbf{H}_{iv}^T \mathbf{R}^{-1} (\mathbf{y}_i - \mathbf{H}_{iv} \begin{bmatrix} \hat{\mathbf{x}}_i \\ \hat{\mathbf{x}}_{vi} \end{bmatrix})$$

$$\dot{\mathbf{P}}_i = \mathbf{Q} - \mathbf{P}_i \mathbf{H}_{iv}^T \mathbf{R}^{-1} \mathbf{H}_{iv} \mathbf{P}_i$$

- The consensus \mathbf{x}_{vc} is the best estimation from the weighted average that maximizes the likelihood

$$\mathbf{x}_{vc} = \left(\sum_{i \in O} \Sigma_{vi}^{-1} \right)^{-1} \sum_{i \in O} (\Sigma_{vi}^{-1} \mathbf{x}_{vi})$$

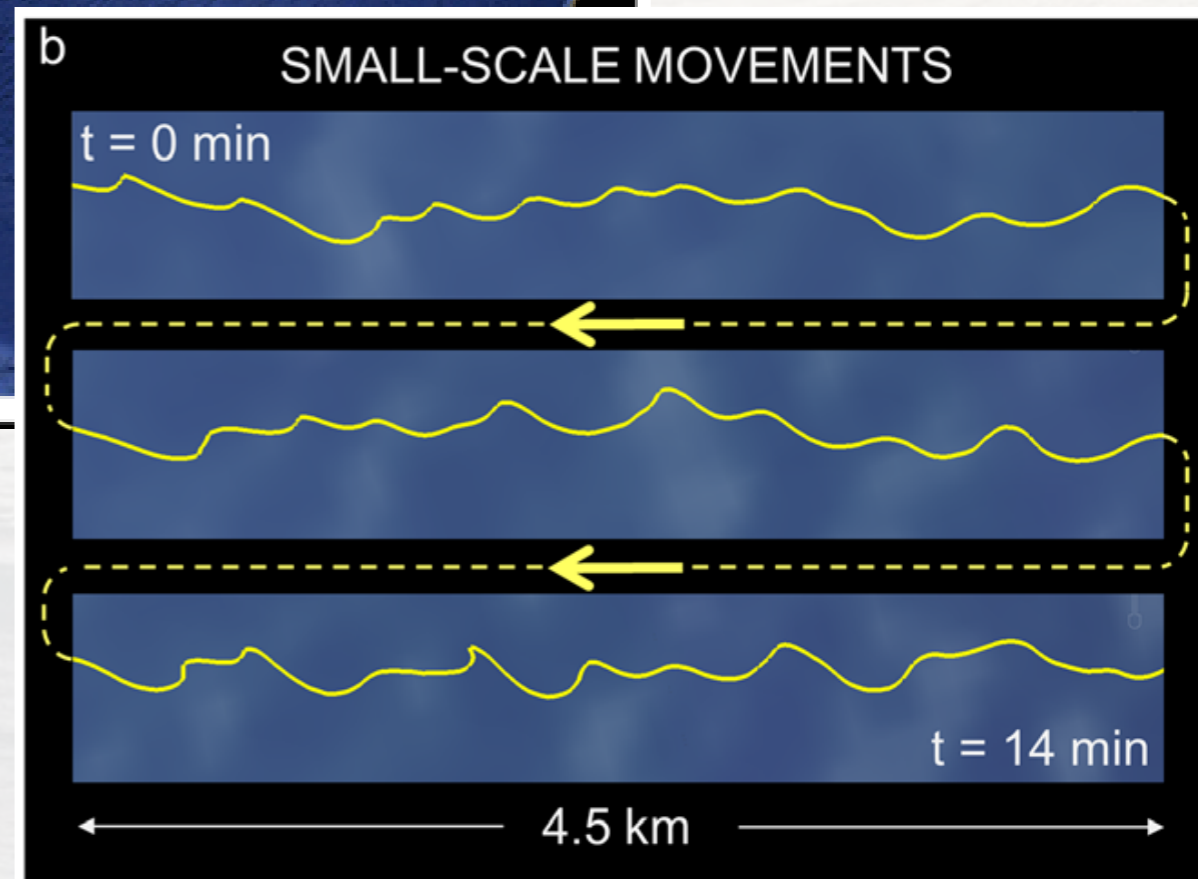
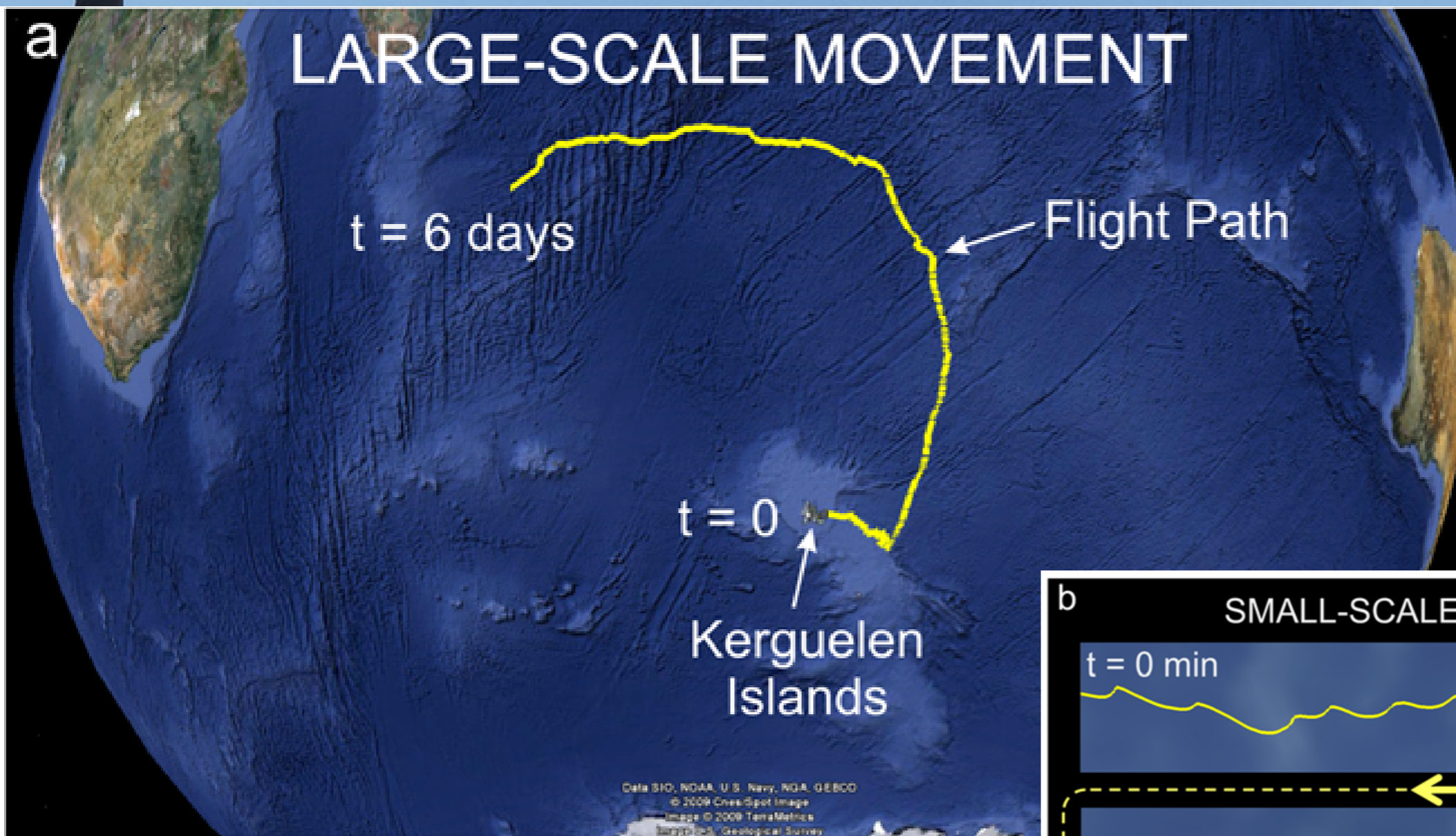


A silhouette of an albatross in flight, positioned to the left of the title text.

Dynamic soaring beyond biomimetics: design and control of an albatross-inspired wind-powered system

Gabriel Bousquet
Michael Triantafyllou
Jean-Jacques Slotine

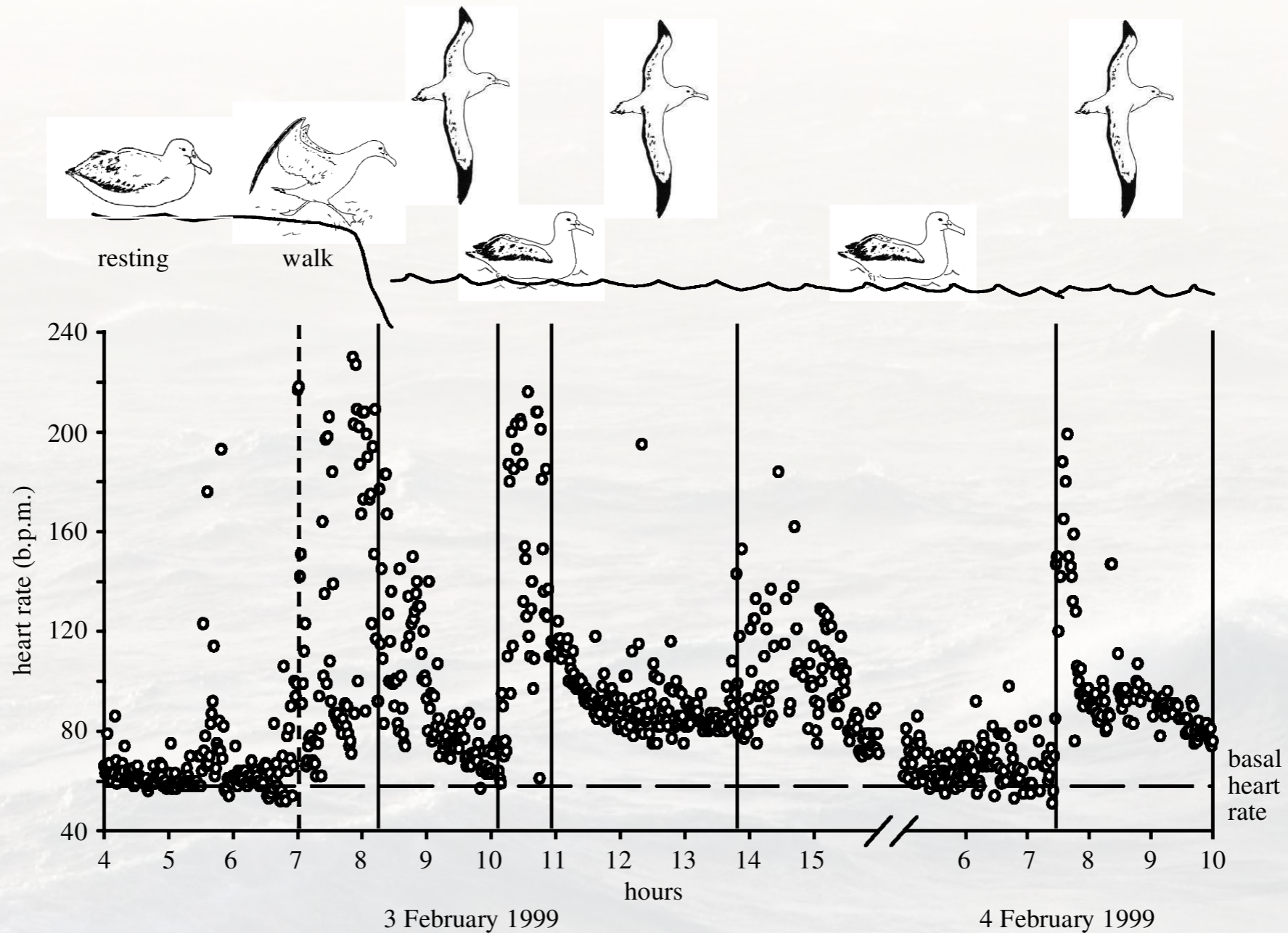
GPS recordings of Albatrosses



The albatross, a wind-powered bird

Fast and fuel efficient? Optimal use of wind by flying albatrosses

H. Weimerskirch^{1*}, T. Guionnet¹, J. Martin¹, S. A. Shaffer² and D. P. Costa²

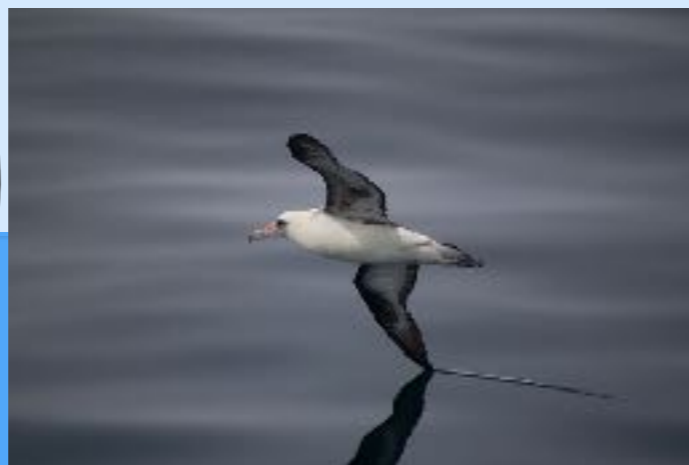
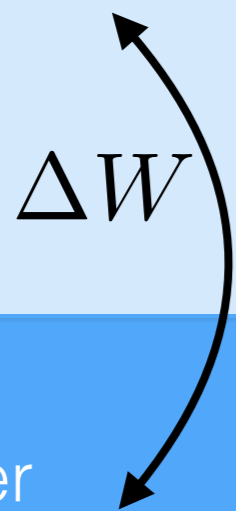


Performance of wind-powered travel

$$V = \frac{\text{Sail lift}}{\text{Total drag}} \Delta W \cos \Psi$$

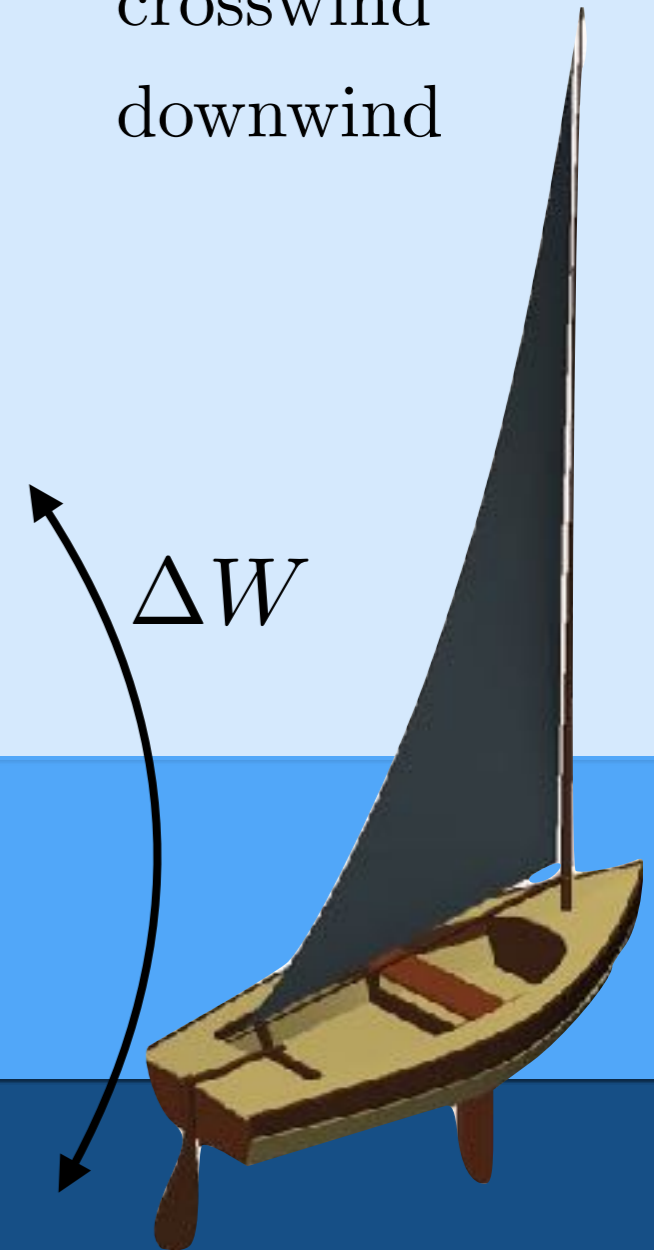
$$\begin{cases} \Psi = \pi/2 & \text{upwind} \\ \Psi = 0 \text{ or } \pi & \text{crosswind} \\ \Psi = -\pi/2 & \text{downwind} \end{cases}$$

Wind layer
100%



Air
boundary layer
75%

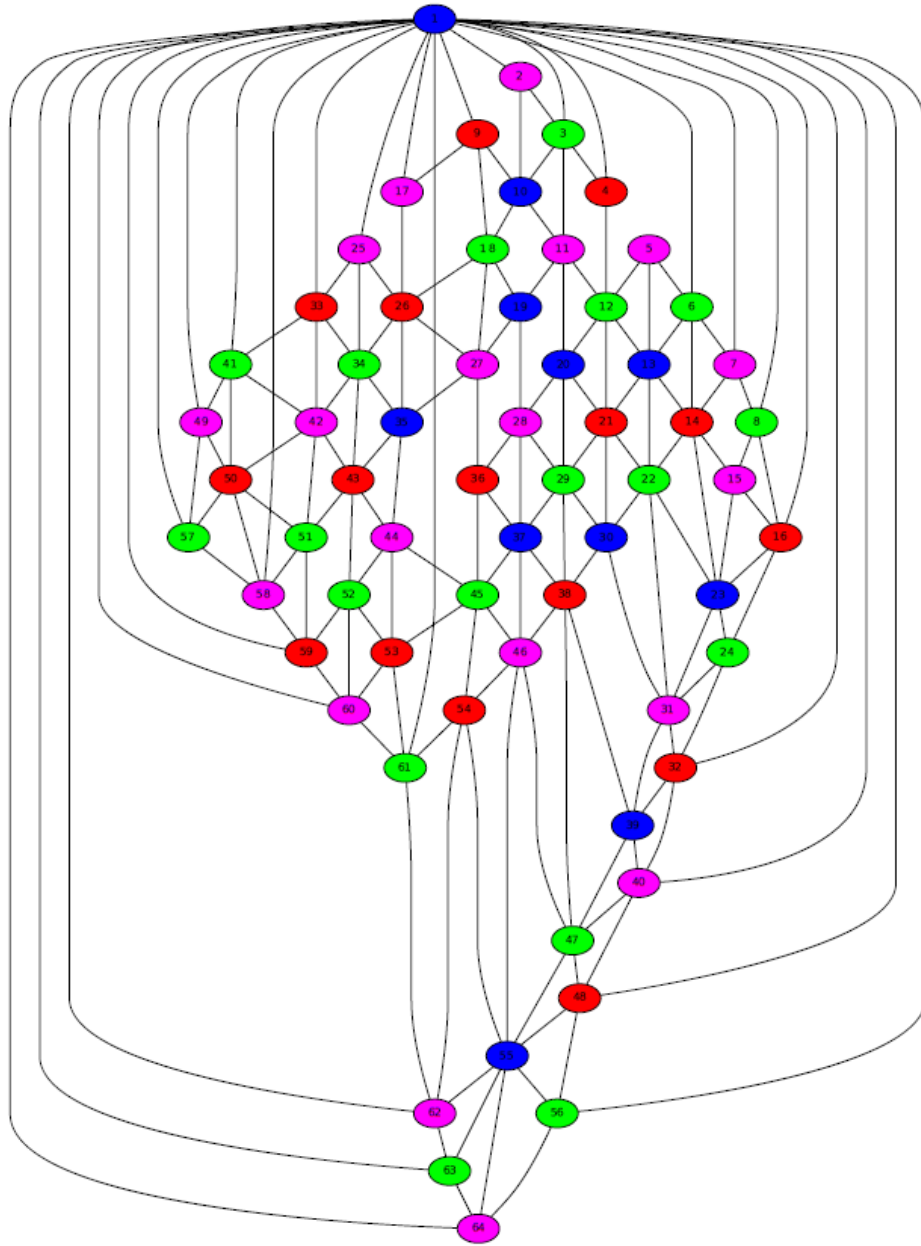
Water
0%



Controlled Instabilities

- Fast maneuvering
- Path of least resistance
- Constraint satisfaction
- Expansion and pruning

Graph Coloring



64 Nodes, 4 states each, 400 constraints

Sudoku

4	9	6	8	7	1	5	3	2
8	7	1	5	3	2	6	4	9
2	5	3	6	9	4	1	7	8
7	8	2	3	5	9	4	6	1
1	4	5	2	6	7	9	8	3
3	6	9	4	1	8	7	2	5
9	1	8	7	4	3	2	5	6
5	2	4	9	8	6	3	1	7
6	3	7	1	2	5	8	9	4

red = Constraint black = Solution

81 Nodes, 9 states each, ~10,000 constraints

An electrifying way to
make ammonia pp. 609 & 637

Breaking symmetry in a
plant embryo pp. 622 & 636

Driving genes into
ecosystems? p. 626

Science

\$10
8 AUGUST 2014
science.org

AAAS

Brain inspired chip

Neuroscience provides a
path toward energy-efficient
computing pp. 614 & 668

nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

TAMING COMPLEXITY

The mathematics of network control — from
cell biology to cellphones **PAGES 158 & 167**

ANTIBIOTICS

A SPOONFUL OF SUGAR

Carbohydrate 'helpers' boost
antibacterial efficacy

PAGE 216

POLICY

WHO NEEDS CHANGE

A World Health Organization
for today's world

PAGE 143

GEOLOGY

WAITING FOR VESUVIUS

What to do when Europe's
most deadly volcano blows

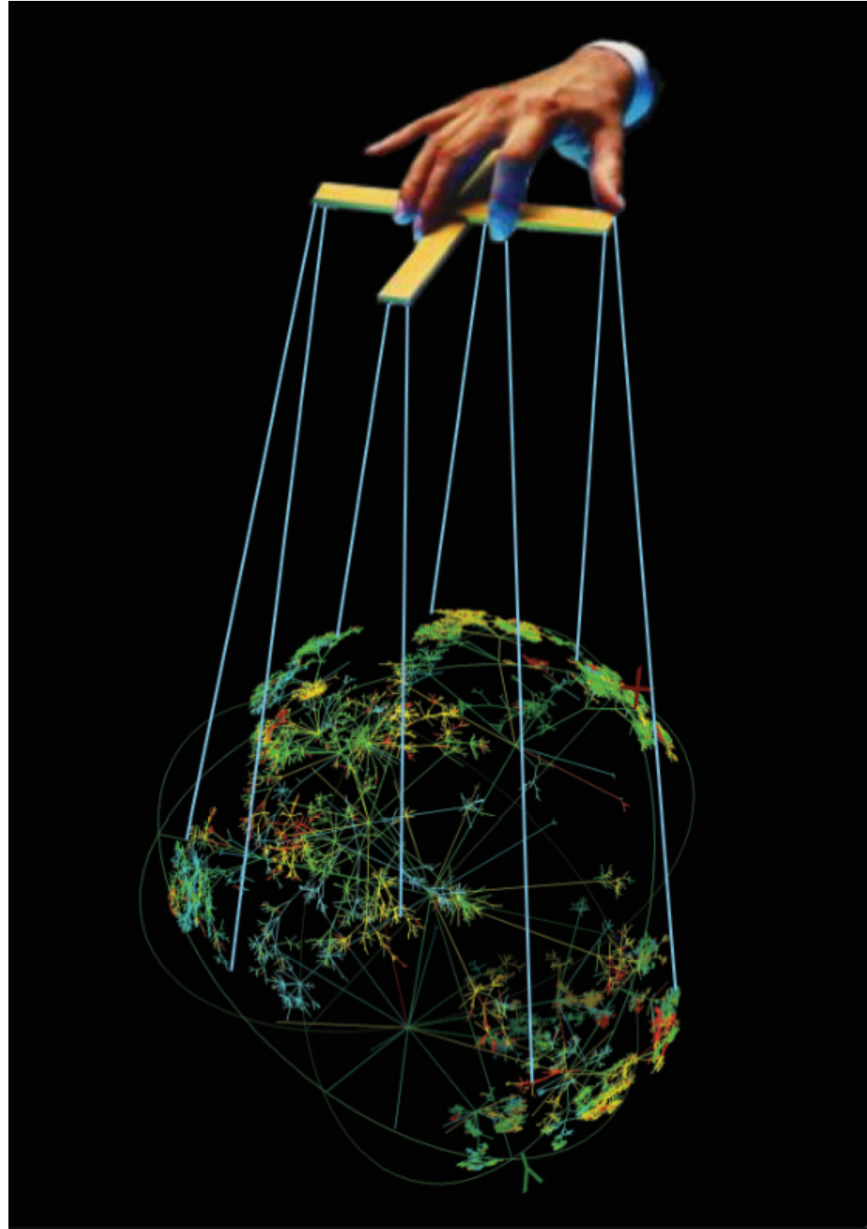
PAGE 140

NATURE.COM/NATURE

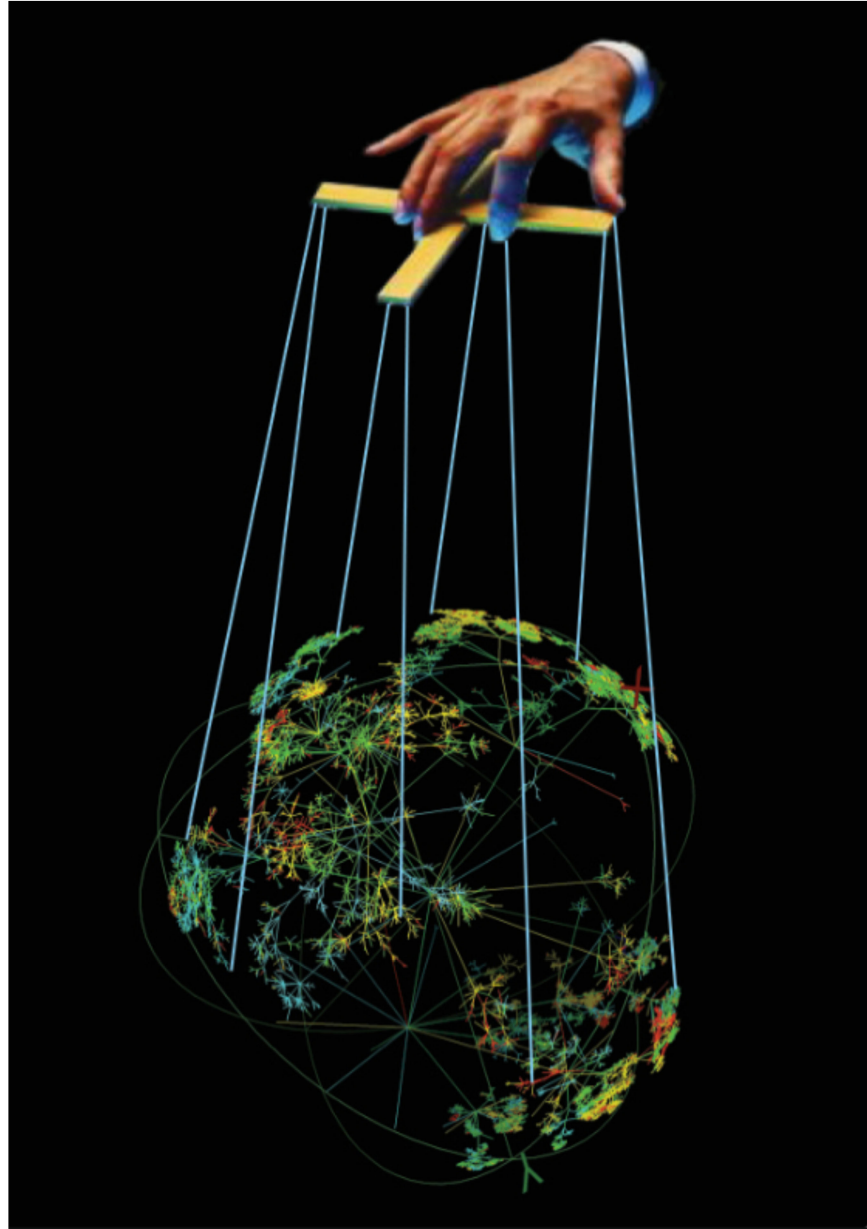
12 May 2011



Controllability of Complex Networks



Controllability of Complex Networks



Which nodes should one drive in order to control the **entire** system?

Observing complex systems

The background of the cover is a complex network visualization. It consists of numerous small, glowing green and yellow nodes connected by thin, faint lines. The nodes are arranged in a way that suggests a hierarchical or interconnected structure, with some clusters and some long, thin vertical chains. The overall effect is that of a dynamic, complex system being observed.

Transition to sustainable energy

Explaining marine ecosystem variability

Evolution of fairness

Mechanisms of cocaine-seeking behavior

Facilitated Variation

Gerhart and Kirschner, P.N.A.S. 104, 2007

Three billion years ago, in early prokaryotic organisms

Components of energy metabolism, biosynthesis of the 60 building blocks, DNA replication, DNA transcription to RNA, translation of RNA to protein, lipid membrane synthesis, transmembrane transport

Two billion years ago, in early eukaryotic cells

Components of the formation of microfilament and microtubule cytoskeletons, motor proteins moving materials along the cytoskeletons, contractility processes, movement of the cell by cilia and ruffling membrane action, shuttling of materials between intracellular organelles, phagocytosis, secretion, chromosome dynamics, a complex cell cycle driven by protein kinases and protein degradation, sexual reproduction with meiosis and cell fusion

One billion years ago, in early multicellular animal life forms

Components of 15–20 cell–cell signaling pathways, cell adhesion processes, apical basal polarization of cells, junction formation, epithelium formation, specialization of cells toward physiological ends, some developmental processes of the single-celled egg to the adult

Near pre-Cambrian, in animals with early body axes

Components of complex developmental patterning, such as anteroposterior axis formation (Wnt/Wnt antagonist gradients) and dorsoventral axis formation (Bmp/antagonist gradients), inductions, complex cell competence, additional specialized cell types, formation of the body plan's map of selector gene compartments (both transcription factors and signaling proteins), various regulatory processes

Link Control

Nepusz and Vicsek, Controlling edge dynamics in complex networks
Nature Physics, 8(7), 2012

Dual problem and properties, simple algorithm.

Slotine and Liu, Complex networks: The Missing Link
Nature Physics, 8(7), 2012

Evolution may be based on ancient optimized components whose connections are the main target of natural selection (facilitated variation). From this perspective, controllability of evolution and development is primarily a link-based concept.

<http://mit.edu/nsl>