AS-74.4192 Elementary Cybernetics

Lecture 7:

Emergent Models

HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Automation and Systems Technology
Cybernetics Group

- Often, the systems are seen from above, and the abstracted individual agents cannot easily be distinguished
- The experiences propose that the system variables can be analyzed in terms of PCA or related feature extraction
- So, the cyberneticity reduces to data preprocessing using traditional data compression means?

NO – the claim here is that truly new thinking is needed

- It is not only preprocessing – now *the whole chain of modeling changes, as well as the end results, or the models and their interpretations, and ways of their application.*

Once more: About cybernetic systems

- Abstract over individuals spatially and temporally
- Cybernetic system is a complex system that is characterized by dynamic equilibrium among opposing tensions
- The balances characterize dynamic attractors that are visible in the data and thus relevant in that domain
- Interacting systems are reactive, controlling each other, the overall dependencies becoming pancausal
- The system gets towards better and better coupling with its environment, meaning more fluent information flow
- During evolution (natural or not) the controls become more and more stringent and the overall system becomes stiffer
- Final result: "Degrees of freedom are eliminated" – **WHAT?**

Cost criteria characterizing behaviors

- The cost for social (clever) agents is

$$J(u) = \frac{1}{2}\bar{x}^T \left(\mathrm{E}\left\{ \overline{xx^T} \right\} \right) \bar{x} - \bar{x}^T \mathrm{E}\left\{ \overline{xu^T} \right\} u$$

- Correspondingly, the cost for selfish agents is

$$J(u) = \frac{1}{2}\bar{x}^T \left(\mathrm{E}\left\{ \overline{xx^T} \right\} + Q^{-1} \right) \bar{x} - \bar{x}^T \mathrm{E}\left\{ \overline{xu^T} \right\} u$$

This can be written also as

$$J(u) = -\frac{1}{2}\bar{x}^T \left(\mathrm{E}\left\{ \overline{xx^T} \right\} + Q^{-1} \right) \bar{x}$$

← Maximize "emergy"

or

$$J(u) = -\frac{1}{2}\bar{x}^T \mathrm{E}\left\{ \overline{xu^T} \right\} u.$$

← Maximize mutual information (as defined here)

Why? Check the optimum (also see Lec. 11) !

About the emergent patterns

- The cost criterion characterizing cybernetic agents

$$J(u) = \frac{1}{2}\bar{x}^T \left(\mathrm{E}\left\{ \overline{xx^T} \right\} + Q^{-1} \right) \bar{x} - \bar{x}^T \mathrm{E}\left\{ \overline{xu^T} \right\} u$$

can be rewritten to read (because $\bar{x} = \phi^T \bar{u}$)

$$J(u) = \frac{1}{2}\bar{x}^T \phi^{\mathrm{T}} \mathrm{E}\left\{ \overline{uu^T} \right\} \phi \bar{x} - \bar{x}^T \phi^{\mathrm{T}} \mathrm{E}\left\{ \overline{uu^T} \right\} u + \frac{1}{2}\bar{x}^T Q^{-1} \bar{x}$$

- A new formulation for the "emergent pattern" is found:

$$J(u) = \frac{1}{2}\left(u - \phi\bar{x}\right)^T \mathrm{E}\left\{ \overline{uu^T} \right\} \left(u - \phi\bar{x}\right) + \frac{1}{2}\bar{x}^T Q^{-1} \bar{x} - \frac{1}{2} u^T \mathrm{E}\left\{ \overline{uu^T} \right\} u$$

Vanishes for clever agent

Constant – no effect

Pattern matching

- One can also formulate the cost criterion as

$$J(x,u) = \frac{1}{2}(u - \phi x)^T \, \mathrm{E}\{uu^T\} (u - \phi x)$$

- This means that the neuron grid carries out *pattern matching* of input data
- Note that the traditional maximum (log)likelihood criterion for Gaussian data (suffering of invertibility problems) would be

$$J(x,u) = \frac{1}{2}(u - \phi x)^T \, \mathrm{E}\{uu^T\}^{-1} (u - \phi x)$$

- **Now:** More emphasis on the most visible directions, in the direction of *freedoms*

Starting point towards the new view of systems

Models of today's systems: *Constraints*

- How can a (locally linear) model be described?
- Traditional analysis (modeling) and design (synthesis) methods are based on models of *constraints*

  $$y = \theta^T \mu$$

  Here, θ is the vector of parameters, μ contains the variables, and *y* is the output
- It is assumed that the data are somehow bound together, and it is this bond that captures the essence of the system
- Reason for this thinking is the dominant role of *natural language* when describing nature and natural laws (?)
- To "cybernetize" this, study a practical example ...

Wittgenstein: "World = totality of states of affairs"

Example

- Take traditional *system identification*:

  Simple static matching between time-series data is done – giving constraint equations between signals in the form

$$y(k) = \sum_{i=1}^{n} a_i \, y(k-i) + \sum_{j=1}^{n} b_j \, u(k-j)$$

- However, huge amounts of theory has been devoted to this – mainly due to two reasons:
  - The model structure does not exactly hold because there is noise; and these noise properties need to be analyzed as separate dynamical systems
  - All model parameters are assumed to be equally "visible" in data; as this is not the case, the algorithms can have lousy numerical properties

- Now: Both of these problems will be solved (by accident!?)

Towards homogeneity

- Augment data vector to have a "homogeneous" view – include y among other data:

$$u = \begin{pmatrix} \mu \\ \text{----} \\ y \end{pmatrix} \quad \text{and} \quad \Theta = \begin{pmatrix} \theta \\ \text{------} \\ -1 \end{pmatrix} \Big/ \sqrt{1+\theta^T\theta}$$

$$y = \theta^T \mu \quad \Leftrightarrow \quad 0 = \Theta^T u$$

- Here all variables have an identical role
- Representation is non-unique – to reach uniqueness, Θ can be normalized to unit length above

The same variables are inputs to some and outputs to some other subsystems!

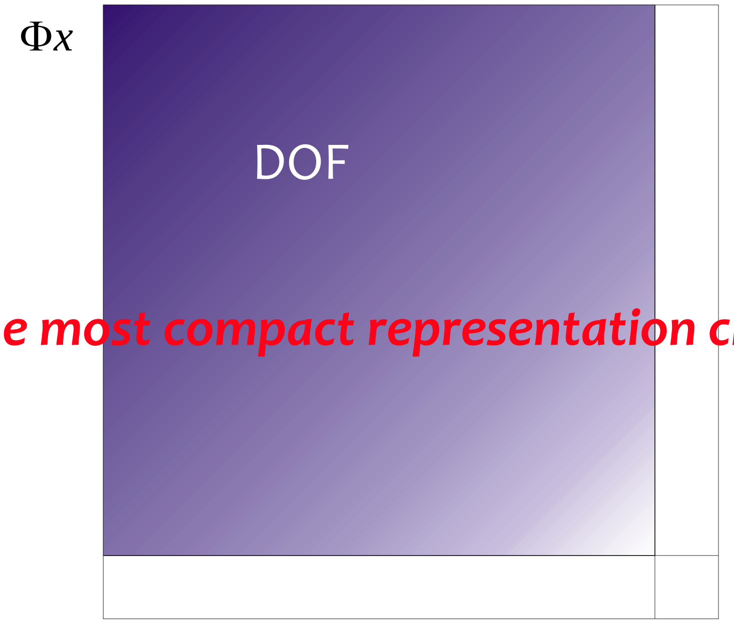
Degrees of freedom

- In principle, if there are n independent variables, there are n degrees of freedom in the data space
- Traditional view: Each constraint equation decreases the degrees of freedom exactly by one (any one of the variables can be expressed as a linear combination of the others)
- However, in practice, the degrees of freedom differ from any integer number
  - Noise increases DOF back to n
  - Interdependencies (more or less explicit) decrease DOF in *practice*
- Modern view: DOF should be studied numerically rather than symbolically!
  - Compare to controllability/observability: Exactly zero determinants are never found from data – but in practice problems often emerge

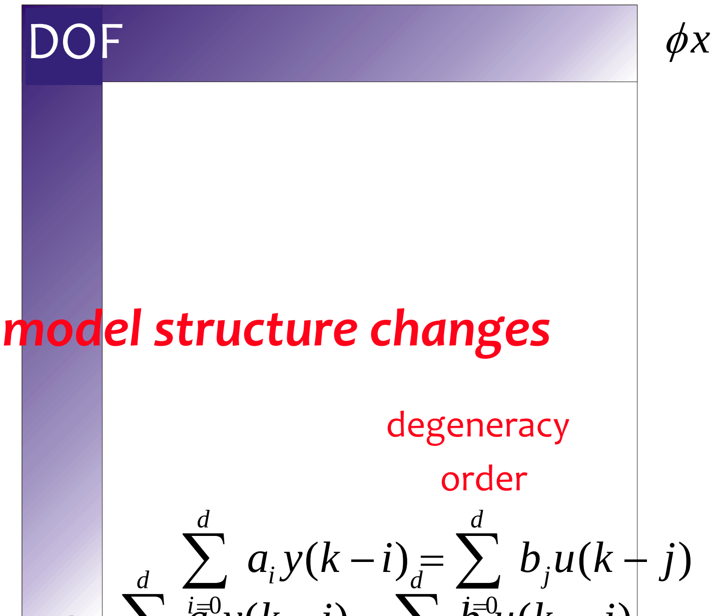
View from above: “Emergent Models”

- Data high-dimensional
- Few connections = constraints
- Many degrees of freedom left

- Data equally high-dimensional
- Many constraints
- Few degrees of freedom (right!)



$$\sum_{i=0}^{d} a_i y(k-i) = \sum_{j=0}^{d} b_j u(k-j)$$



The most compact representation changes = model structure changes

degeneracy

order

$$\sum_{i=0}^{d} a_i y(k-i) = \sum_{j=0}^{d} b_j u(k-j)$$

Constraints vs. *freedoms*

determinism vs. stochastics?

- Claim: The degrees of freedom are more characteristic to a system than the constraints are
- Reason: In deeply interconnected systems, emphasis on freedoms is a more compact representation of the system
- The constraint model determines a line in the data space – "null space", where there is no freedom among data
- "Axes of freedom" = remaining subspace that is *orthogonal to the null space* = basis of a NEW MODEL STRUCTURE
- The eigenvalue decomposition of the data covariance matrix reveals in which directions there is variation in the data and how much: *Eigenvectors = axes of freedom*, and *eigenvalues = their relevances*

Remember that it was the principal subspace that was the key issue also in cybernetic populations!

Example

- Assume that

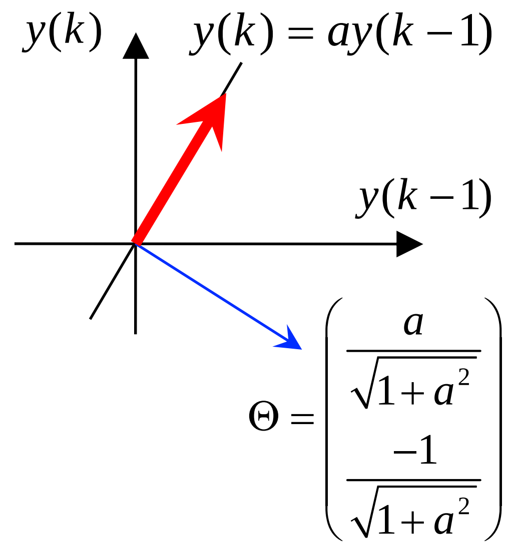
$$y(k) = ay(k-1).$$

Now

$$\Theta = \begin{pmatrix} \dfrac{a}{\sqrt{1+a^2}} \\ \dfrac{-1}{\sqrt{1+a^2}} \end{pmatrix}, \qquad u = \begin{pmatrix} y(k-1) \\ y(k) \end{pmatrix},$$

so that

$$S = \left(\Theta \;\vdots\; \Phi\right) = \left(\begin{array}{c:c} \dfrac{a}{\sqrt{1+a^2}} & \dfrac{1}{\sqrt{1+a^2}} \\ \dfrac{-1}{\sqrt{1+a^2}} & \dfrac{a}{\sqrt{1+a^2}} \end{array}\right).$$



Normalized basis vectors spanning the whole space S:

Constraint

Axis of freedom

- Now extend (defining redundancy among variables)

$$\begin{cases} y(k) = ay(k-1) \\ y(k+1) = ay(k). \end{cases}$$

In this case (without normalization):

$$\Theta' = \begin{pmatrix} a & 0 \\ -1 & a \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad u = \begin{pmatrix} y(k-1) \\ y(k) \\ y(k+1) \end{pmatrix}.$$

The constraint span a two-dimensional subspace in the three-dimensional variable space – one degree of freedom remains

- Orthogonalization of basis Θ' (Gramm-Schmidt procedure):

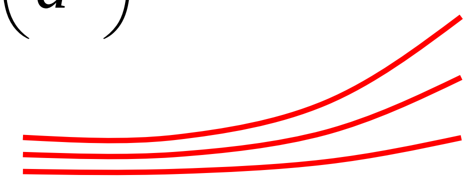
"Deflation"

$$\begin{pmatrix} a & 0 & \vdots & 1 \\ -1 & a & \vdots & 0 \\ 0 & -1 & \vdots & 0 \end{pmatrix} \rightarrow \begin{pmatrix} a & \frac{a^2}{1+a^2} & \vdots & \frac{1}{1+a^2} \\ -1 & \frac{a^3}{1+a^2} & \vdots & \frac{a}{1+a^2} \\ 0 & -1 & \vdots & 0 \end{pmatrix}$$

(Where is this vector from?)

$$\rightarrow \begin{pmatrix} a & \frac{a^2}{1+a^2} & \vdots & \frac{1}{1+a^2+a^4} \\ -1 & \frac{a^3}{1+a^2} & \vdots & \frac{a}{1+a^2+a^4} \\ 0 & -1 & \vdots & \frac{a^2}{1+a^2+a^4} \end{pmatrix} \rightarrow \Phi = \begin{pmatrix} 1 \\ a \\ a^2 \end{pmatrix} \Big/ \sqrt{1+a^2+a^4}$$

"Axis of freedom"
exponential form
(*eigensignal*)



Prototypes

HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Automation and Systems Technology
Cybernetics Group

- Yet another example: Three variables, only one constraint

$$y(k) = a_1 y(k-1) + a_2 y(k-2)$$

or

$$y(k) - a_1 y(k-1) + a_2 y(k-2) = 0$$

$$\downarrow$$

$$a_0 y(k) - a_1 y(k-1) - a_2 y(k-2) = 0 \nearrow \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}^T \begin{pmatrix} y(k) \\ y(k-1) \\ y(k-2) \end{pmatrix} = \Theta^T u(k) = 0$$

- Interpretation of the constraint: *decaying harmonic wave?*
- Interpretations of the degrees of freedom:
  - First: Filtered mean value, level $\phi_1 = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^T$
  - Second: Filtered trend, direction $\phi_2 = \begin{pmatrix} -1 & 0 & 1 \end{pmatrix}^T$

Towards pattern matching

- Use of the model becomes an *associative pattern matching process* against data (exponential curve in the example)
- Linearity – patterns can be freely scaled and added together
- Vector x is the vector of scaling factors = *latent variables* (note that generally Φ is a *matrix*, containing several "axes of freedom" as collected together)

$$x(k) = \left(\Phi^T \Phi\right)^{-1} \Phi^T \cdot u(k)$$

- The *reconstruction* where noise is filtered is given as

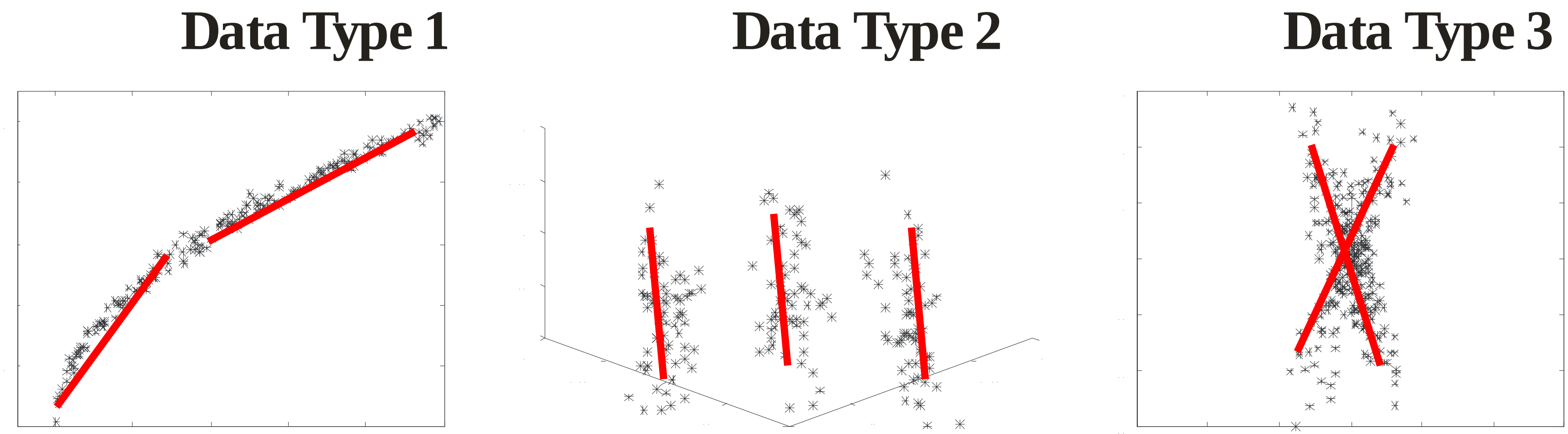
$$\hat{u}(k) = \Phi \cdot x(k)$$

- The more there are internal constraints (feedbacks, etc.), the more efficient the freedoms-oriented approach becomes

If the degrees of freedom are design parameters, compression of the search space can be reached

"Natural data" as sparse-coded features

- Many non-trivial domains can be modeled in terms of *Gaussian mixture models* – mutually exclusive Gaussians
  - *Smooth nonlinearities = linear models around the operating point*
  - *Independent (sparse) components = overlapping data clusters*



Smooth nonlinearities | Clustered data | Independent components

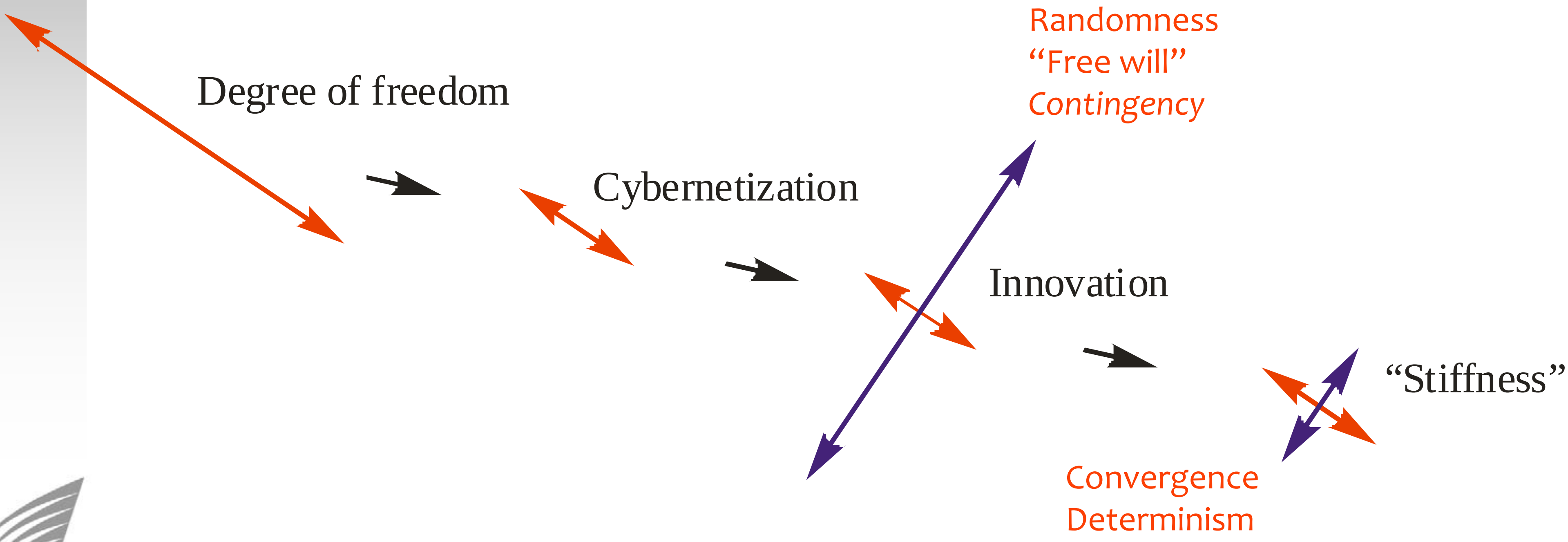
Rules of game vs. *strategies*

- Constraints are the natural way to see the world – partly because language defines connections among entities
- Wittgenstein said that *all logical reasoning only consists of uninteresting tautologies*
- Similarly in all domains, for example in mathematics, the axioms span the space of trivialities – it takes ingenuity to escape the constraints and detect the freedoms
- In some formal environments nontrivial DOF's can be found: For example, if A is a payoff matrix, and x and y are vectors containing choice probabilities of opponents, so that xAy is the average gain, the *degree of freedom reveals the optimal zero-sum game strategy*, A containing the rules (constraints).

Forms of "life" are *constraints-induced emergent strategies* in an environment

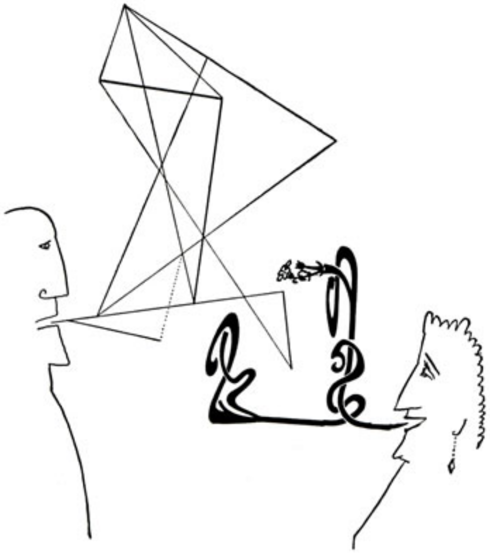
Evolution in a new perspective

- As degrees of freedom become modeled and controlled, becoming new constraints, new innovations are perpetually needed to define new degrees of freedom – otherwise the system dimensions "collapse"

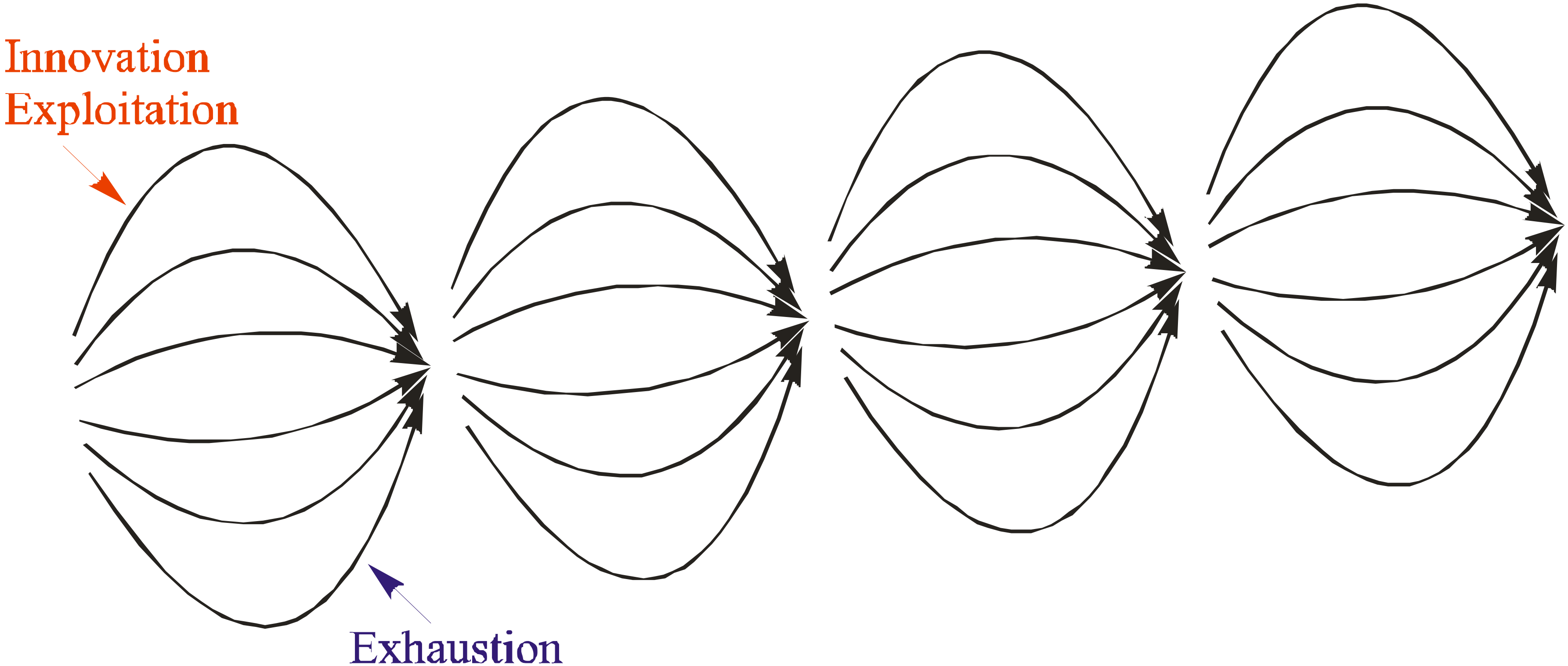


"New Dialectics"

- Always two ends needed to define an *axis*
- For example, study the *intellectual domain*
  - According to Hegel (+ Marx + Kuhn + …): First there is the Thesis – then an Antithesis is proposed
  - The Antithesis determines the "alternative direction", new way to see things
  - When there is enough discussion, and tensions are released, a balance is found: the "correct" location among the ends gets fixed = the Synthesis
  - In other words, the freedom gets controlled …
  - … and changes to rigid "standard science".
- Also, study a *dialog* among (two) persons: To understand each other, mental realms need to get coupled and balance be found



- An evolutionary process is a “saltationistic” alternation of chaotic divergence & deterministic convergence



More colours to Darwin!

- How about *mating*?
- Darwinian theory: everybody wants to be the winner, all others are losers
- But only the winner can marry the winner; what about the others?
- Now: one tries to find a good match, one tries to find a mate that is similar, maximizing ones degrees of freedom
- *Optimality criteria are personal*

Fast differentiation of individuals / species

- In a physical system, there can be room for many in one niche; in higher-level systems, one is enough to exhaust it
- “Goal of life” is then to

  find *your own degrees of freedom* and exploit that variation

- Here a degree of freedom is interpreted as

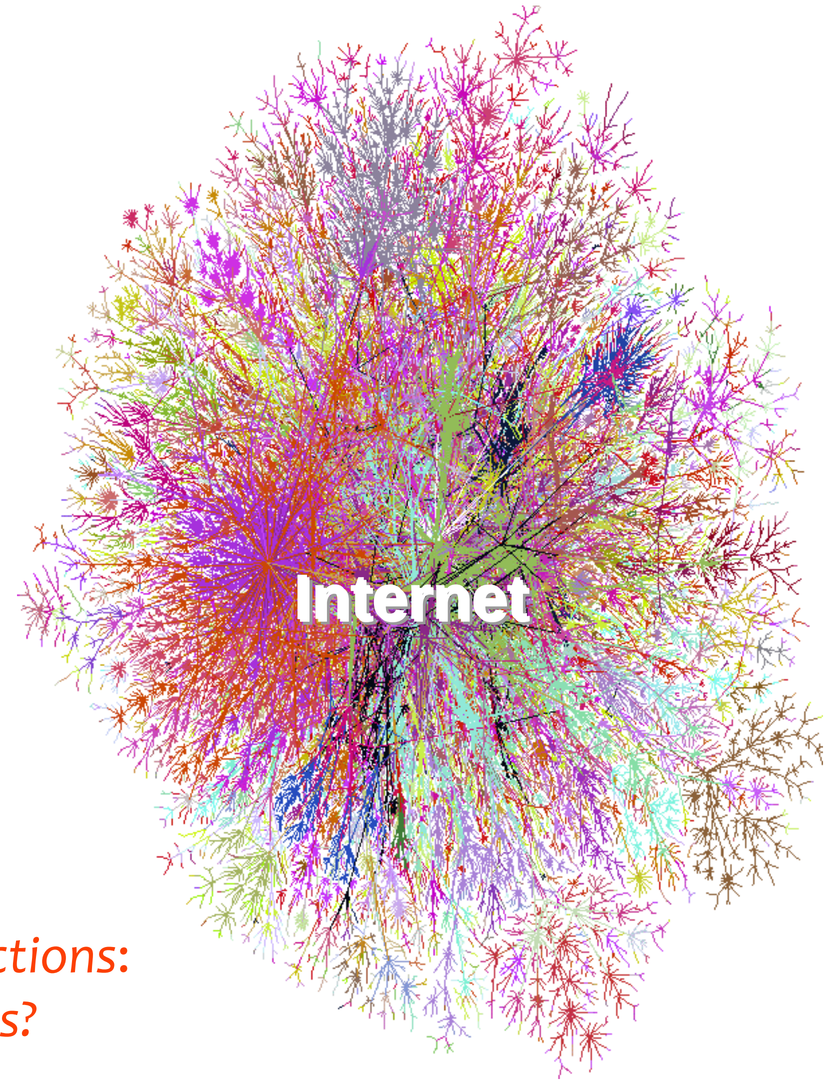
  a way data can be seen as information

- So that

  observations/experiences become relevant/reasonable in your own world!

Further: Complex networks

- Internet, human networks, ...
- Complex networks are perhaps the most potential area of new methodologies
- However, the population thinking does no more hold: How to extend the framework?
- *What does this mean from the point of "practical semiosis"?*



Network structure should reflect *functions*:
How to capture the net of *interactions*?

HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Automation and Systems Technology
Cybernetics Group

Approaches to networks

- Graph theory
  - Connections between nodes are “crisp”
  - However, there is a continuum of interaction effects: The connections in reality are not of “all-or-nothing” type
- Bayesian networks
  - Strong probabilistic theory – *assuming that assumptions hold...*
  - However, the “nodes” in real networks are often not independent of each other: Loops and alternative paths exist in complex networks
- Now: **Neocybernetic framework**
  - Numeric, non-crisp connections, **fully connected**
  - **“Pancausality”** taken as the starting point: It is assumed that, in equilibrium, **all** nodes are causes and **all** are effects – opposite approach!

From emergent level back to the agents?

- As compared to earlier studies, *inverse analysis* now needed
- The network is system as seen from above, afterwards, as an end-effect of many components interacting
- One knows that a dynamic, yielding network is self-controlled – result of a neocybernetic “stupid agents”
- The variables in the internal closed loops are already massively modified by the balancing interactions

How to get back to the lower level, to the agents?

- Start from the beginning – applying neocybernetic modeling principles once more!

Cybernetic intuition #1: Stationarity & statistics

- Abstract away individual actions and realizations of interactions in the network
- Assume that the stationary state has been reached
- What are the statistical properties of the system?

- As advertised by Barabasi etc., the emergent phenomena in the networks are characterized by the *power law*

$$y = z^D$$

"SISO case"

- As observed before, this dependency seems to govern all structures with *fractal* and *self-organized* structure
- This is taken as starting point here – and extended.

Cybernetic intuition #2: Multivariate nature

- Assume there are many variables of power law behavior:

$$\begin{cases} y = c_1 z_1^{D_1} \\ \vdots \\ y = c_n z_n^{D_n} \end{cases}$$

← Parameter c_1 constant with respect to z_1

These can be combined:

$$\frac{y}{\bar{y}} = \left(\frac{z_1}{\bar{z}_1}\right)^{D_1} \cdots \left(\frac{z_n}{\bar{z}_n}\right)^{D_n}$$

- Further, there can exist various such dependencies
- Variables can be rearranged; assume there are (normalized) input variables u and internal variables x (activities):

$$\begin{cases} x_1^{a_{11}} \cdots x_n^{a_{1n}} = u_1^{b_{11}} \cdots u_m^{b_{1m}} \\ \vdots \\ x_1^{a_{n1}} \cdots x_n^{a_{nn}} = u_1^{b_{n1}} \cdots u_m^{b_{nm}} \end{cases}$$

Cybernetic intuition #3: Linearity pursuit

- The same dependencies can be expressed in various ways; the equivalent static set of equations (after taking logarithms) is

$$\begin{cases} a_{11}\log x_1 + \cdots + a_{1n}\log x_n = b_{11}\log u_1 + \cdots + b_{1m}\log u_m \\ \vdots \\ a_{n1}\log x_1 + \cdots + a_{nn}\log x_n = b_{n1}\log u_1 + \cdots + b_{nm}\log u_m \end{cases}$$

or, in matrix form

$$A\log x = B\log u$$

Nonunique representation of dependencies

where the logarithms are calculated elementwise.

There is a close connection to model structures found earlier

Semiosis: $\mathrm{E}\{\log u_i \log u_j\}$ carries information

Cybernetic intuition #4: Dynamicity vs. staticity

- Rather than being a static balance, the variable values result from a dynamic equilibrium among tensions caused by interactions
- The above set of equations is the dynamic balance of the following system (assuming that $-\Gamma A$ is stable)

$$\frac{d\left(\log x\right)}{dt} = -\Gamma A \log x + \Gamma B \log u$$

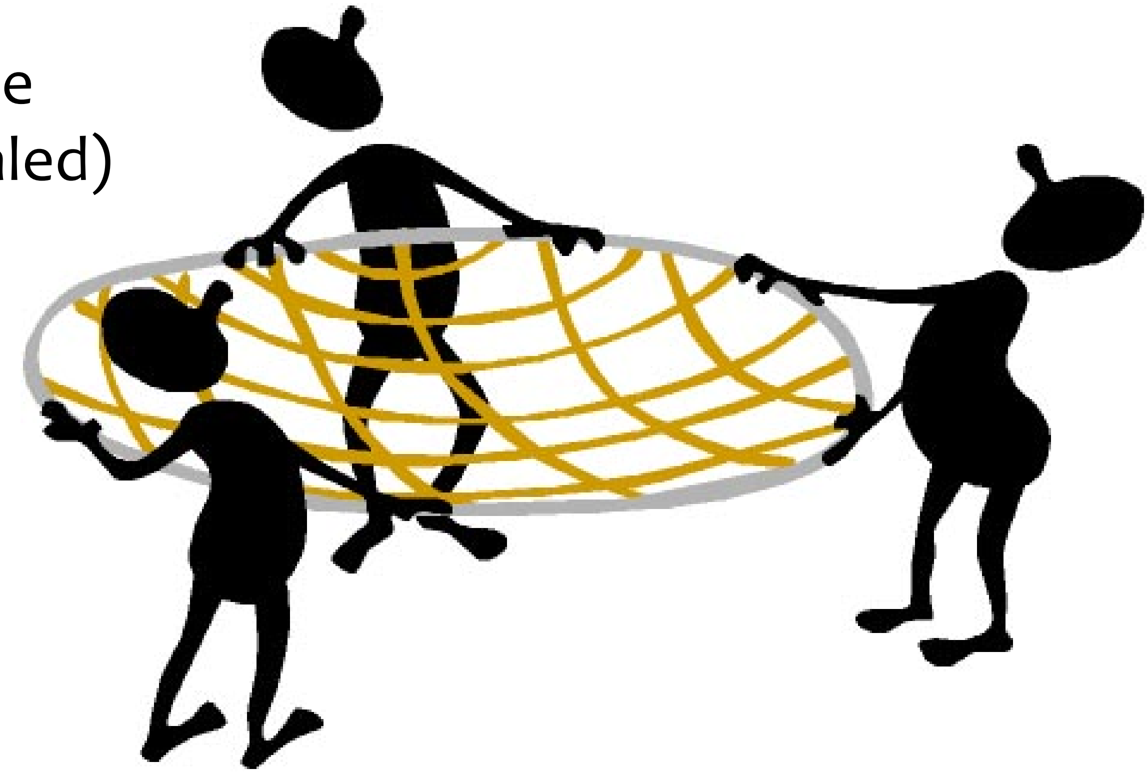
Intuitions available concerning internal interactions in the complex network

... The familiar model again!

- Difference: Now *logarithmic* variables $\log x$ and $\log u$
- The balance based on local interactions can be returned to the neocybernetic framework

We already know how matrices A and B are selected in a system!?

- *Model is multiplicative* rather than *additive* – log variables
- Dynamics is caused by *all components interacting* rather than by individual agents
- The variables have the interpretation of (scaled) probabilities



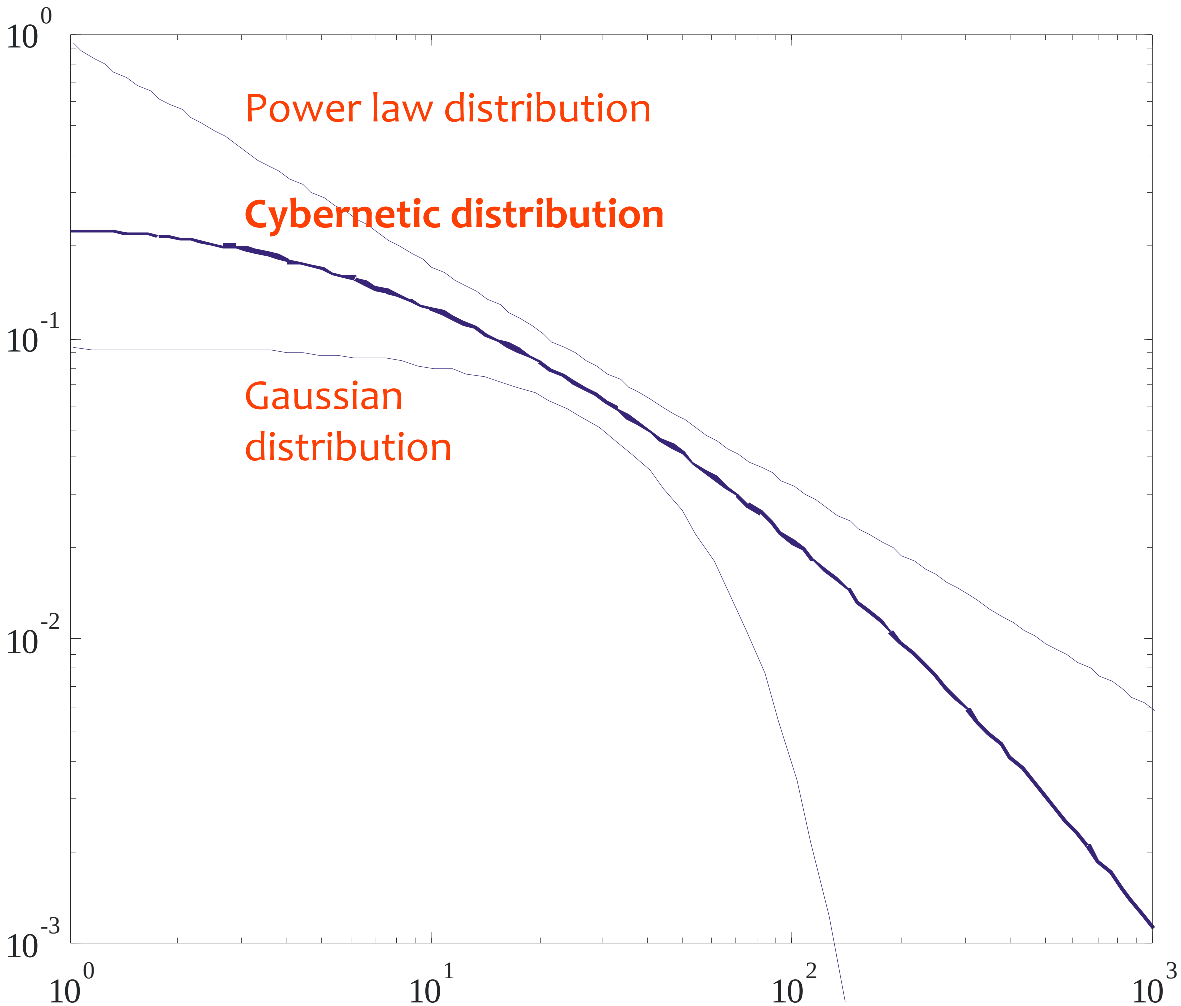
Closer look at distributions

- "Logarithm of a quantity is a sum of many other logarithms"
- Assume the numbers being summed are probabilistic
- If they have the same distribution, the central limit theorem applies: Their sum has approximately normal distribution

$$p\left(\sum_j \log u_j\right) = c' \exp\left(-\left(\sum_j \log u_j - \mu\right)^2 / 2\sigma^2 \right)$$

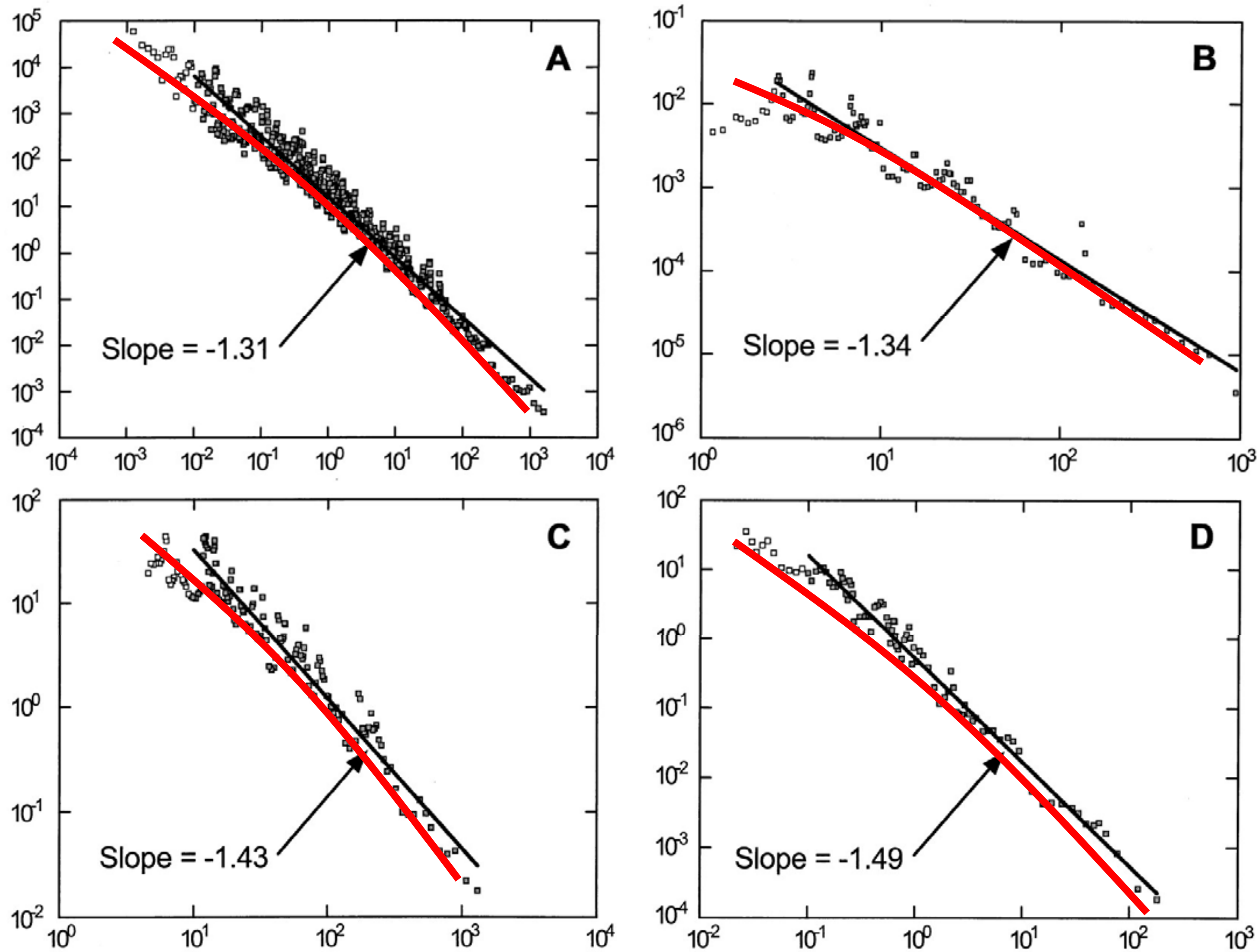
- The sum has *log-normal distribution*: On the log/log scale, the distribution of a "multivariate fractal" quantity behaves quadratically rather than linearly!

$$\log\left(p\left(\sum_j \log u_j\right)\right) = c - \left(\sum_j \log u_j - \mu\right)^2 / 2\sigma^2$$



- Longer “tails” than in normal distribution
- Ends not so emphasized as in power law distribution

- Quadratic curves **are** better than linear!



Extents of forest fires (from Science)

Networks – systems as seen from outside

- Many things are changed, how about adaptation principles?
- As seen from above, the system tries to become better controlled, maximum variation directions being emphasized; optimization can be implemented by local actors familiarly…
- It seems that the *Hebbian law is inverted* now: When $\log x_i$ and $\log u_j$ correlate, their coupling is tuned *down* rather than up, high correlations meaning strong adaptation tension
- On the microscale, this emergent learning rule is manifested in variations becoming equalized + stiffnesses q_i increasing

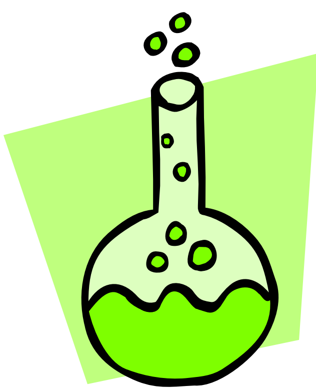
- Opposite views: The environmental variation is (naturally!) minimized as the system-level variations are maximized

Concrete motivation: *Chemical systems*

- Can chemical systems be seen as such “action networks”?!
- Prototypical reaction

$$a_1 \mathrm{A}_1 + \cdots + a_N \mathrm{A}_N \quad \longrightarrow \quad b_1 \mathrm{B}_1 + \cdots + b_M \mathrm{B}_M , \qquad \Delta H$$

- First, a more general formulation for this is needed – the reaction has to be presented in vector form, etc. ...




Intuition #1: Problem formulation

- First augment the reaction:

Enthalpy

$$a_1 C_1 + \cdots + a_m C_m \underset{k_\mathrm{b}}{\overset{k_\mathrm{f}}{\rightleftarrows}} b_1 C_1 + \cdots + b_m C_m, \qquad \Delta H$$

here, there are all chemicals on both sides; a_i and b_j can be zeros. Reactions are assumed reversible (k_b can be zero).

- Collect all chemical concentrations in a single data matrix u; then one can write $\Delta u = r\,\theta$ where r is *reaction rate*, and

$$\Delta u = \begin{pmatrix} \Delta C_1 \\ \vdots \\ \Delta C_m \\ \hline \Delta T \end{pmatrix} \quad \text{and} \quad \theta = \begin{pmatrix} b_1 - a_1 \\ \vdots \\ b_m - a_m \\ \hline c_T \end{pmatrix}$$

- If there are many simultaneous reactions, the changes in the system state can be expressed in the matrix form

$$\Delta u = r^T \Theta$$

- This kind of approach is known as "flux balance analysis" (also compare to *reaction invariants*)
- However, it is difficult to keep track of all fluxes (for example, to master temperatures, the system should be isolated)
- Flux balance captures the *stoichiometric balance* = more or less *formal balance*
- There is no information of whether the reactions actually take place or not – one needs the functional or dynamic balance

Intuition #2: Thermodynamic equilibrium

- Reaction speed k_f is related to probability of unit reaction is related to probability of the constituents to be located near enough each other is related to chemical *concentrations*
- In strong liquids *activities* substitute concentrations
- Reaction speed is also dependent of the temperature (Arrhenius law) – altogether

$$k_\mathrm{f} = c_\mathrm{f}\; e^{-a_T/T}\; C_1^{a_1} \cdot \cdots \cdot C_n^{a_n} \qquad\qquad k_\mathrm{b} = c_\mathrm{b}\; e^{-b_T/T}\; C_1^{b_1} \cdot \cdots \cdot C_n^{b_n}$$

- In equilibrium, the reactions forward and backward are equal, and there holds

$$K = \frac{e^{-b_T/T}}{e^{-a_T/T}} \frac{C_1^{b_1} \cdot \cdots \cdot C_n^{b_n}}{C_1^{a_1} \cdot \cdots \cdot C_n^{a_n}}$$

Intuition #3: Linearity

- Again, the function is purely multiplicative – take logarithms:

$$\log K = (a_T - b_T)1/T + (b_1 - a_1)\log C_1 + \cdots + (b_n - a_n)\log C_n$$

- To get rid of constants and logarithms, it is also possible to differentiate the expression

$$0 = (b_T - a_T)\Delta\left(\frac{1}{T}\right) + (b_1 - a_1)\frac{\Delta C_1}{\bar{C}_1} + \cdots + (b_n - a_n)\frac{\Delta C_n}{\bar{C}_n}$$

where the variables are deviations from the nominal values, divided by those nominal values

- The differentiated model is only locally applicable, valid in the vicinity of the nominal value

- Acidity is logarithmic measure, and its absolute value can be directly included in data:

$$\mathrm{pH} = -\lg C_{\mathrm{H}^+}$$

- Non-balance compounds can be included in data: Assume that G denotes the rate of change, or flow, into / out from the system, so that in balance, for example

$$\frac{\Delta \dot{C}_0}{\bar{\dot{C}}_0} = b_T \Delta\left(\frac{1}{T}\right) + b_1 \frac{\Delta C_1}{\bar{C}_1} + \cdots + b_n \frac{\Delta C_n}{\bar{C}_n}$$

$$u' = \begin{pmatrix} 1/T \\ \hdashline \mathrm{pH} \\ \hdashline \log C_1 \\ \vdots \\ \log C_n \\ \hdashline \log G_1 \\ \vdots \\ \log G_1 \end{pmatrix}$$

$$u = \begin{pmatrix} \Delta\left(\frac{1}{T}\right) \\ \hdashline \Delta \mathrm{pH} \\ \hdashline \frac{\Delta C_1}{\bar{C}_1} \\ \vdots \\ \frac{\Delta C_n}{\bar{C}_n} \\ \hdashline \frac{\Delta G_1}{\bar{G}_1} \\ \vdots \\ \frac{\Delta G_m}{\bar{G}_m} \end{pmatrix}$$

Relative change in flow

Search for “equilibrium dissipation”!

Intuition #4: Multiple reactions

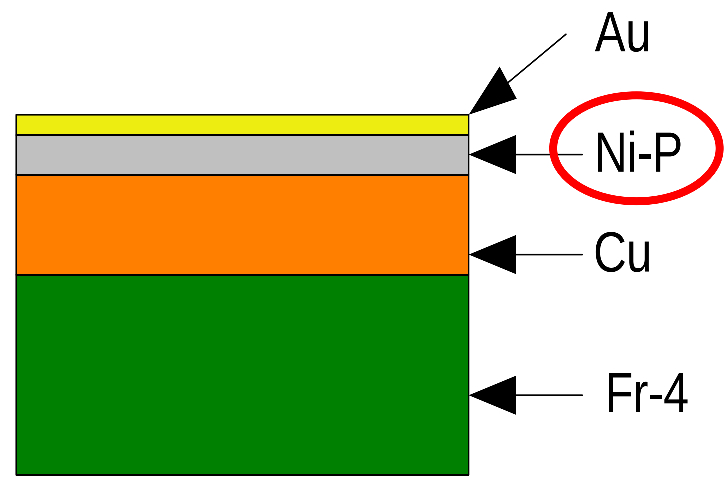
- Now, when the reaction parameters are collected in vector ϕ, there holds

  $$0 = \Phi^T \; u$$

  This holds also if there exist *simultaneous reactions*, so that Φ is a matrix
- Compare to flux balance analysis: Now one only needs to study levels (causing "chemical pressures"), not changes
- This is essential in complex chemical systems: The levels can better be controlled than the individual reactions
- Linear emergent models of balances are not only models for the data but *system models*

Example: Printed Circuit Board manufacturing

- In PCB manufacturing extra nickel layer is used
  - as an oxidation barrier between copper and gold
  - to bring wear resistance to the boards
- Crucial parameters:
  - Nickel layer thickness 4.5 μm
  - Phosphorous content 8.5 wt.%
    - corrosion resistance
    - solderability
- How to supervise and control these parameters?
  - No on-line measurements available
  - Time delay of laboratory measurement considerable



Getting into details ...

- Electrochemical reaction mechanism (one out of seven models proposed!)
  - Anodic reaction:

$$H_2PO_2^- + H_2O \rightleftarrows H_2PO_3^- + 2H^+ + 2e^-$$

  - Cathodic reactions:

$$H_2PO_2^- + 2H^+ + e^- \rightleftarrows P + 2H_2O$$

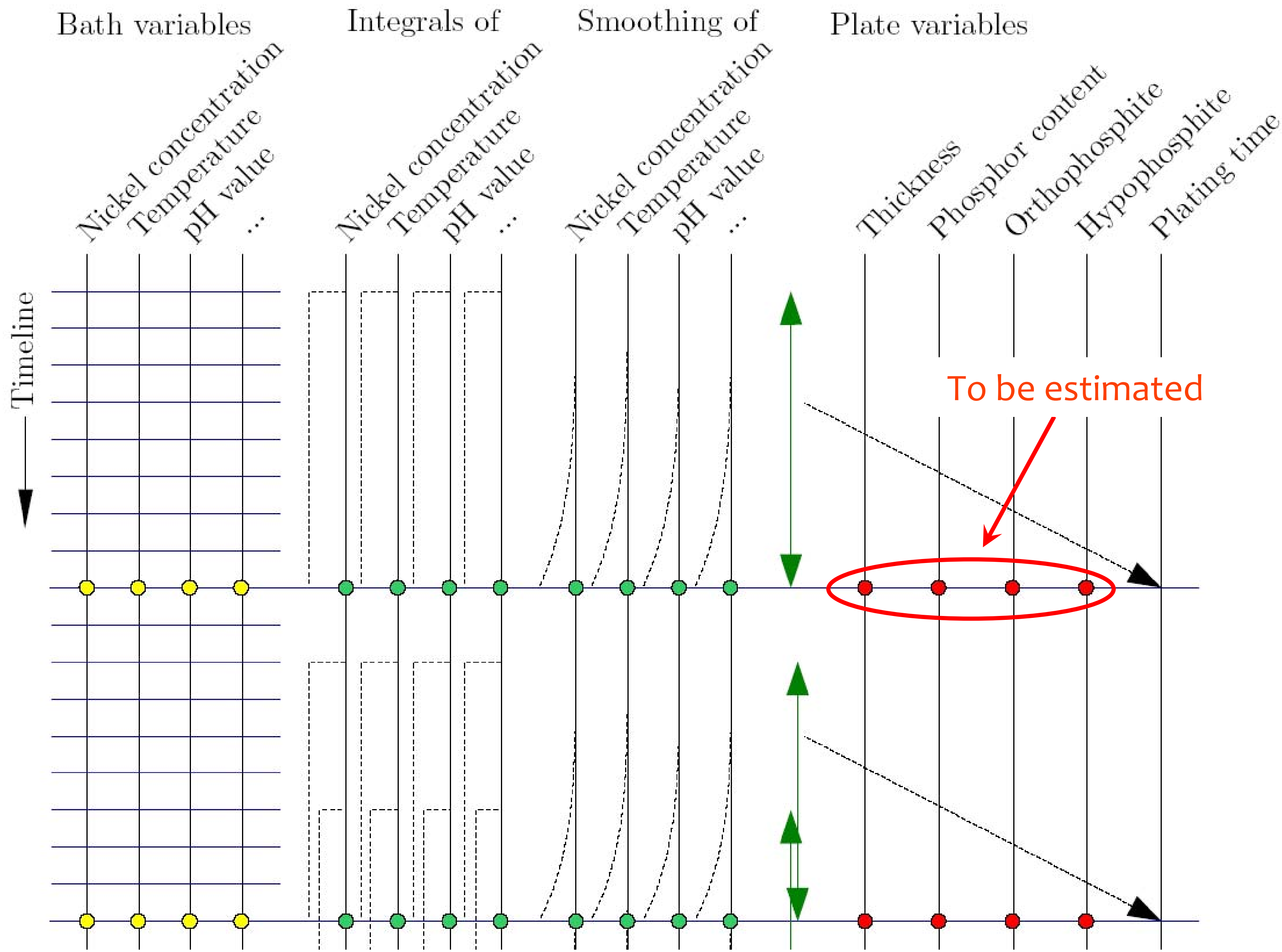
$$2H^+ + 2e^- \rightleftarrows H_2$$

$$Ni^{2+} + 2e^- \rightleftarrows Ni$$

- For each of these reactions the current densities in different locations can be calculated from Buttler-Wolmer equation

$$\mathrm{i_n} = i_{0n}\mu_n \left\{exp(\nu\alpha_{an}p_n k\eta_n) - exp(-\nu\alpha_{an}p_n k\eta_n)\right\}$$

To some extent unknown chemistry (company secrets)!



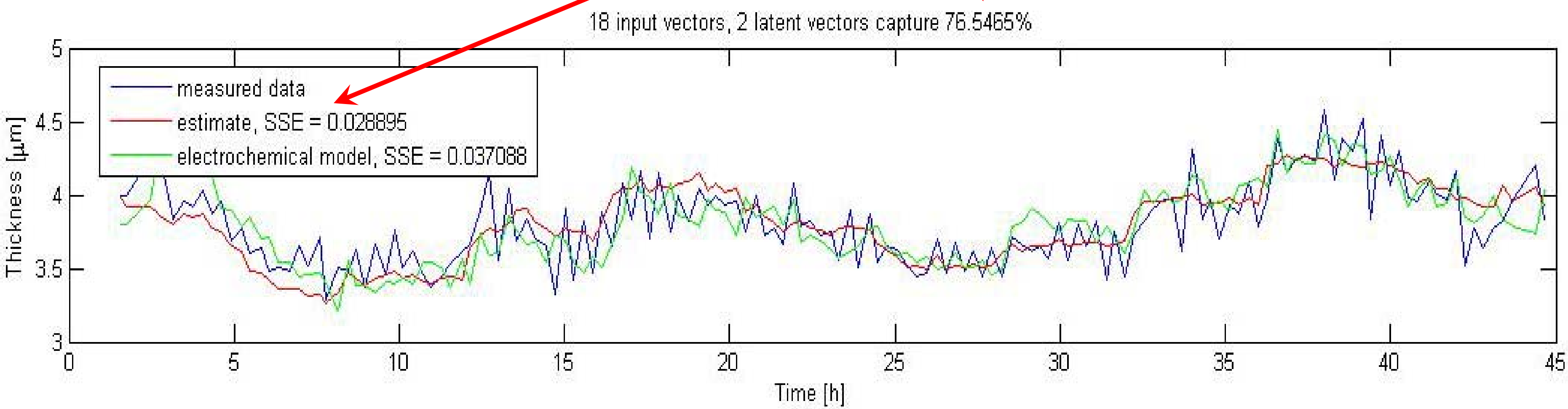
Linearity – integrals traverse through the model

PCR results for layer thickness

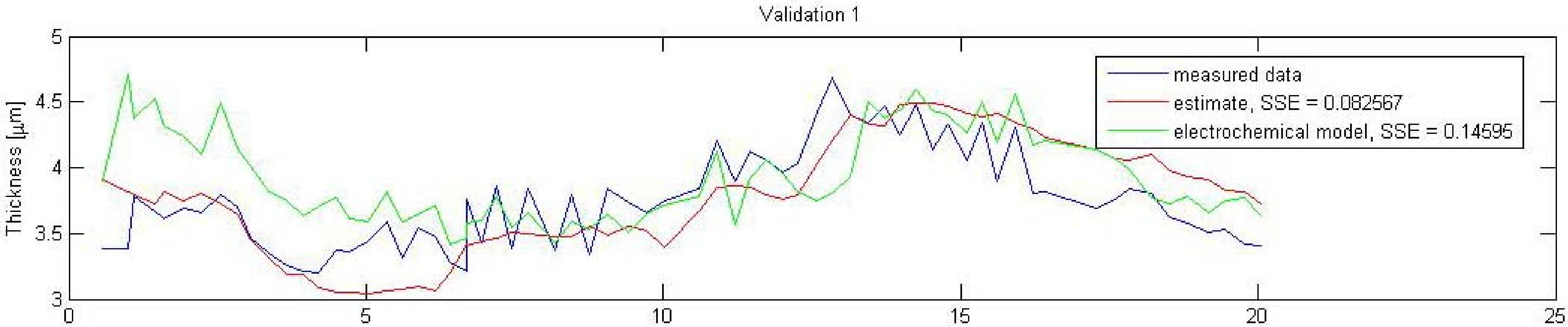
- Available output data parts
  - 1 estimation set
  - 2 validation sets
  - Only two latent variables applied

- Balance assumed
- Logarithmic variables
- Linear PCR model

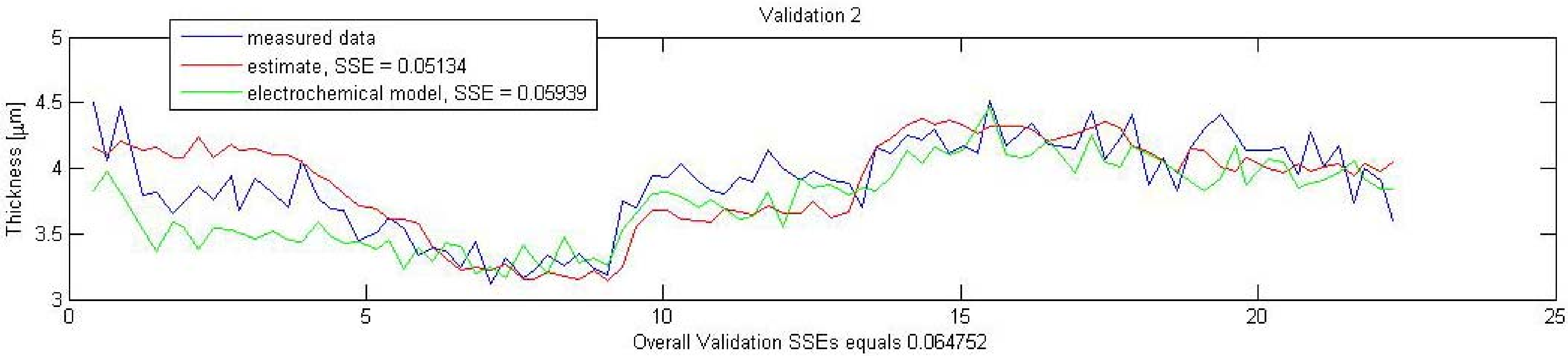
- Estimation:



Validation

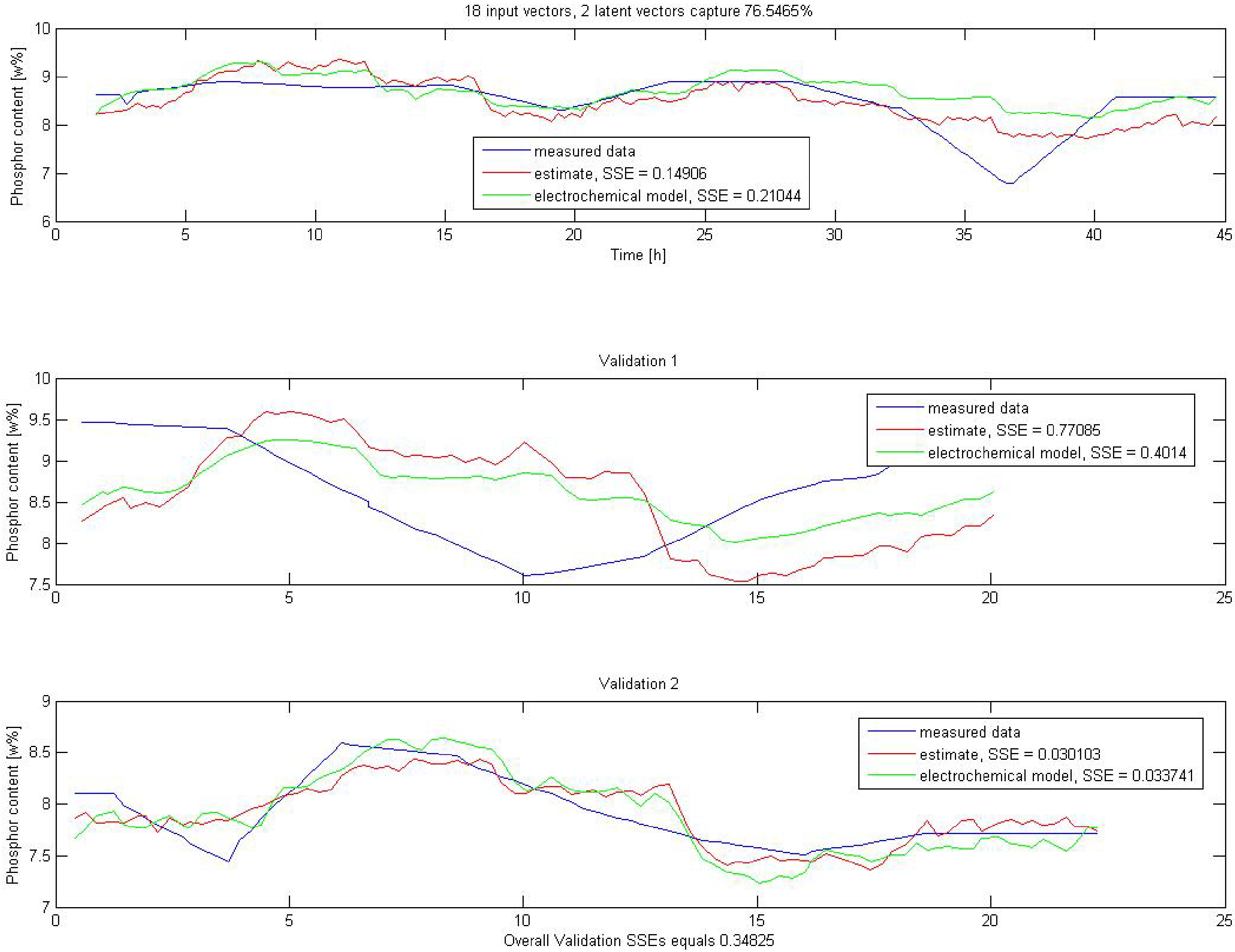


Validation 1
Thickness [μm]
measured data
estimate, SSE = 0.082567
electrochemical model, SSE = 0.14595
0
5
10
15
20
25
3
3.5
4
4.5
5



Validation 2
Thickness [μm]
measured data
estimate, SSE = 0.05134
electrochemical model, SSE = 0.05939
0
5
10
15
20
25
3
3.5
4
4.5
5
Overall Validation SSEs equals 0.064752

PCR results for phosphor



Too little information available!

HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Automation and Systems Technology
Cybernetics Group

- It seems that
  - the neocybernetic model produces accurate estimation/validation results, even better than the electrochemical model
  - it provides an insight into significance of variables
- From the practical point of view
  - the model is easy to implement and maintain, it improves production quality and lowers measurement expenses (?)
  - not all reactions need to be known – ignorance of variables does not matter as long as the system remains stable, one can concentrate on the freedoms
  - the still unbounded degrees of freedom can be regulated – "Superorganisms" can be constructed by external explicit feedbacks!

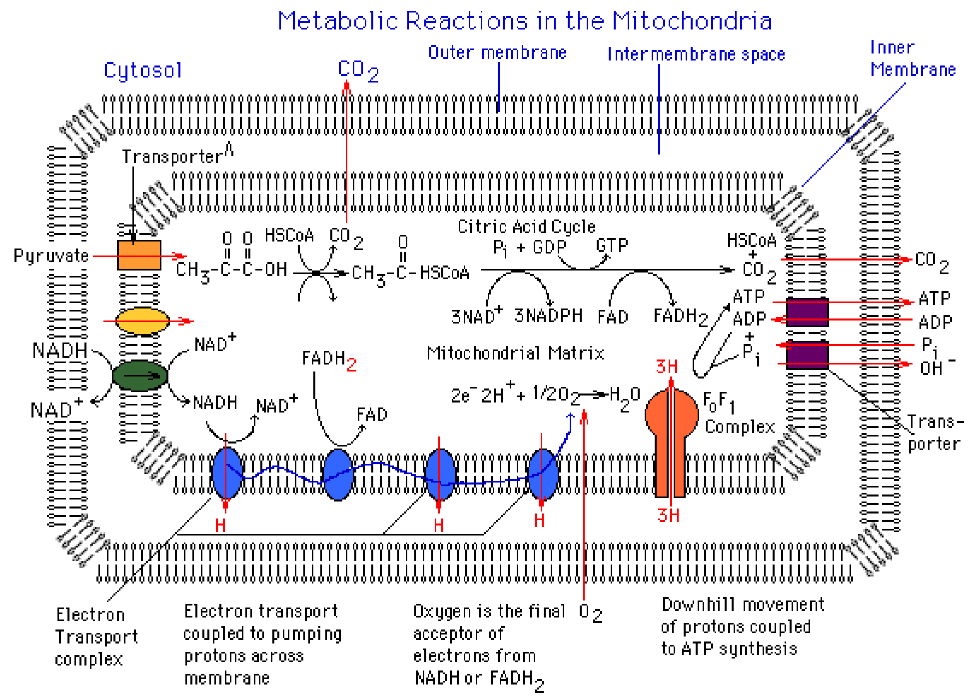
... Next, a more ambitious case ...

Cell level #1: Metabolic system

- Constraints = Balance equations
- DOF's = Metabolic behaviors
- Anthropocentric interpretations: Nutrient, waste product

Very different from flux balance analysis

- When complexity cumulates, the balance reactions start looking goal-oriented, pre-planned, and "clever"
- For example, scarcity of some chemical changes the balance appropriately



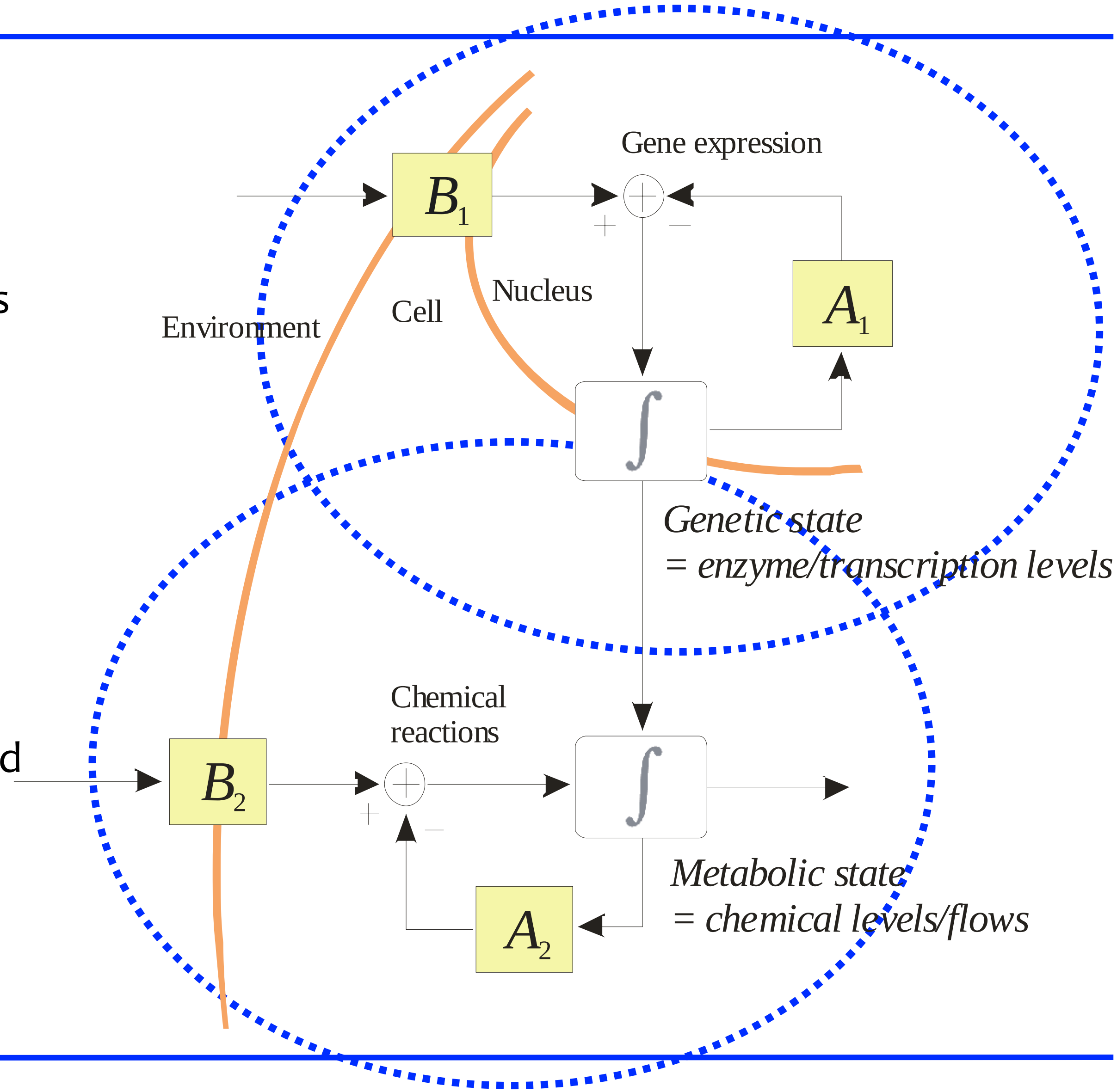
Cell level #2: Genetic system

- Active genes determine the enzymes (proteins) available = the reactions actually taking place in the cell
- Special enzymes act as *transcription factors*, activating (or inhibiting) other genes
- The gene activation relationships constitute a causal network
- Traditional graphs are too "qualitative" (all or nothing), and networks become too dense and intangible

- Alternative approach again: Assume "pancausality"
- In equilibrium, causal "forces" balance each other even though the circumstances differ
- Static model rather than sequential, dynamic ones

… Two cybernetic levels of cell processes

- Appropriate abstractions:
  - Two successive process levels of “generalized diffusion”
  - Metabolic processes fast, genetic ones slow
  - In both cases, forget about the sequential nature
  - Emergent models based on latent (logarithmic) variables
  - Both levels – same approaches!?



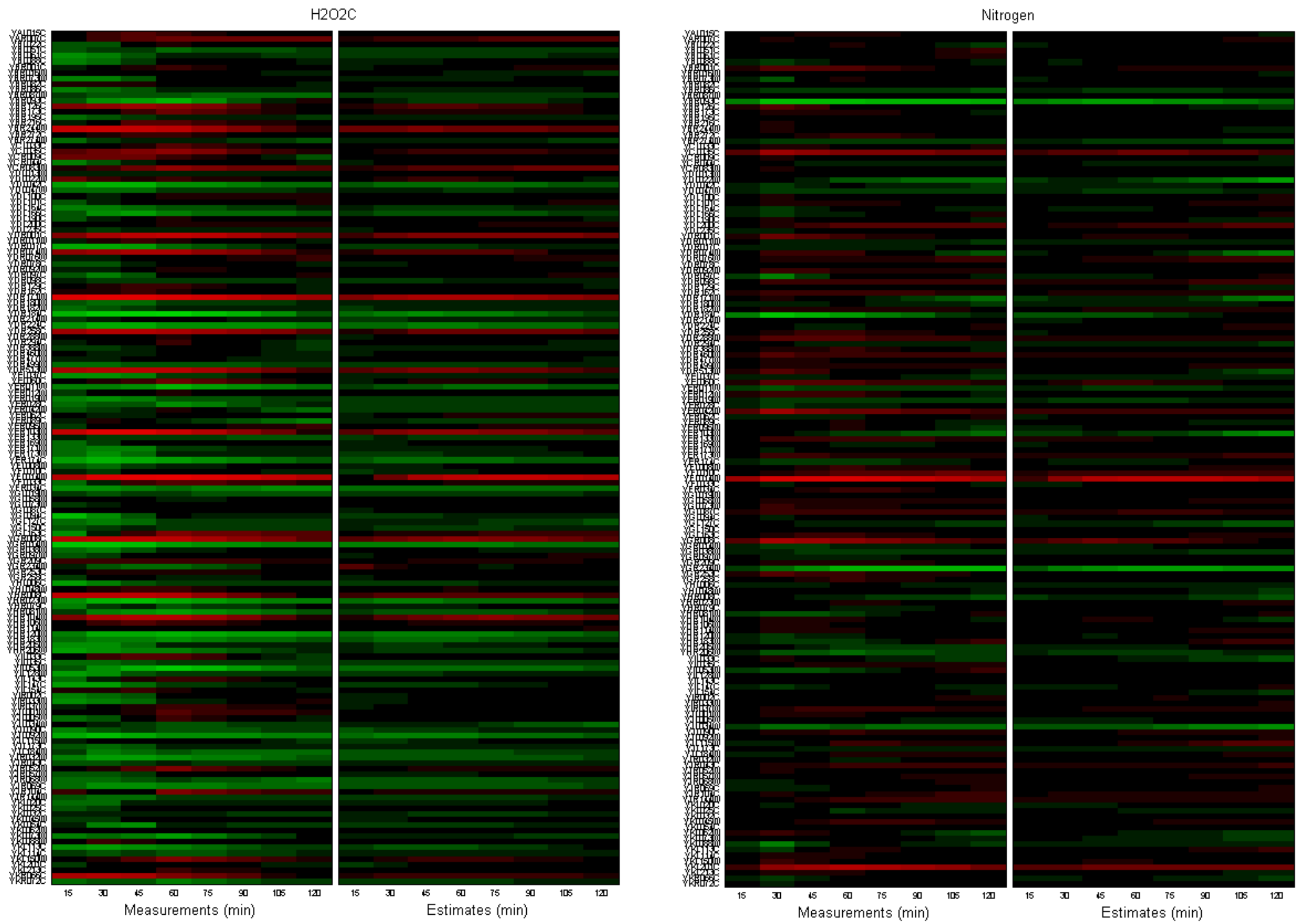
***n* = 4** only!

dim(u) = 10

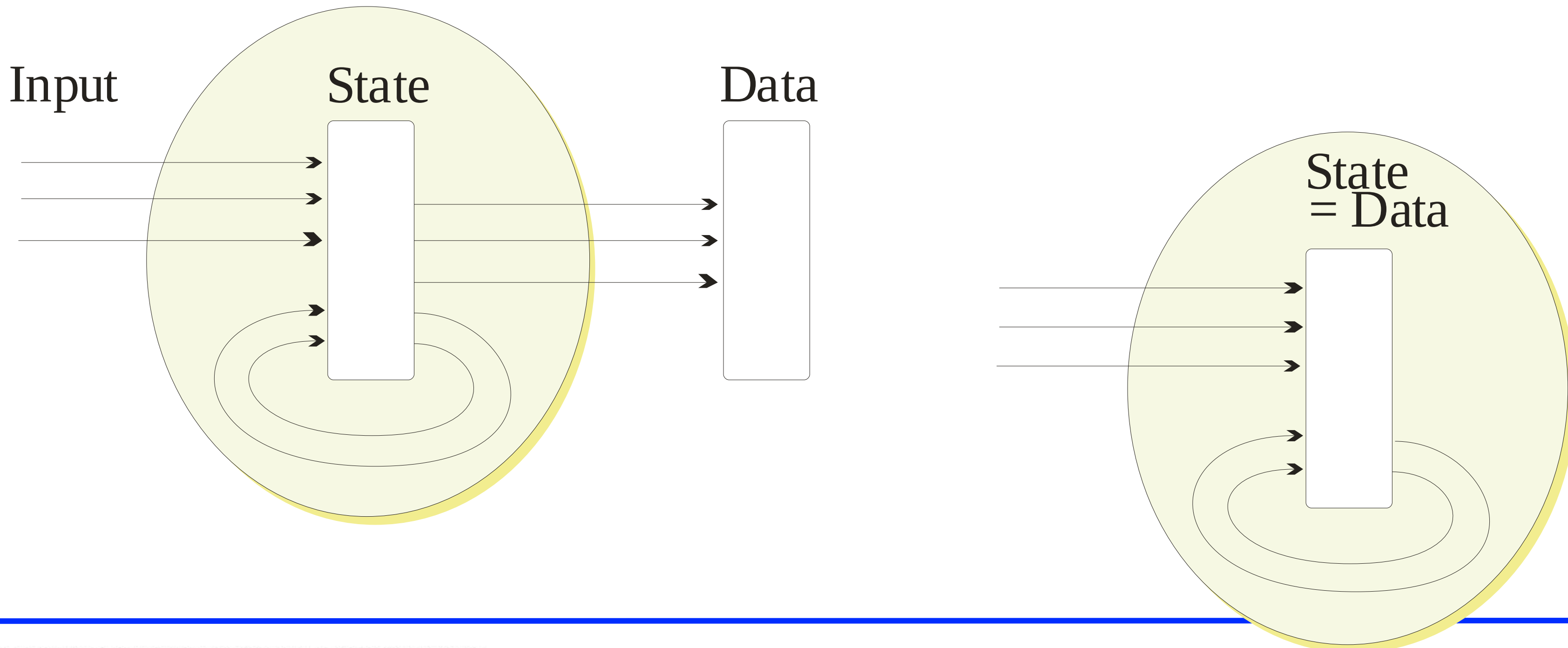
dim(y) = 4135

- Step tests:

  254 “stress genes” shown



- A system model can be applied also for design and control: The observed correlations are also causalities, changing a variable value affects the system, making the other variables search a new balance



Conclusion

- Freedoms define the directions where variations "make a difference that makes a difference" (G. Bateson)

- Traditionally: constraints – world as it is / has to be
- Cybernetically: freedoms – "world as it could be"

- One goes from info transfer to negotiation (feedforward vs. feedbacks); from hard controls to persuasion (imposed vs. natural dynamics)
- In applications, the role of human changes from implementing controls to acting as a *catalyst*

- *One is near practical applications of cybernetics here...*