Hands-on Neural Process Modeling through Dynamic Field Theory (DFT)

Minseok Kang, Raul Grieben, Gregor Schöner
Institute of Neural Computation
Ruhr-University Bochum, Germany
dynamicfieldtheory.org

What is DFT?

- =neural process account of thinking and acting…
  - that may actually bring about thoughts and action
  - that explains the laws of thinking and acting
  - while respecting neural principles

Central hypothesis of DFT: embodiment

- thinking and acting are brought about by the embodied and situated brain that is shaped by evolution and development

Central hypothesis of DFT: embodiment

- thinking and acting are brought about by the embodied and situated brain that is shaped by evolution and development
- => neural processes with continuous state, continuous time, potential coupling to the sensory and motor surfaces, and stability

Autonomy

- the neural principles of DFT ~ connectionism
- but: conceptually, most current neural network accounts are input driven
- while thought and action are driven by the inner state of the mind/brain = autonomous neural processing
- => DFT must address how inner states arise, persist, and evolve in time

Integration

- (embodied) cognition entails many different processes probed in a large variety of paradigms
- DFT is aimed to provide a single theoretical language to understand all these processes and how they interrelate

Dynamic Field Theory (DFT)

- 1 Space: inner states are localized activation patterns in low-dimensional features spaces
- 2 Time: autonomy derives from neural dynamics, attractors and dynamic instabilities
- 3 Coupling: cognition emerges from dynamic coupling across low-dimensional features spaces
- 4 Integration: in DFT architectures

1 Space

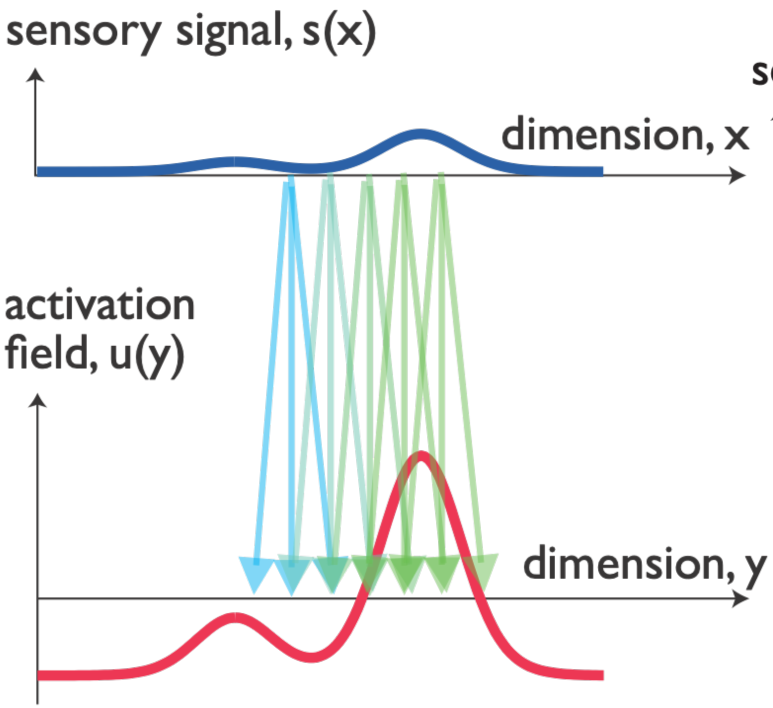
- activation in neural populations carries functional meaning
- activation: $u(x, t)$ where x spans low-dimensional spaces

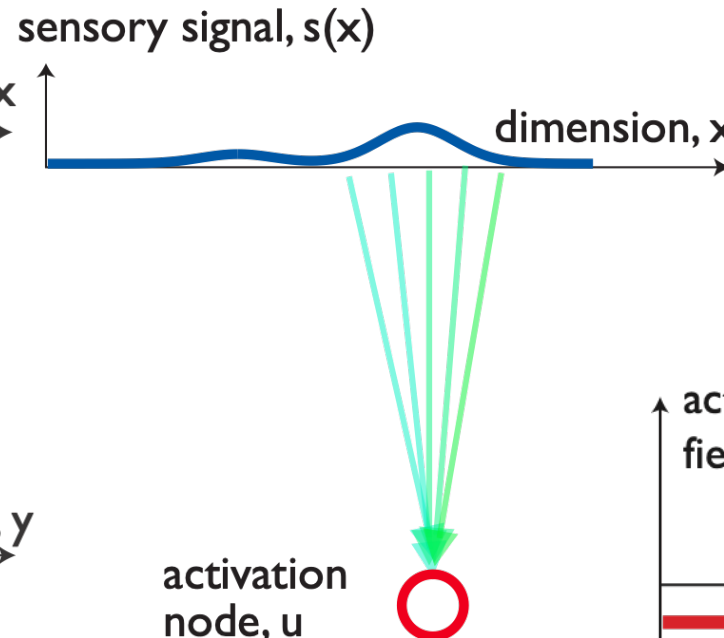
[Schöner TopiCS 2019]

Feature spaces from forward connectivity

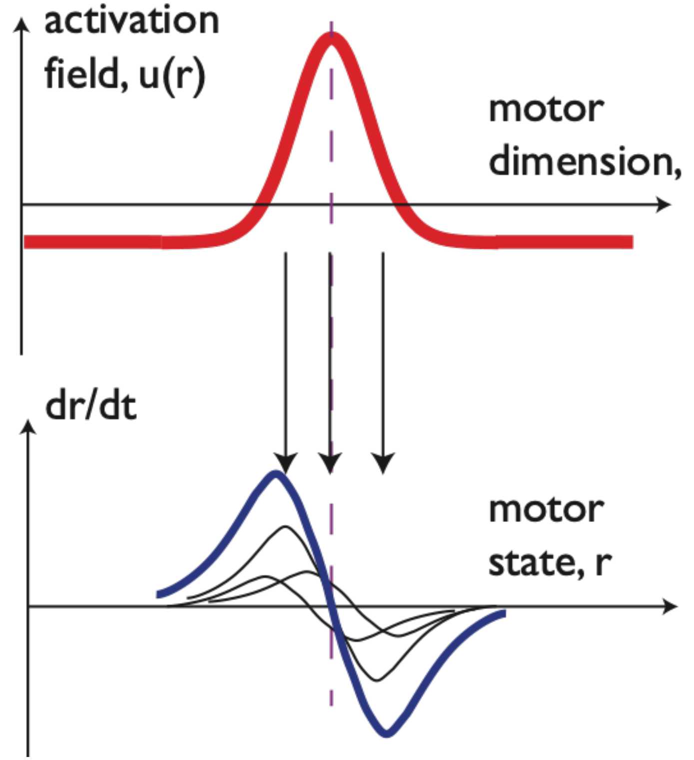
from sensory surfaces

to motor surfaces





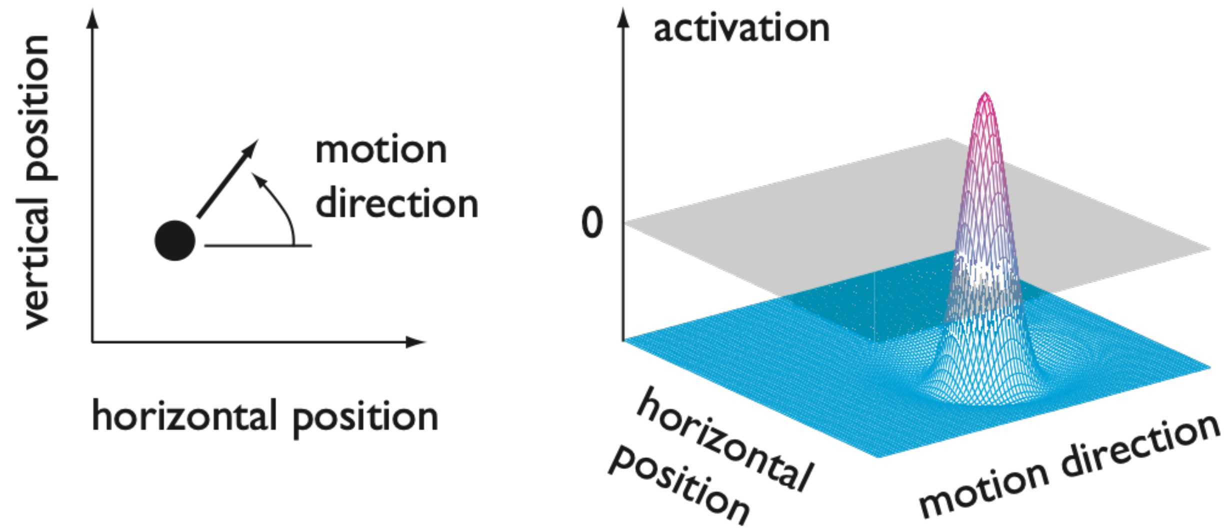
motor space

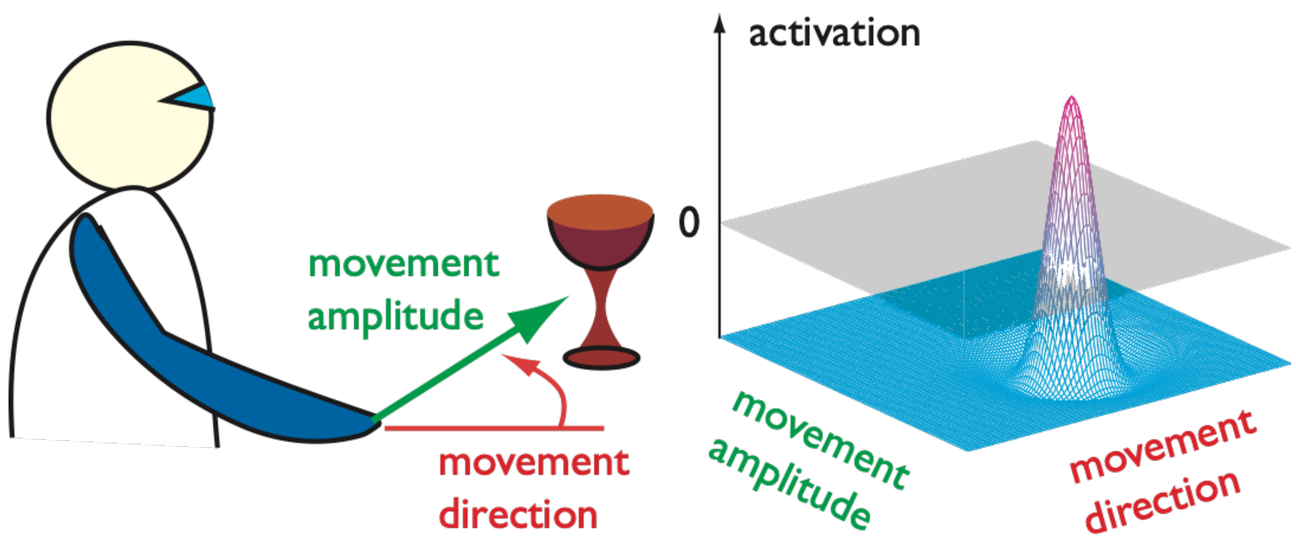


feature space

concepts

Hypothesis: mental states are activation patterns localized in low-dimensional feature spaces





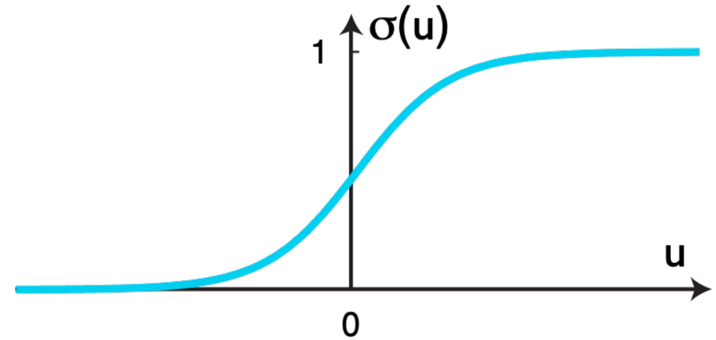
- [~ *Gärdenfors*]

2 Time

membrane dynamics — resting level — input

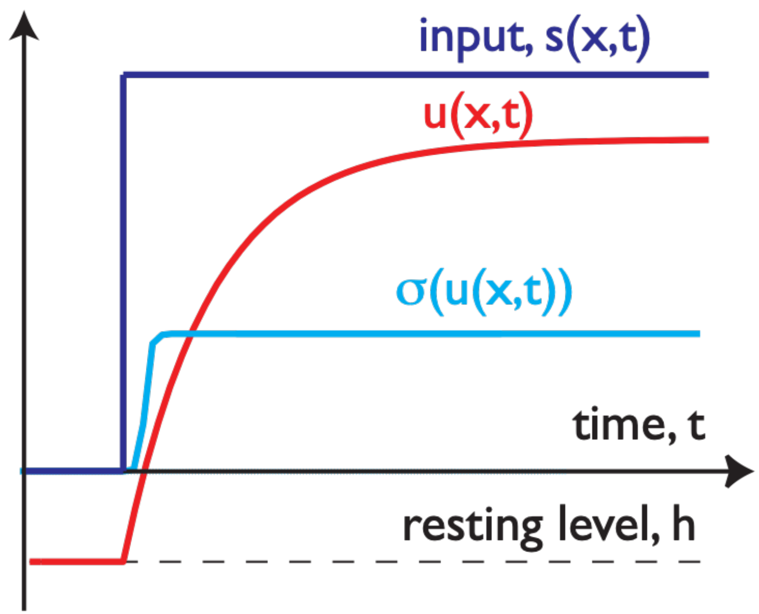
$$\tau\dot{u}(x,t) = -u(x,t) + h + s(x,t)$$

- Neural dynamics: continuous activation evolves in continuous time toward attractors



- [~ *Grossberg, Wilson-Cowan, Amari*]

- but: so far only transmits and smooths input time courses



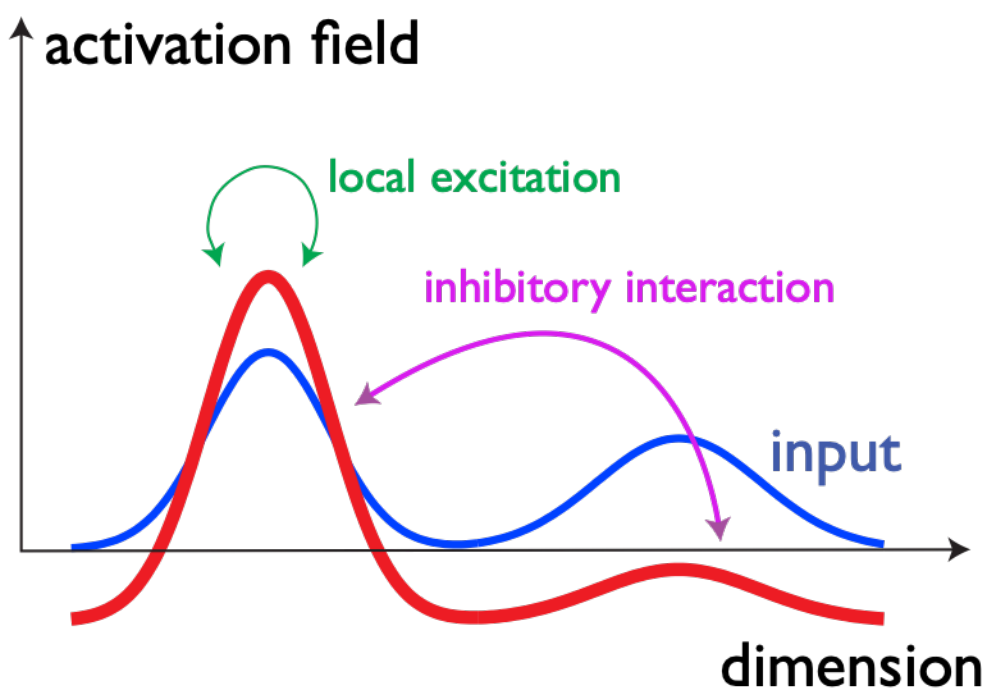
…beyond input driven activation

$$\tau\dot{u}(x,t) = -u(x,t) + h + s(x,t)$$

$$+\int w(x-x')\sigma(u(x',t))dx'$$

interaction

- strong recurrent connectivity within populations
- excitatory: *w>0* for neighbors in space
- inhibitory: *w<0* over larger spatial distance



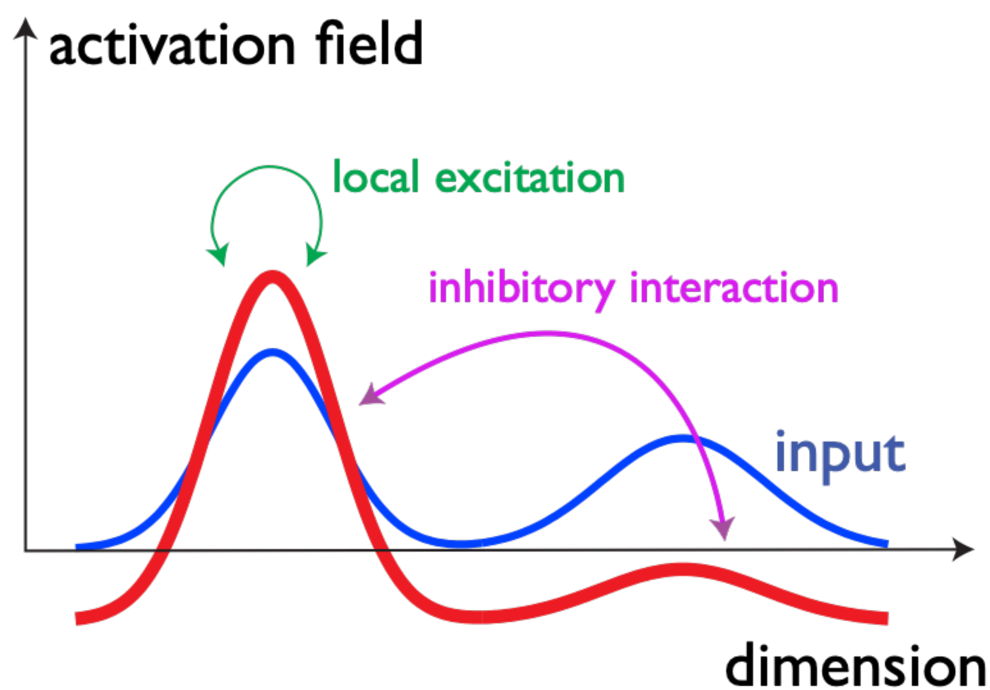
=> localized activation peaks are stable states = attractors

$$\tau \dot{u}(x,t) = -u(x,t) + h + s(x,t)$$

$$+ \int w(x-x')\sigma(u(x',t))dx'$$

interaction

- strong recurrent connectivity within populations
- excitatory: *w>0* for neighbors in space
- inhibitory: *w<0* over larger spatial distance

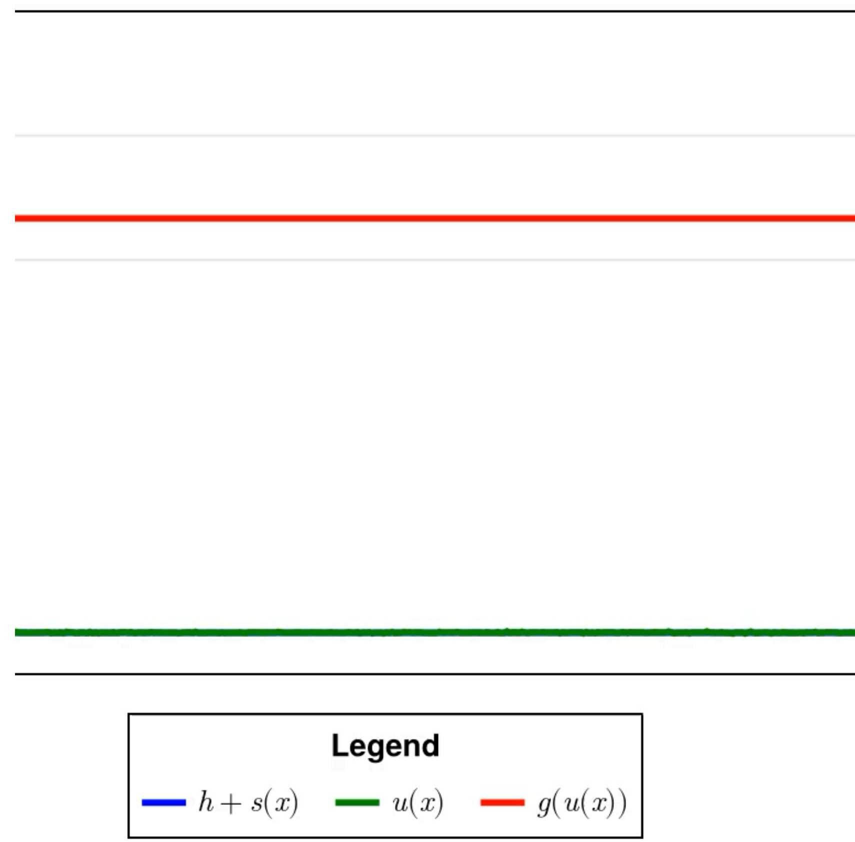


How do peaks arise and disappear?

- detection: a peak arises
- reverse detection: a peak disappears

- detection instability of sub-threshold state => peak
- peak persists below detection instability => bistable
- reverse detection instability of peak => sub-threshold

bistable:
same input
two stable
states

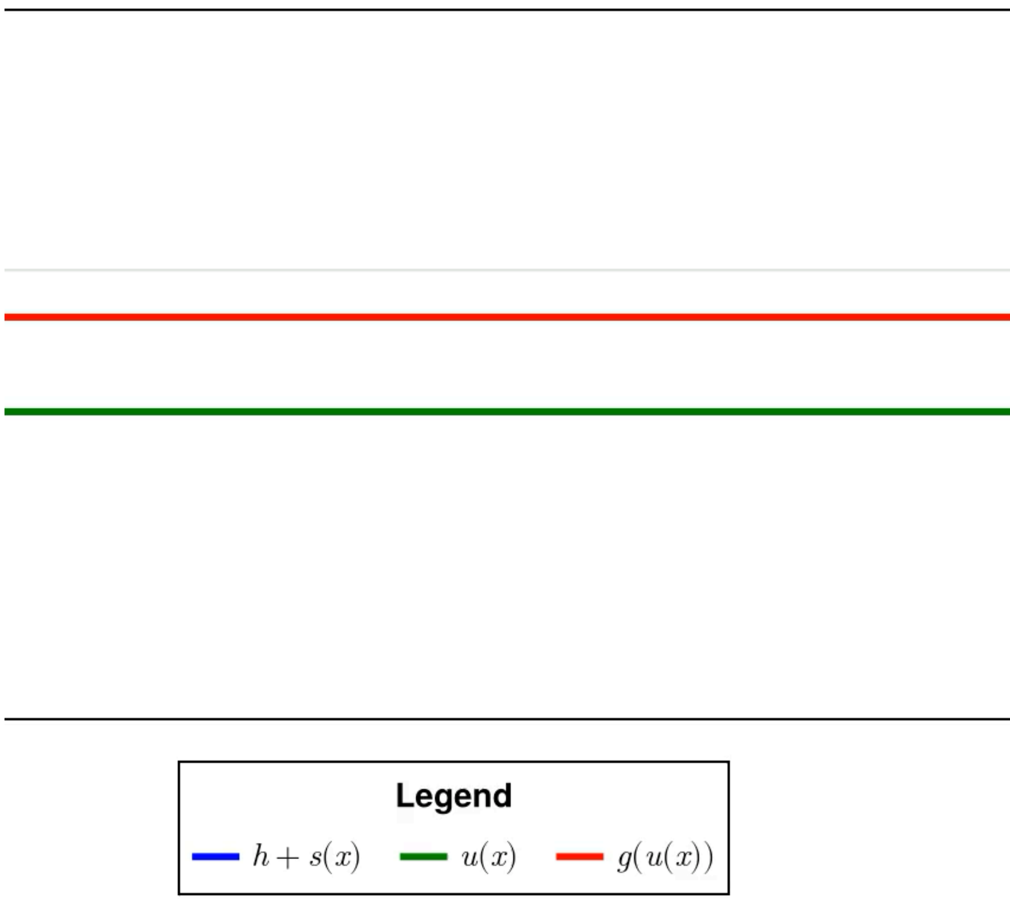


Autonomy from attractors and their instabilities

- detection instability
- reverse detection instability
- sustained activation
- selection
- selection instability
- boost driven detection/selection
- match events and sequences

How does a peak NOT disappear?

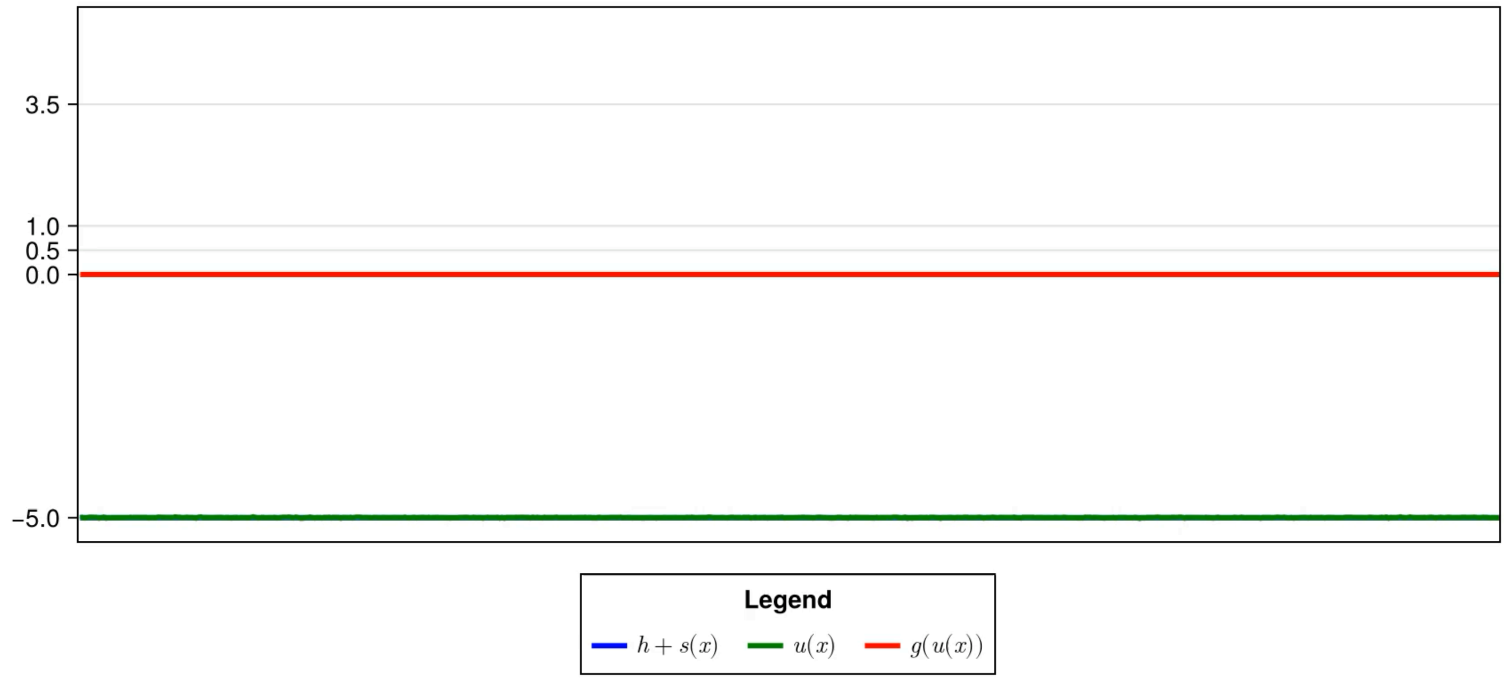
- sustained activation
- emerges when reverse detection instability is not reached as input is removed
- ~foundation of working memory

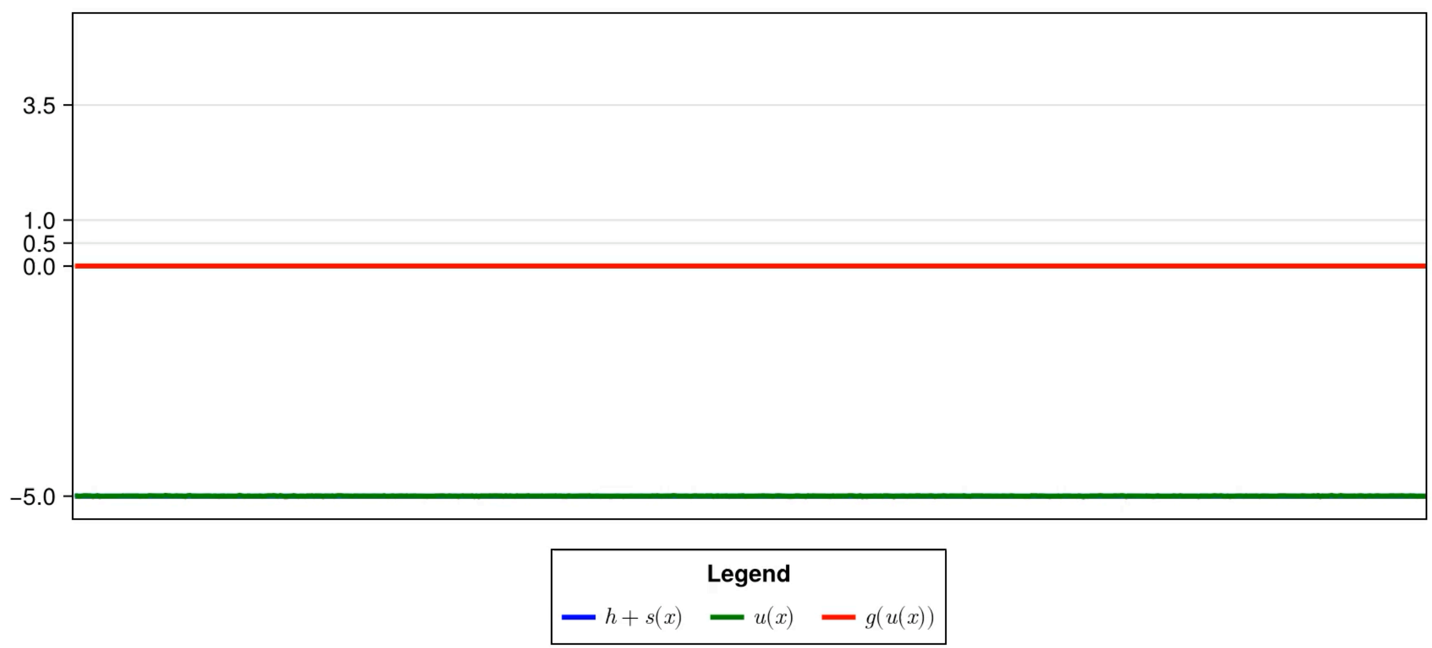


Which peaks arises?

- selection: one peak arises over input at one location rather than input at other locations

- selection
- selection instability





Which peaks arises?

- selection: one peak arises over input at one location rather than input at other locations
- this at any level: gaze, attention, action, thoughts.... => at any moment in time, new mental states compete for activation as thought and behavior evolve in time...
- => selection is foundational for understanding the neural process of decision making

Limit case: detection and selection without a cue

- the opposite limit of input-driven selection… have a set of pre-activated locations and receive a pure “go”-signal to decide
- ~ “guessing the corner” in a soccer goalie metaphor

- detection and selection induced by homogeneous boost
- => peak forms that amplifies small inhomogeneities

- detection and selection induced by homogeneous boost
- => categories emerge from continuous spaces

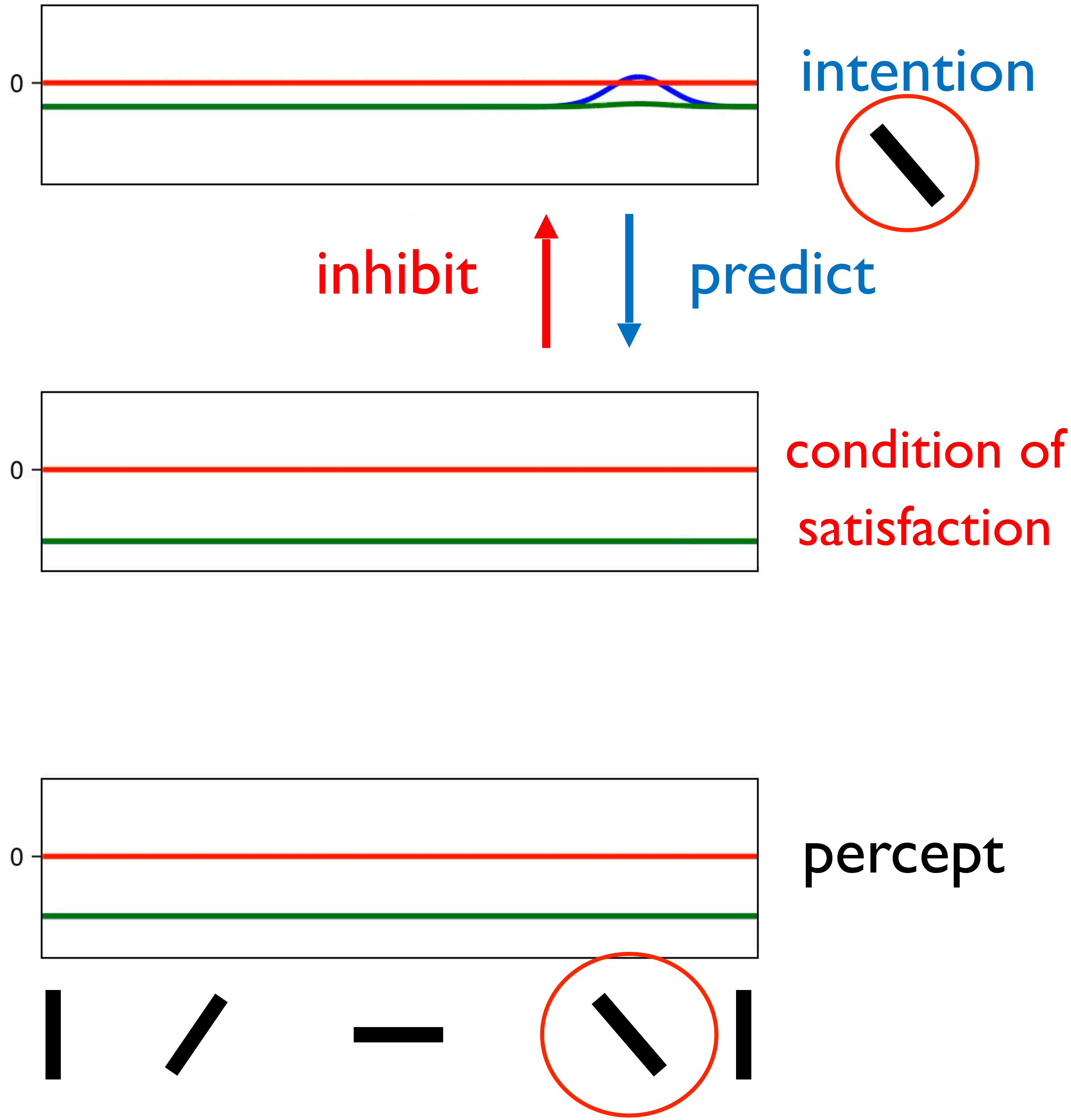
How do events arise?

- peaks arising and disappearing at discrete moments in time (events) in response to time continuous changes
- ~ foundational for autonomous neural processing

- task: match detection
- => event

stimulus

- => basis of autonomous sequence generation



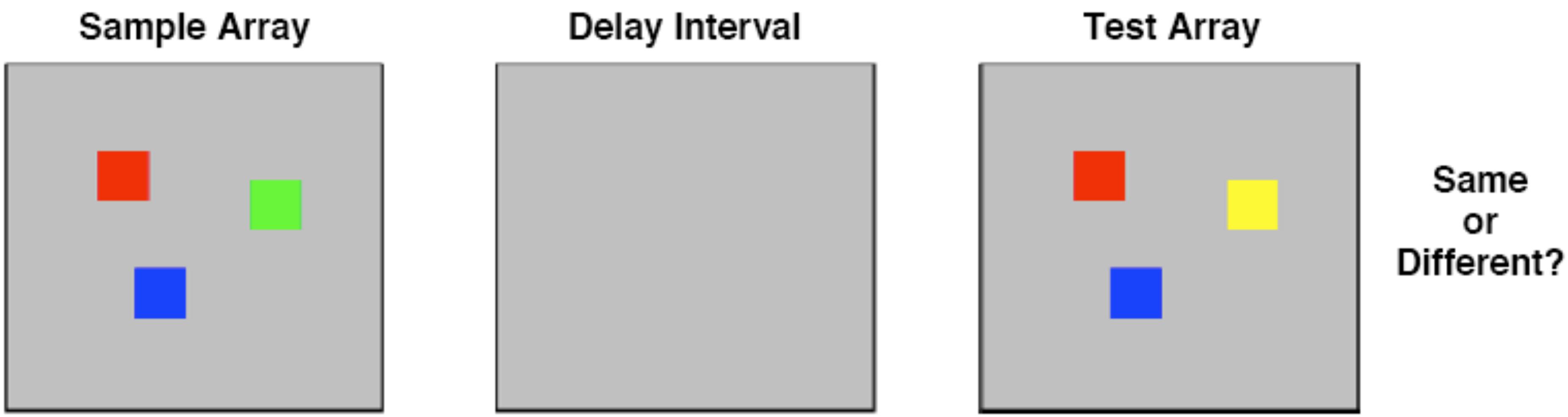
Autonomy from attractors and their instabilities

- detection instability
- reverse detection instability
- sustained activation
- selection
- selection instability
- boost driven detection/selection
- match events and sequences

Empirical evidence

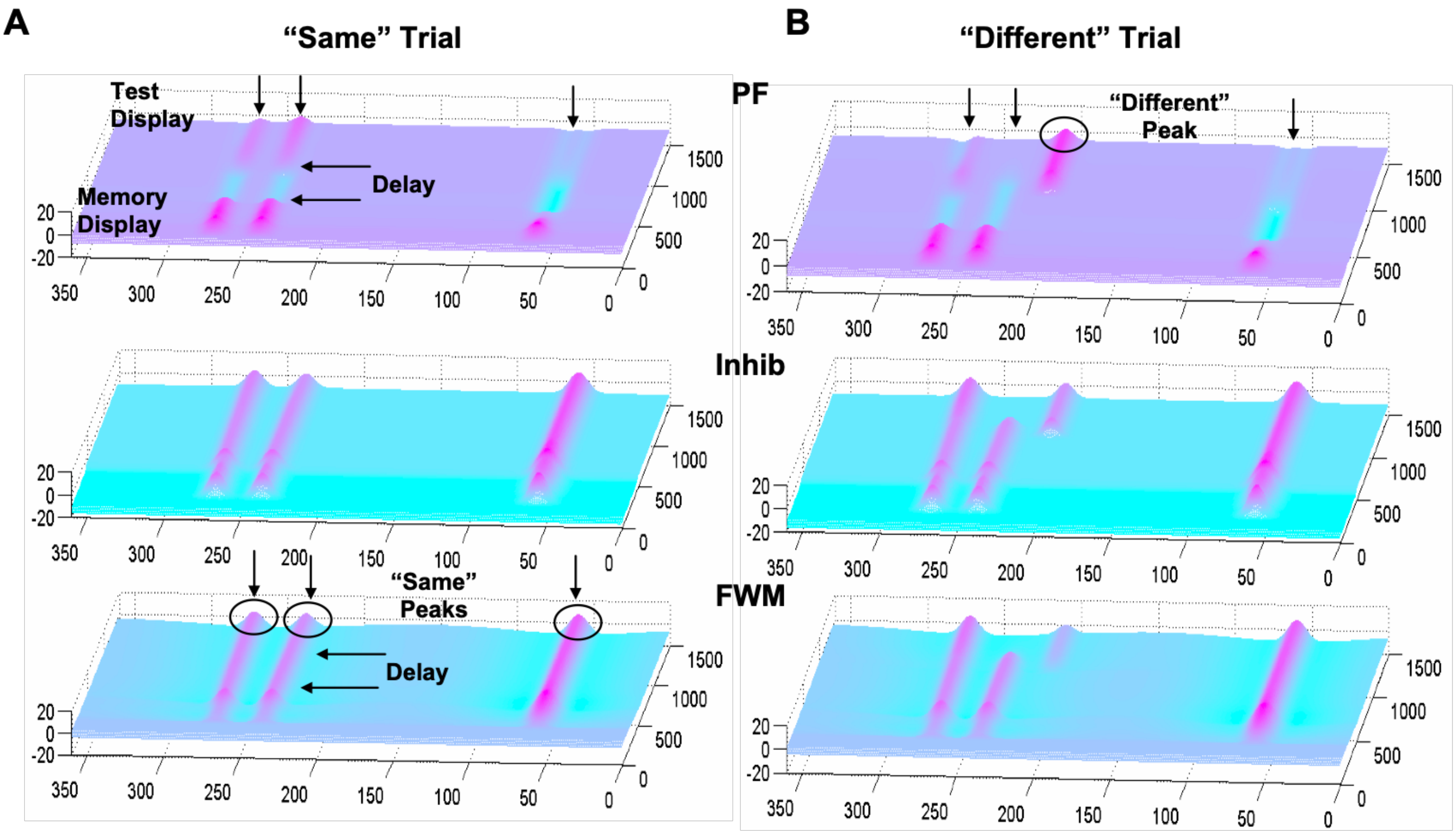
- Seeking empirical evidence for DFT by looking for experimental signatures for core DFT principles:
  - metric effects
  - space-time effects..
- [as contrasted to using DFT to fit data]

Example: visual working memory



[Johnson, Spencer, Luck, Schöner: *Psychological Science* 2008]

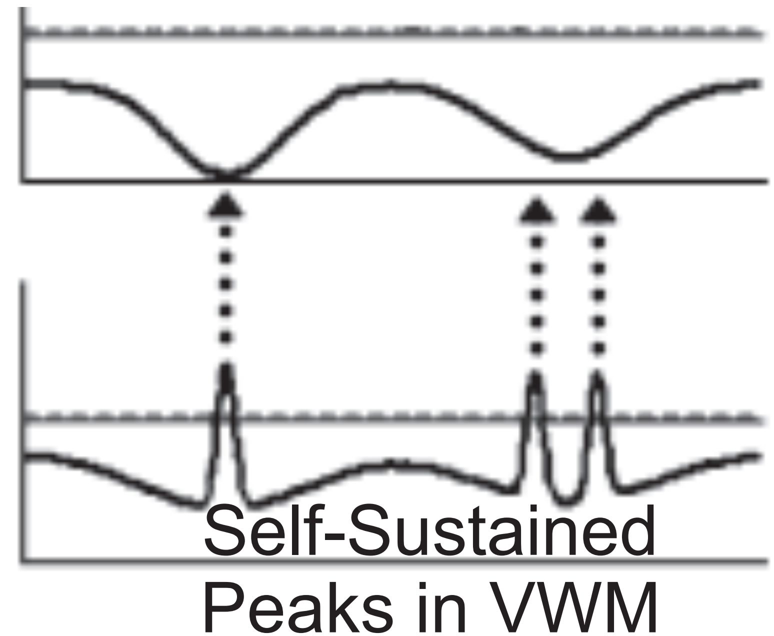
DFT model of change detection



[Johnson, Spencer, Schöner: *New Ideas in Psychology* 2008]

Metric effect

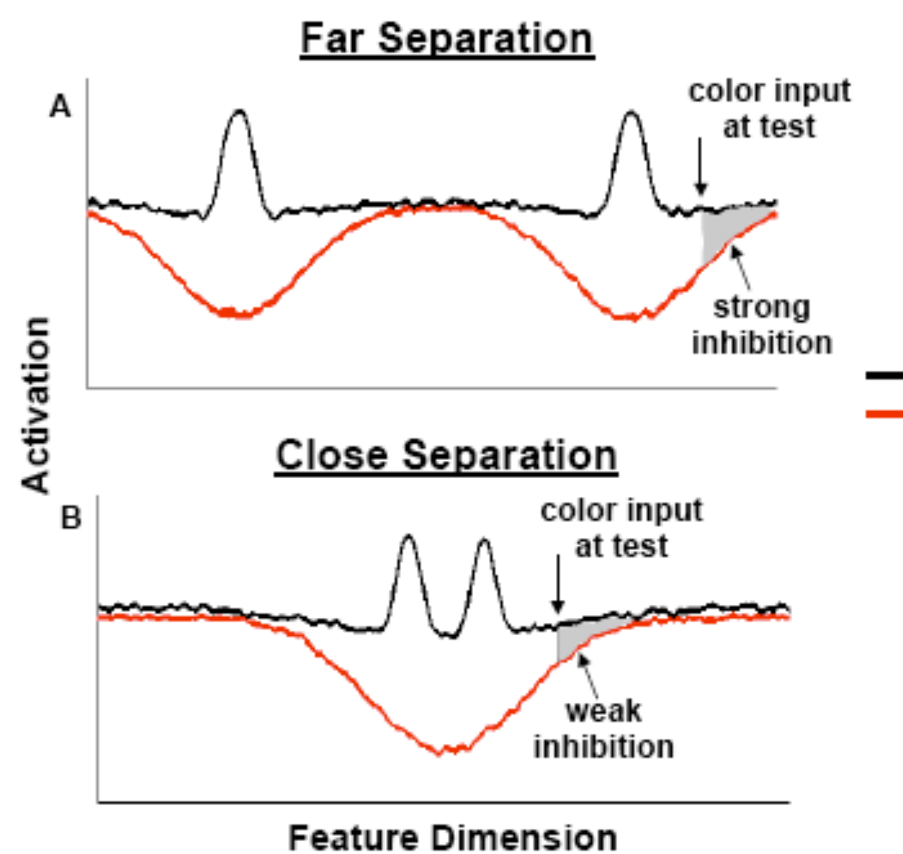
- close metric separation: the WM peaks are weakened by mutual inhibition
- => less inhibition projected to change detection layer



[Johnson, Spencer, Luck, Schöner: *Psychological Science* 2008]

Metric effect

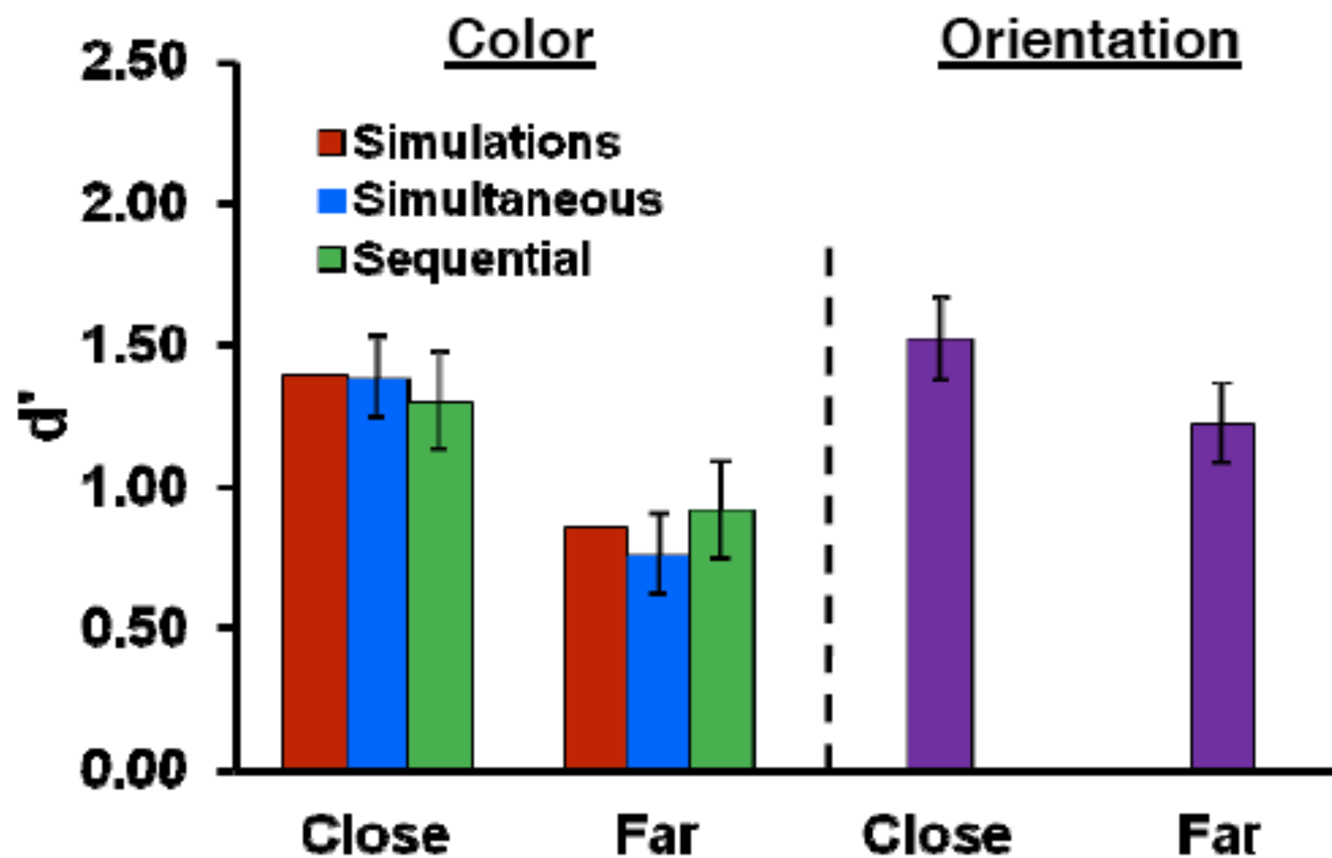
- close metric separation: the WM peaks are weakened by mutual inhibition
- => less inhibition projected to change detection layer
- => lower threshold for change detection



[Johnson, Spencer, Luck, Schöner: *Psychological Science* 2008]

Experimental confirmation

- => predict more sensitive change detection for item that are metrically close!



[Johnson, Spencer, Luck, Schöner: *Psychological Science* 2008]

Such evidence from many different fields

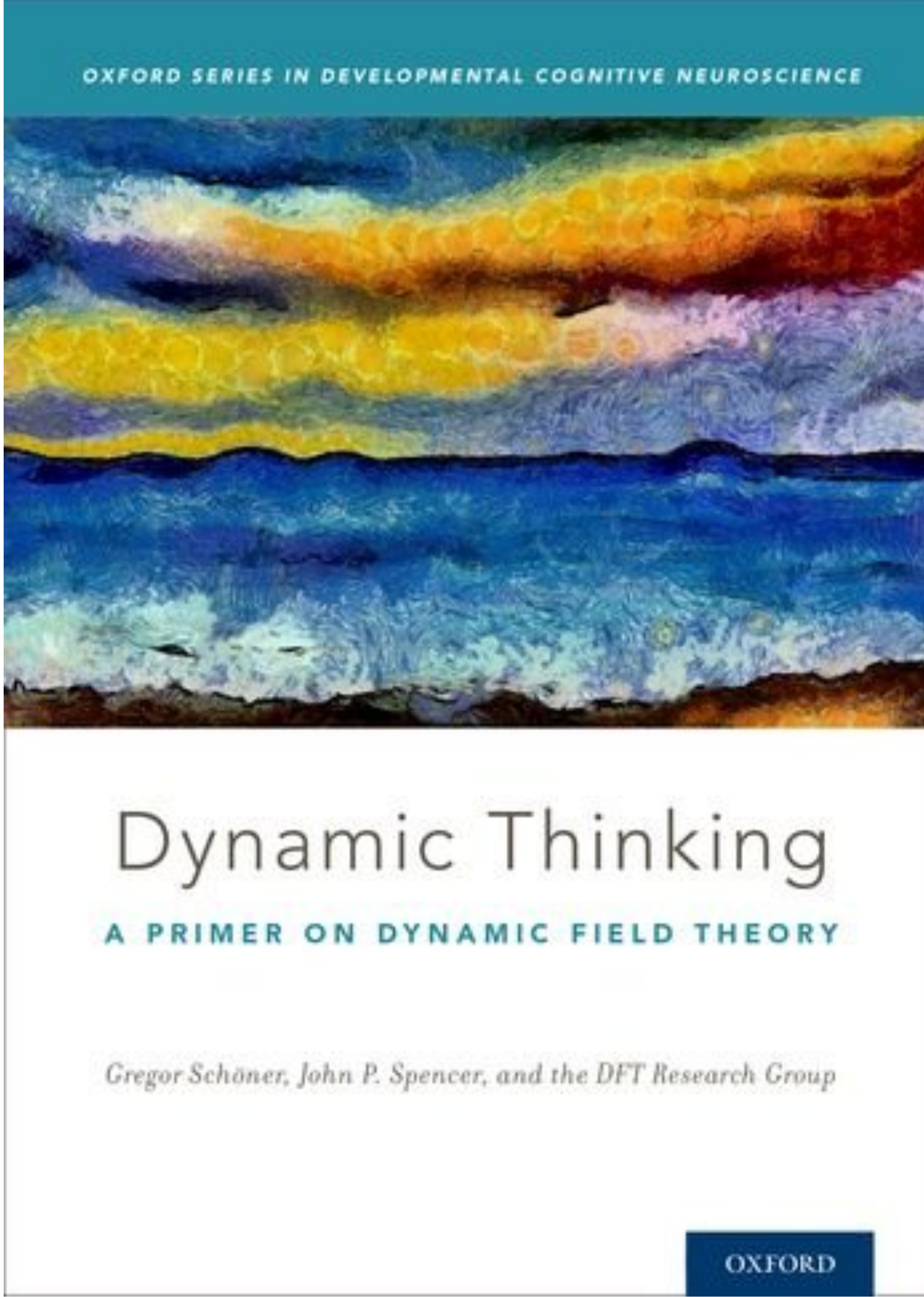
- broad variety
  - of psychological processes
  - of experimental measures

- reaction time classical selection decisions *[Erlhagen, Schöner Psych Rev 2002]*
- perseveration selection decisions (development) *[Thelen, Schöner, Scheier, Smith BBS 2001]*
- spatial and visual working memory (development) *[Spencer, Schutte, Simmering, Johnson JEP, Child development and others]*
- cognitive control (development) *[Buss, Spencer Monographs SRCD]*
- habituation and visual memory (development) *[Thelen, GS Psych Rev 2006; Perone, Spencer, Cog Sci 2013]*

- visual search [*Grieben et al. Att Perc & Psychophysics 2020, CogSci mult*]
- cognitive neuroscience of visual working [*Buss, et al., Psych Rev 2021*]
- situational word learning [*Bath, Spencer, Samuelson, Psych Rev 2021*]
- ideomotor theory [*Vogel-Blaschka, Kunde, Herbort, Scherbaum Psych Rev 2024*]

- perceptually grounding relations *[Richter, Lins, Schöner Cog Sci 2021]*
- perceptually grounding nested phrases [*Sabinasz, Schöner TopiCS 2023; Sabinasz, Richter, Schöner Cog Neurodyn 2023; Sehring et al. CogSci2024*]
- mental mapping *[Kounatidou, Richter, Schöner, CogSci2018]*
- truth value and polarity *[Kati, Sabinasz, Schöner, Gaup CogSci2024]*
- analogical mapping *[Hesse, Sabinasz, Schöner, CogSci 2022; Kang, Sabinasz, Schöner, CogSci 2024]*

- the DFT primer
- resources at:
- dynamicfieldtheory.org



Both strength and challenge

- DFT linking to many different fields of research/sub-disciplines
- at different levels of description
  - behavioral: RT, errors, response metrics, movement
  - neural: population of single units
  - neural: cognitive neuroscience
- demonstration of generative capacity on autonomous agents/robots

Signatures vs. fitting

- DFT can be and has been used to provide quantitative fit of data
- As a process model, it is much more constrained in doing so than more reduced models
    - e.g. diffusion vs. DFT account of decision making
- DFT is NOT aimed to provide data compression

Dynamic Field Theory (DFT)

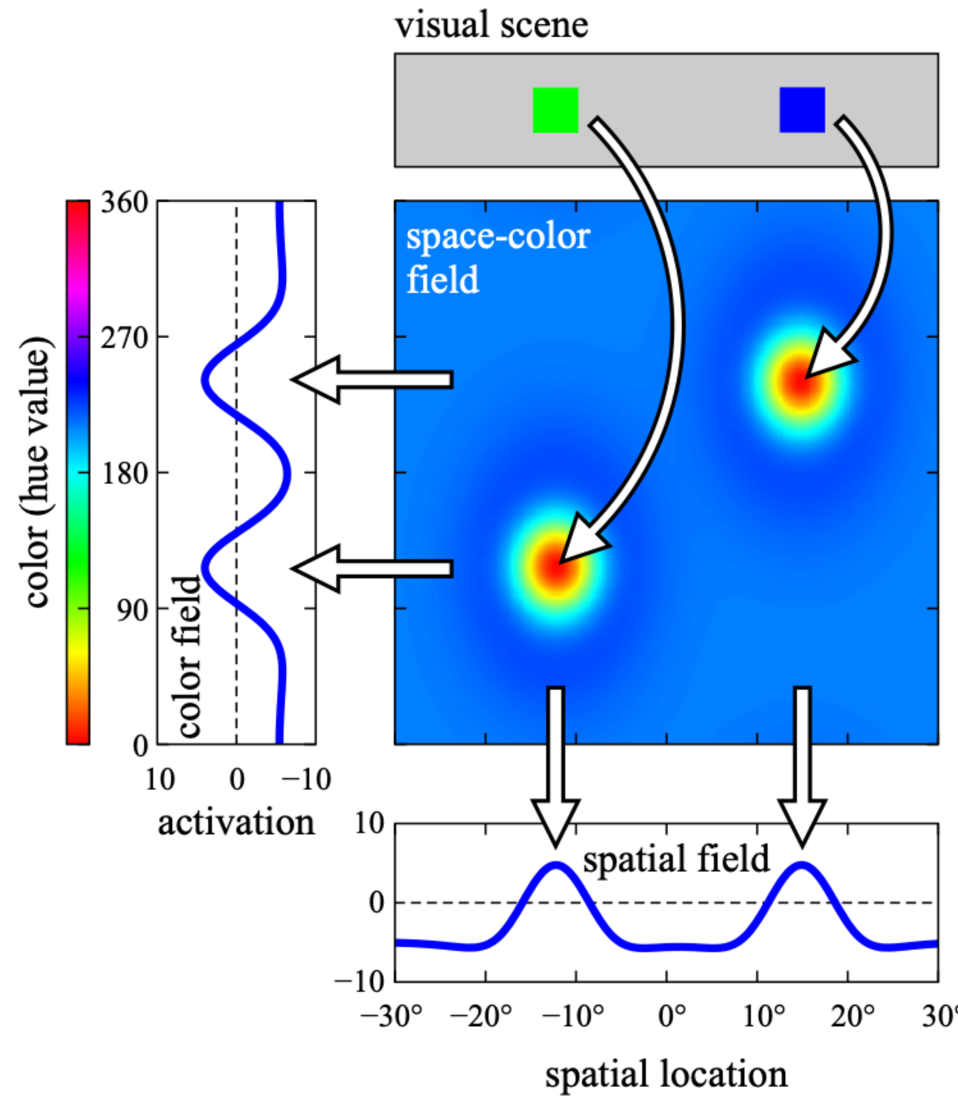
- 1 Space: inner states are localized activation patterns in low-dimensional features spaces
- 2 Time: autonomy derives from neural dynamics, attractors and dynamic instabilities
- 3 Coupling: cognition emerges from dynamic coupling across low-dimensional features spaces
- 4 Integration: in DFT architectures

3 Coupling

- binding, unbinding
- mental maps
- cued selection
- binding through space
- coordinate transforms

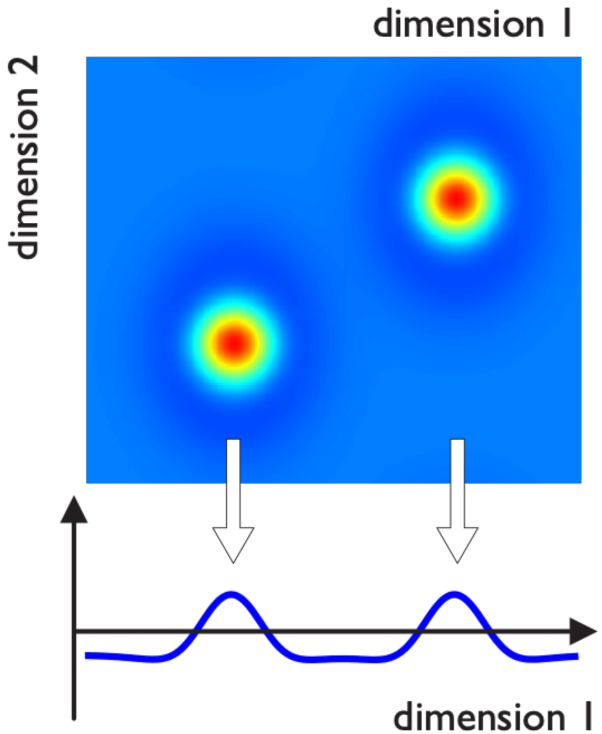
Joint representations

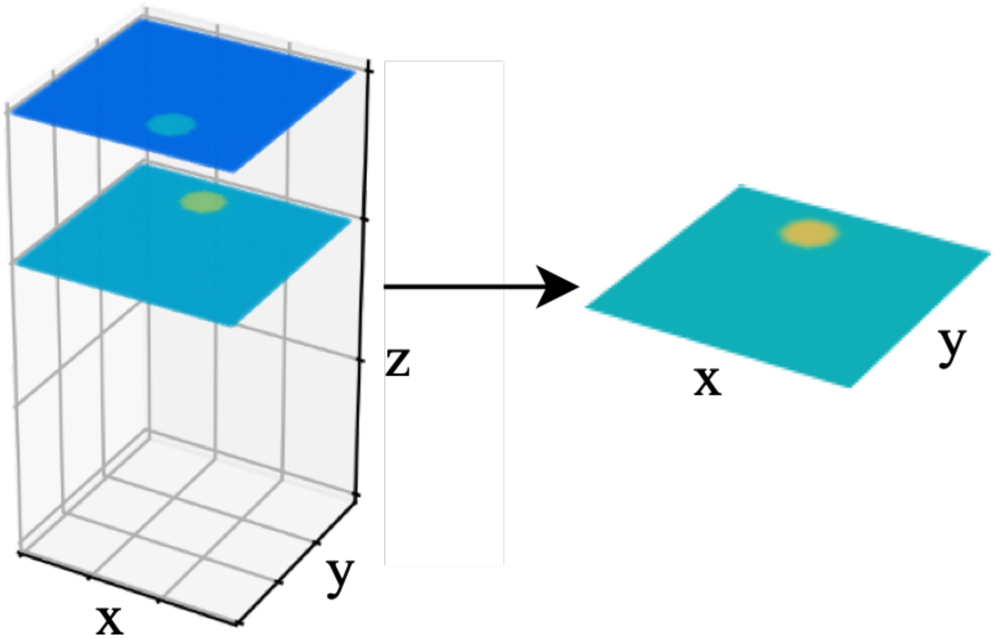
- unbinding = extracting features by contraction coupling



[Schneegans et al., Ch 5 of *DFT Primer*, 2016]

Contraction coupling

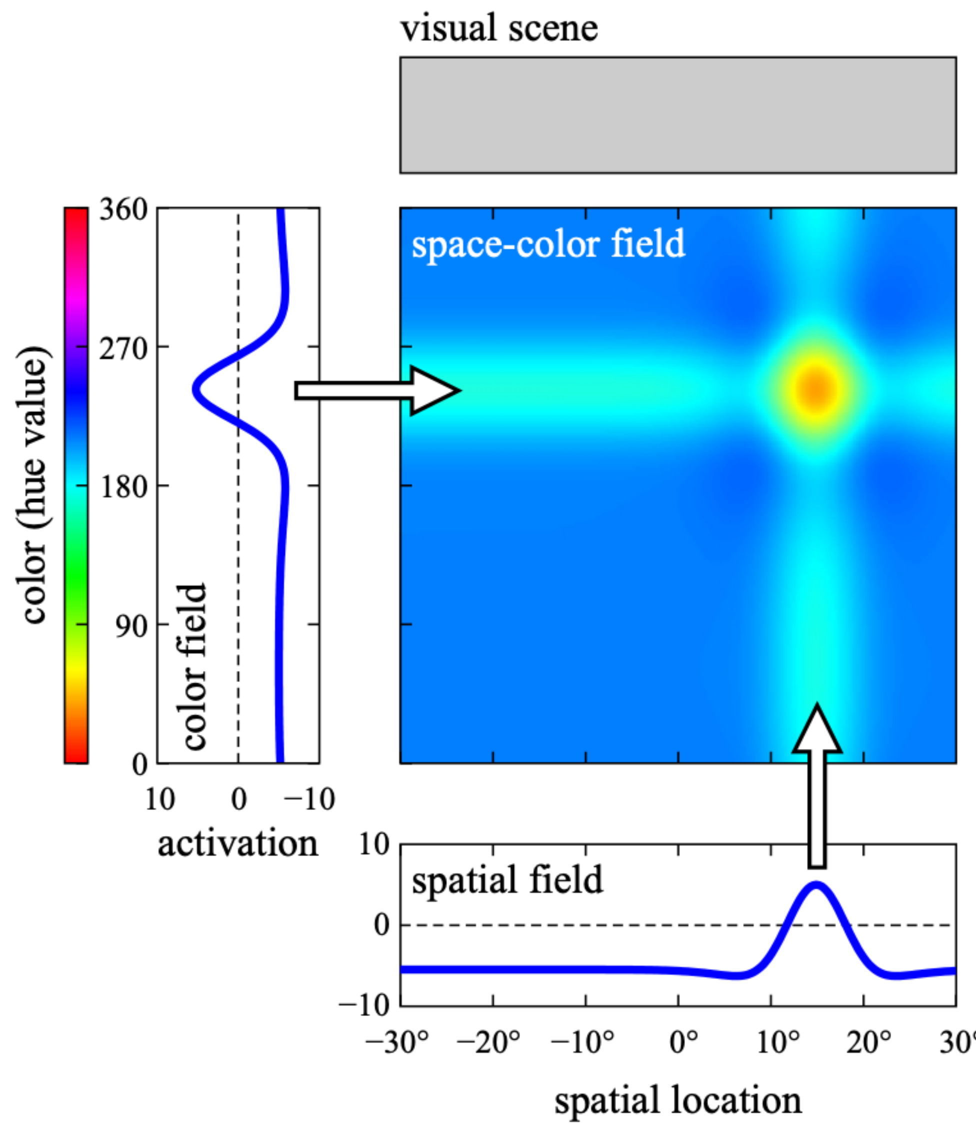




[Sabinasz, Richter, Schöner, *Cog. Neurodyn.* 2023]

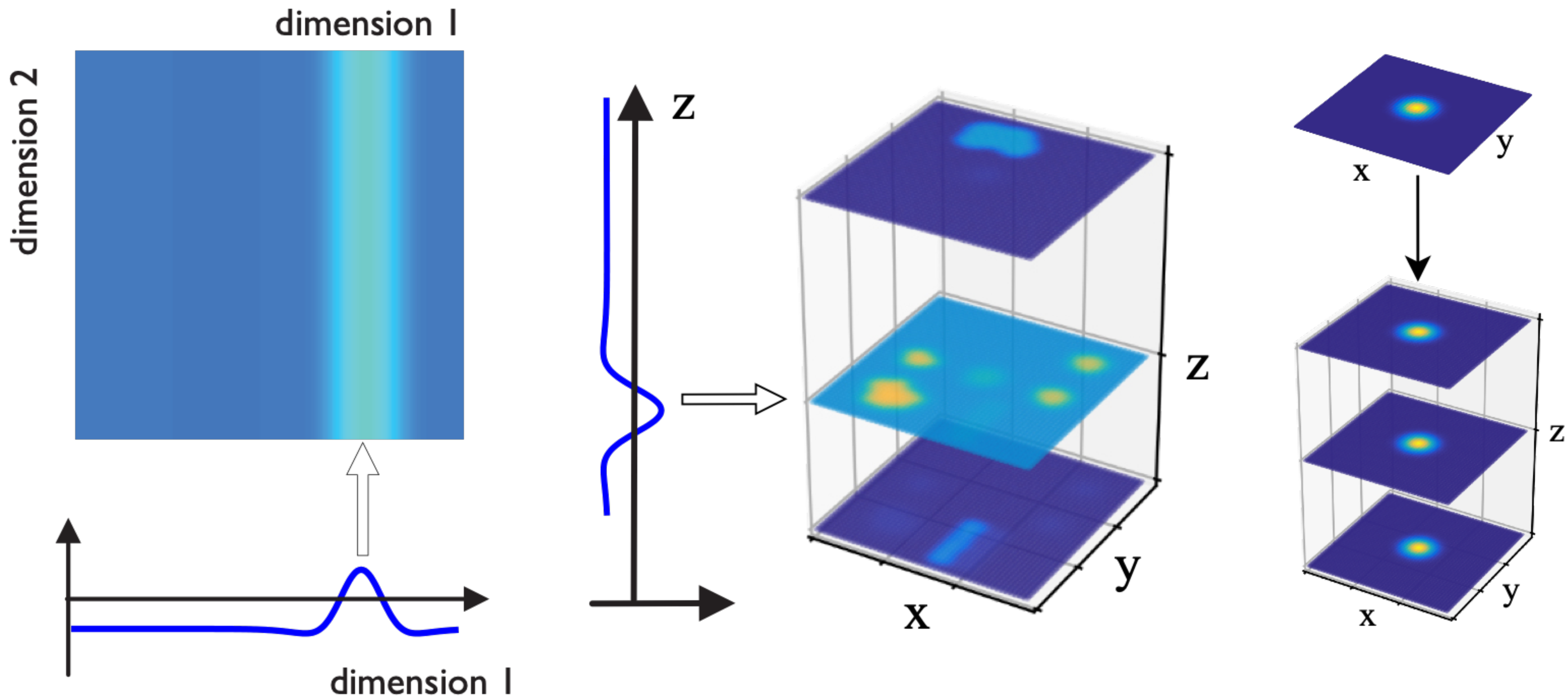
Binding in mental maps

- bind separate features into objects in mental maps
- by expansion mapping



[Schneegans et al., Ch 5 of *DFT Primer*, 2016]

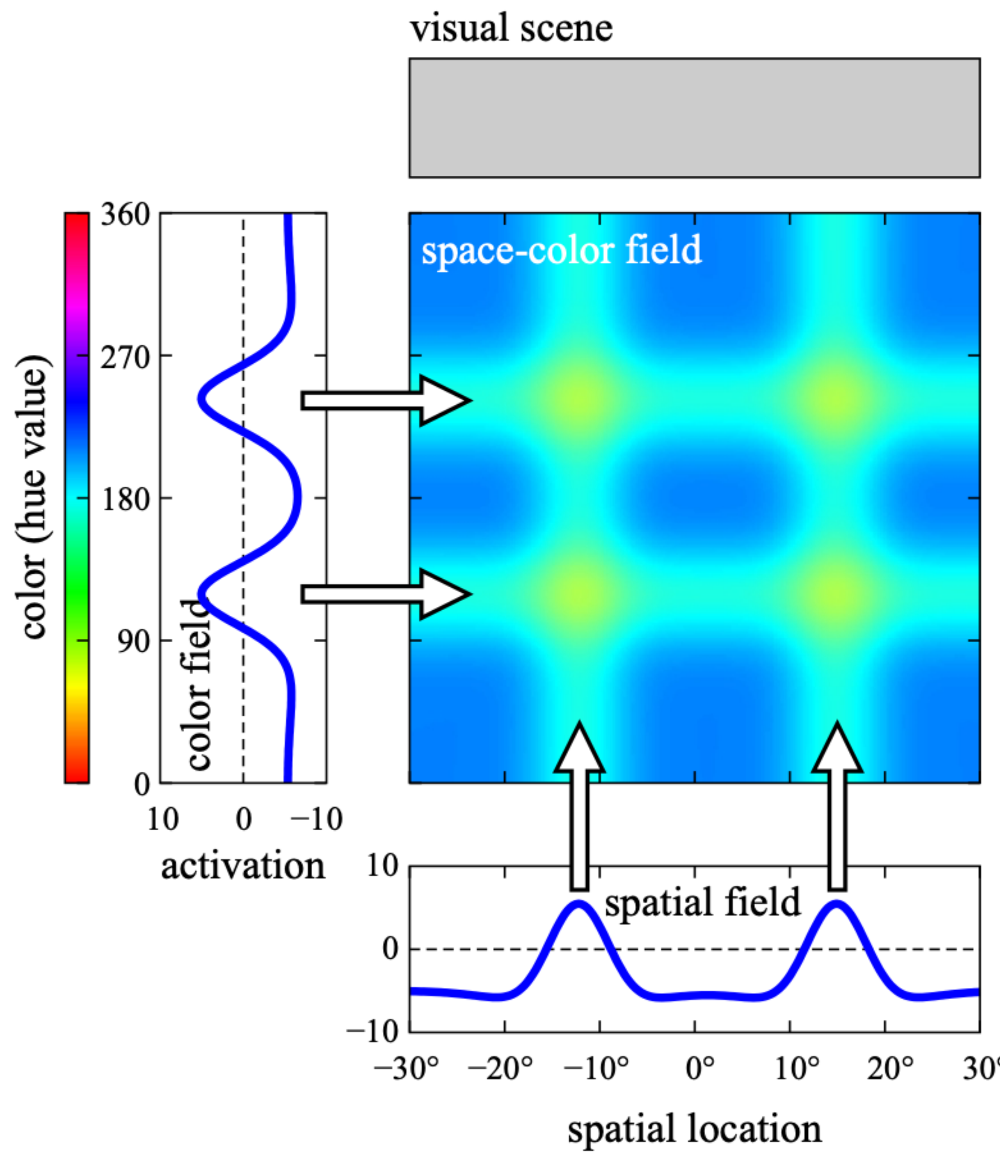
Expansion coupling



[Sabinasz, Richter, Schöner, *Cog. Neurodyn*. 2023]

Binding problem

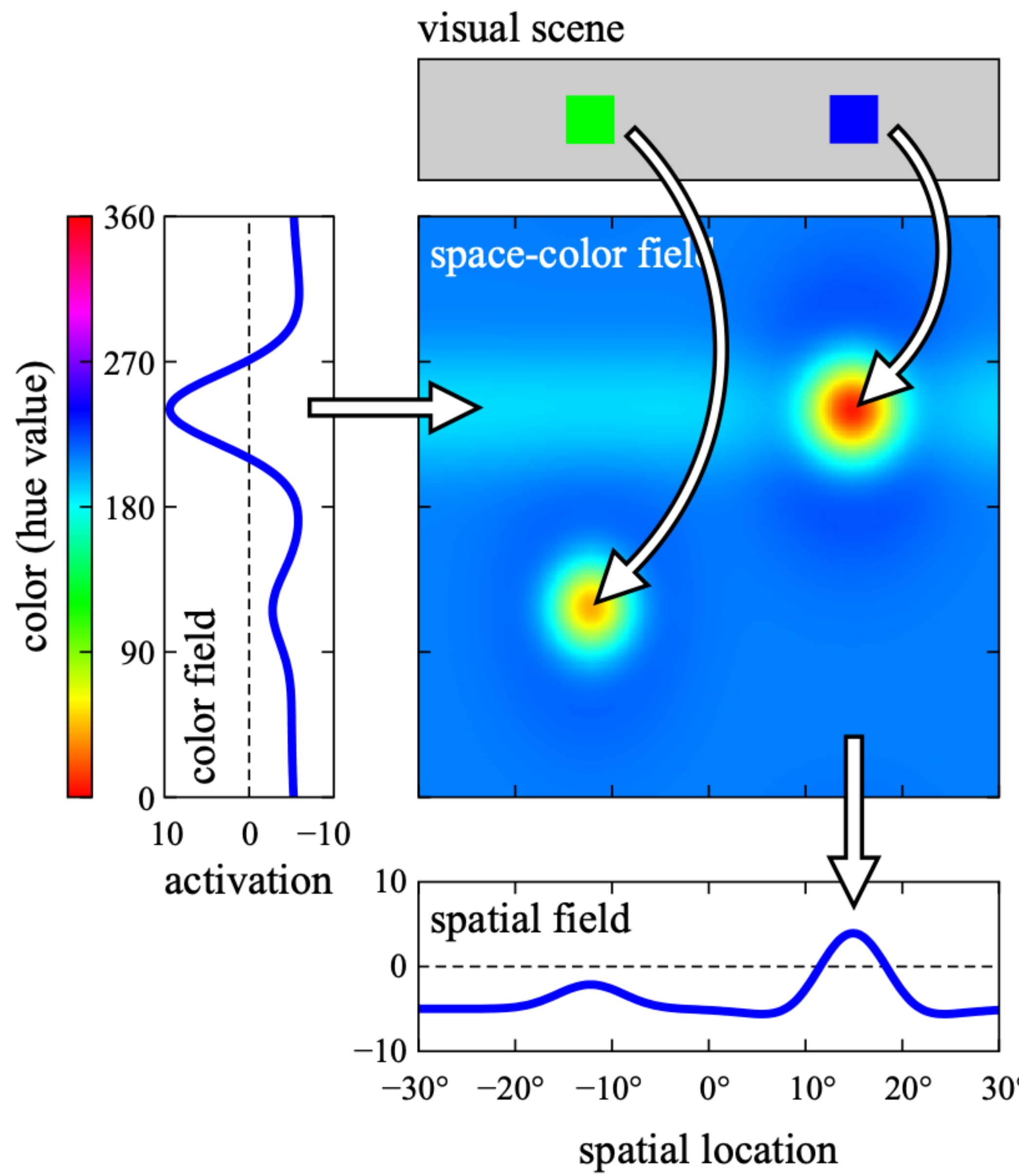
- => bind one object at a time
- => attentional bottleneck



[Schneegans et al., Ch 5 of *DFT Primer*, 2016]

Cued selection

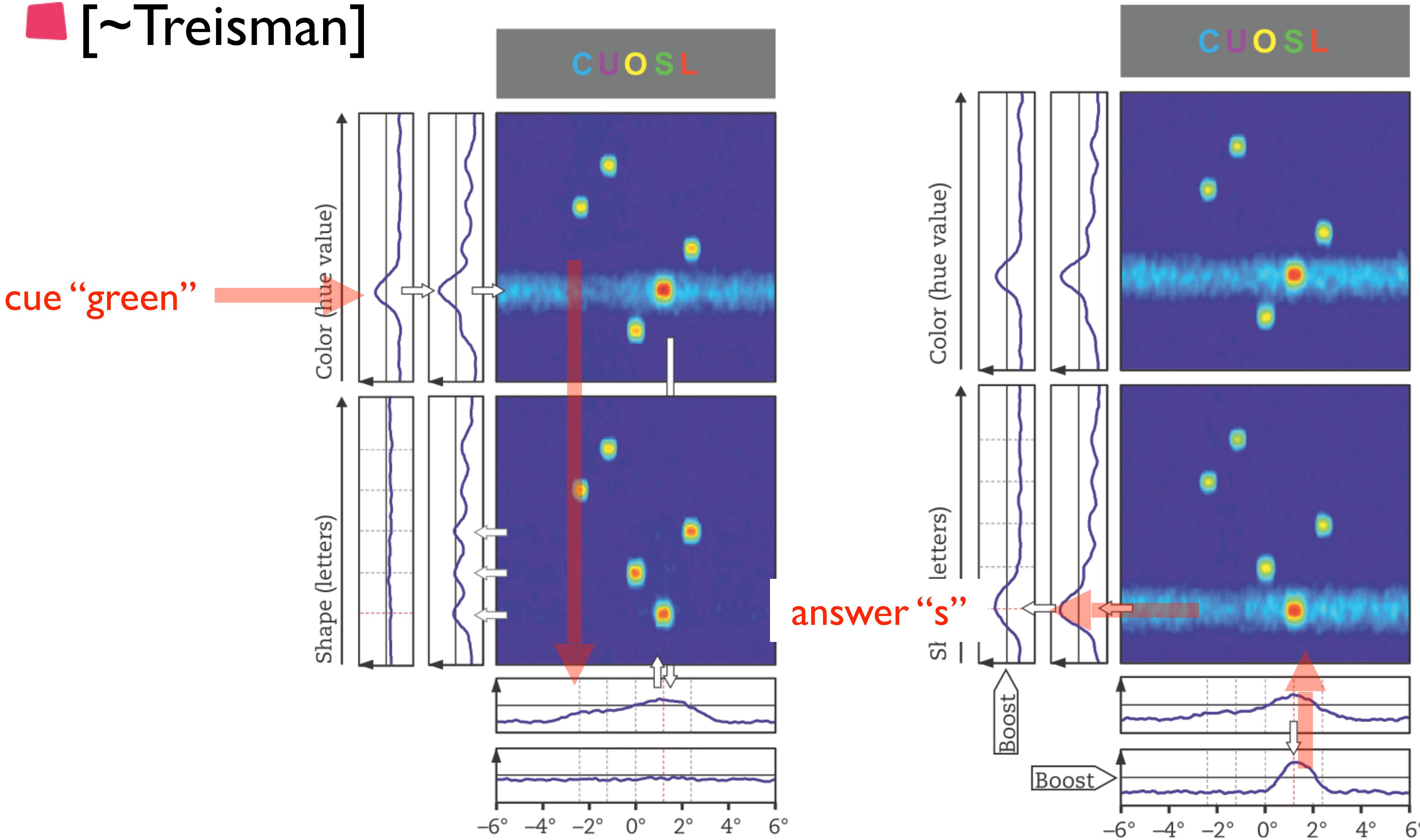
- combines expansion and contraction



[Schneegans et al., Ch 5 of *DFT Primer*, 2016]

Binding through space

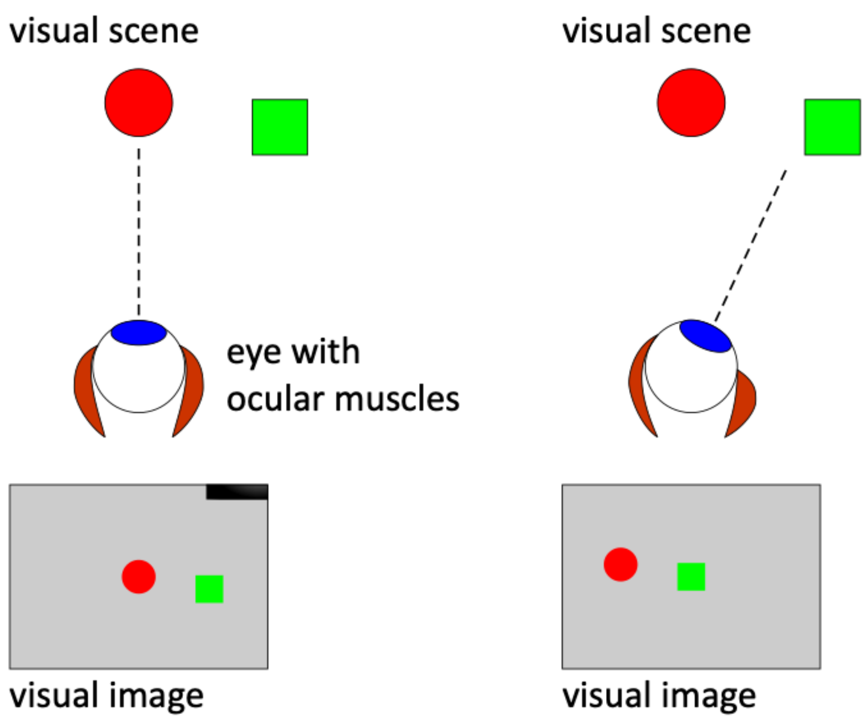
- [~Treisman]



[Schneegans et al.,Ch 5 of *DFT Primer*, 2016]

Coordinate transforms

- enable representations that are more invariant than the sensory-motor surfaces



body/world frame

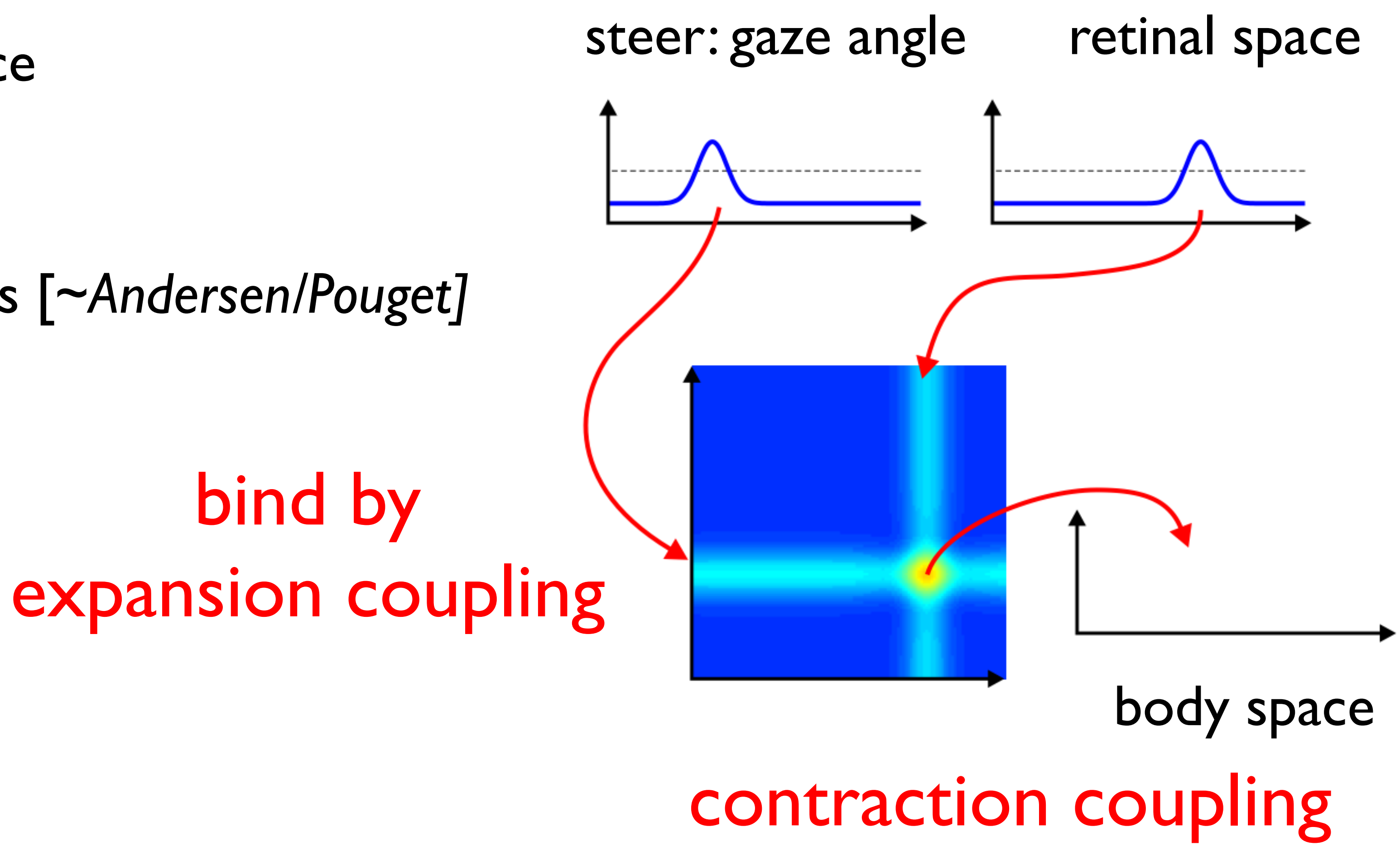
gaze shift

retinal frame

[Schneegans Ch 7, *DFT Primer*, 2016]

Example: retinal <=> body/world space

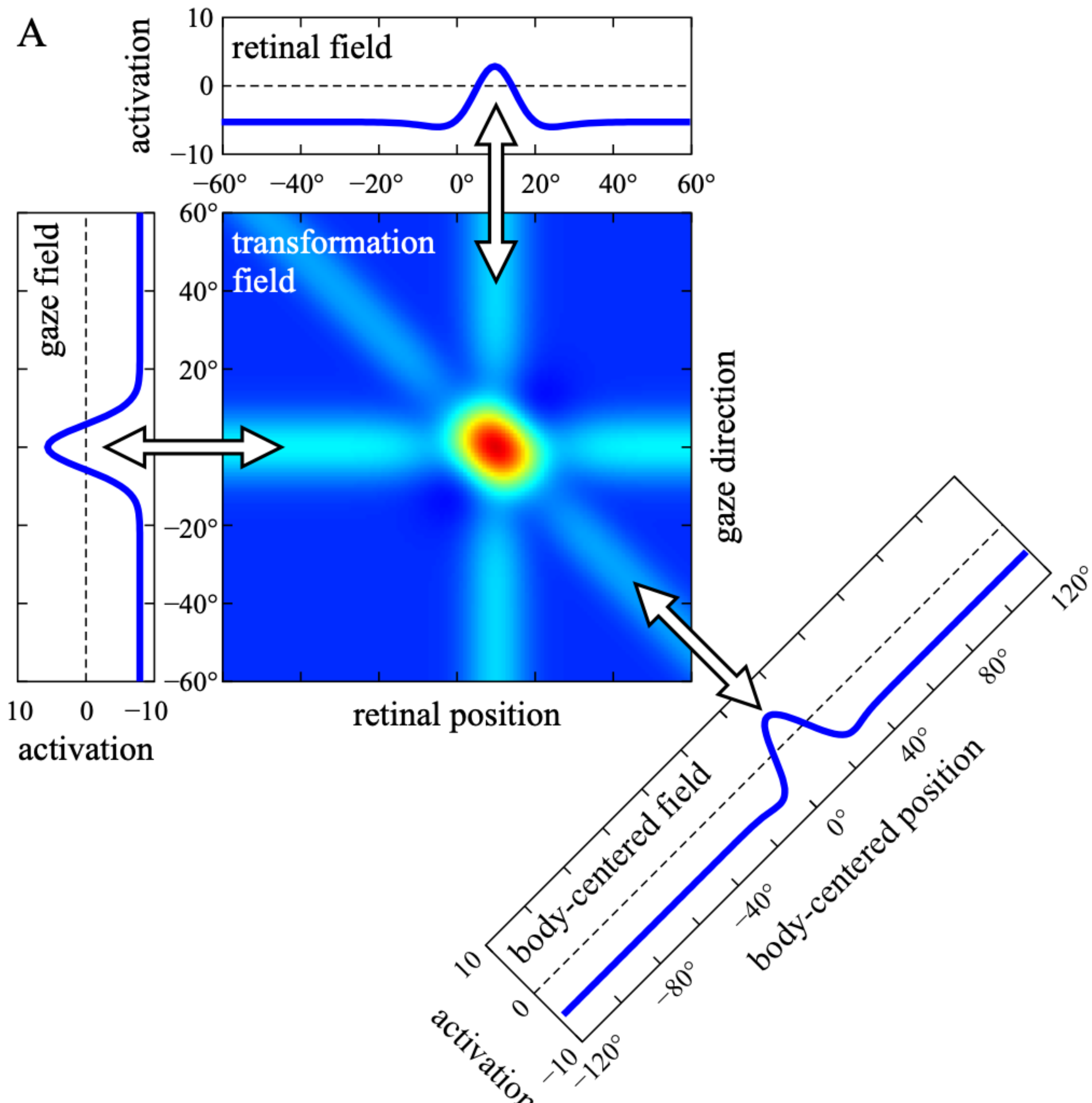
- joint representation of
  - retinal space
  - gaze angle
  - = gain fields [~*Andersen/Pouget*]



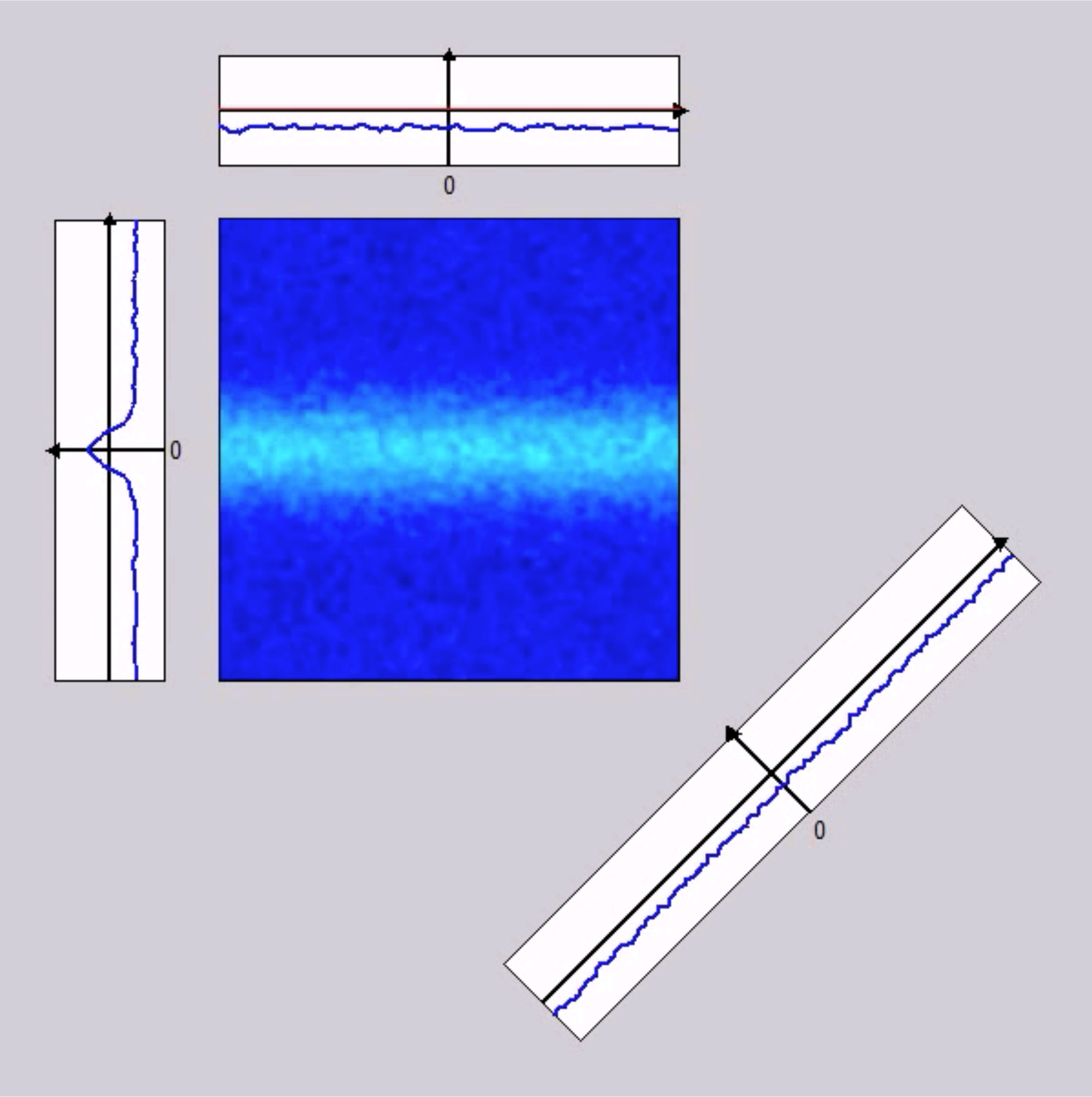
[Schneegans Ch 7, *DFT Primer*, 2016]

Retinal <=> body/world space

- bi-directional coupling



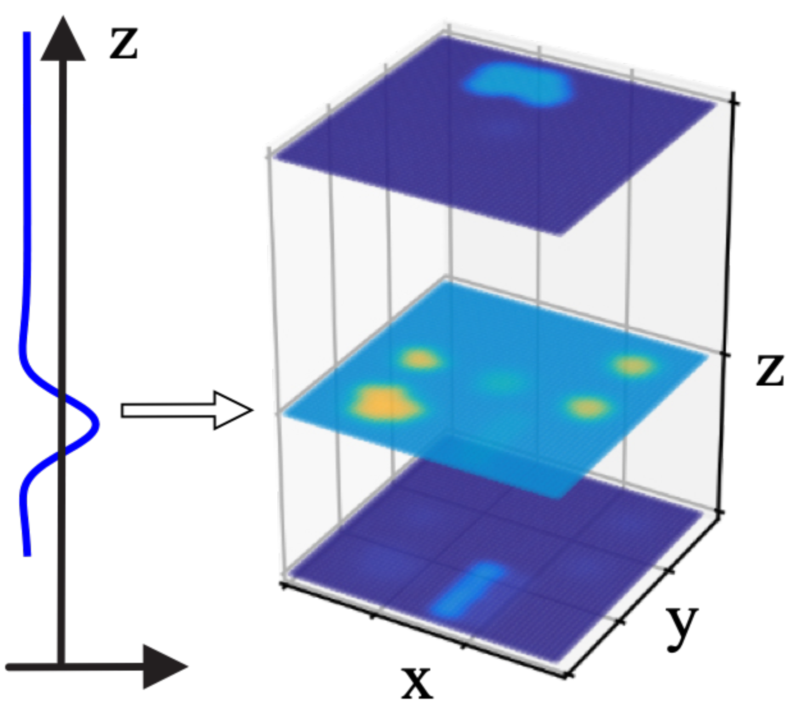
Spatial remapping during saccades

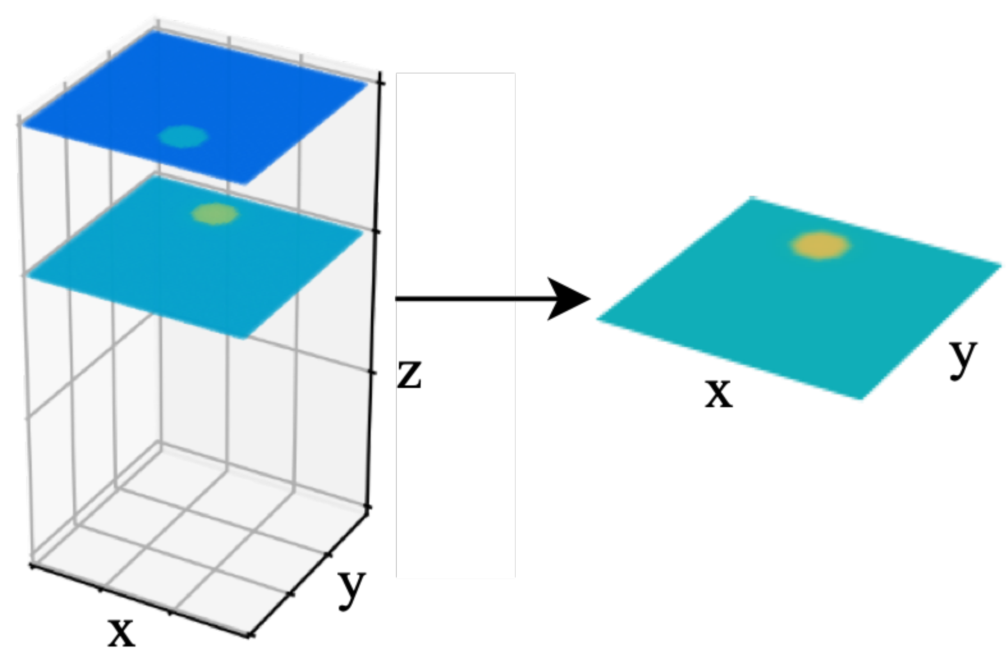


[Schneegans, Schöner *Biological Cybernetics* 2012]

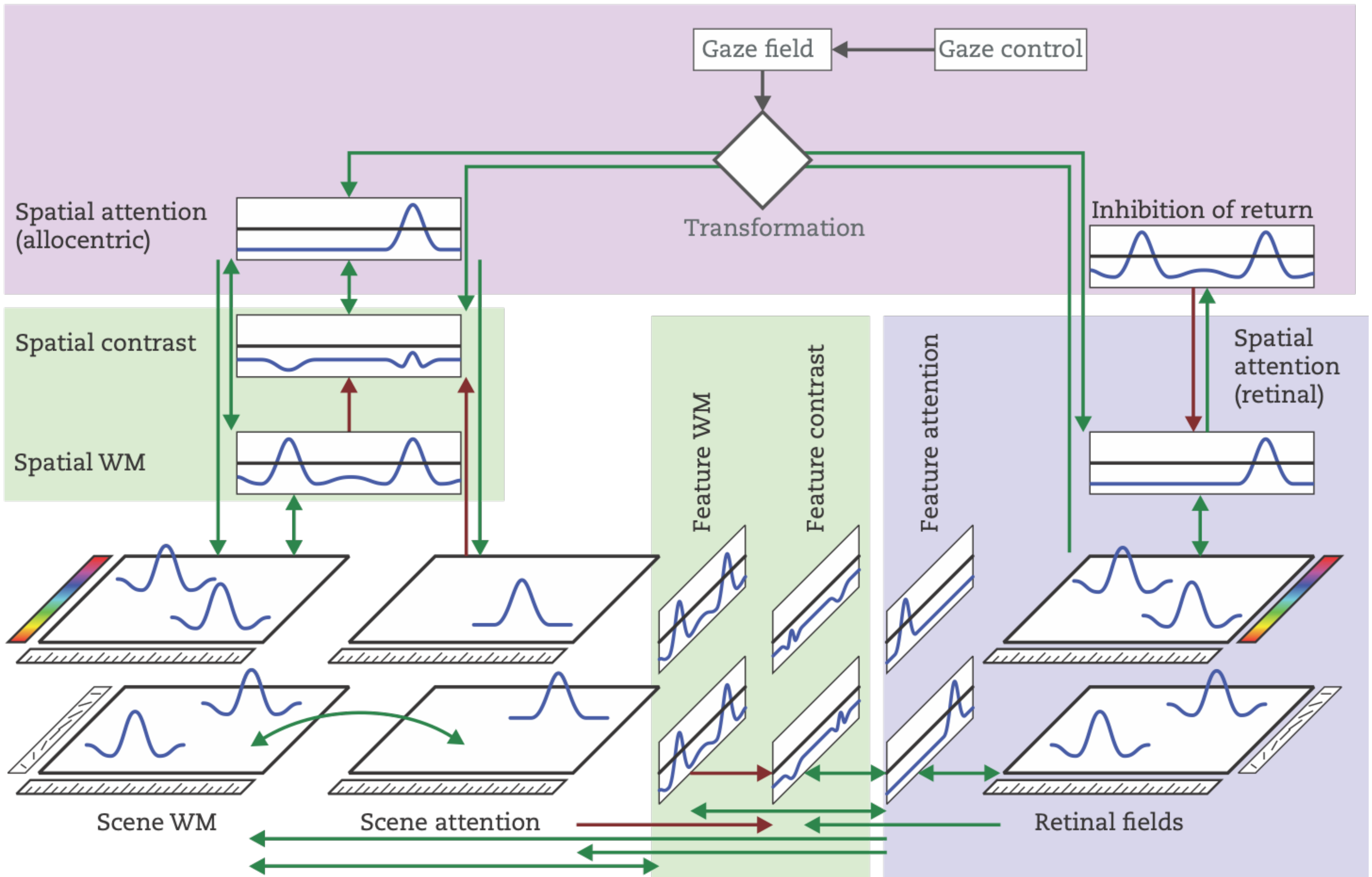
4 Integration: DFT architectures

- dynamic modularity: fields retain their dynamic regime under coupling
- coupling must preserve feature dimensions: "non-synesthesia principle"

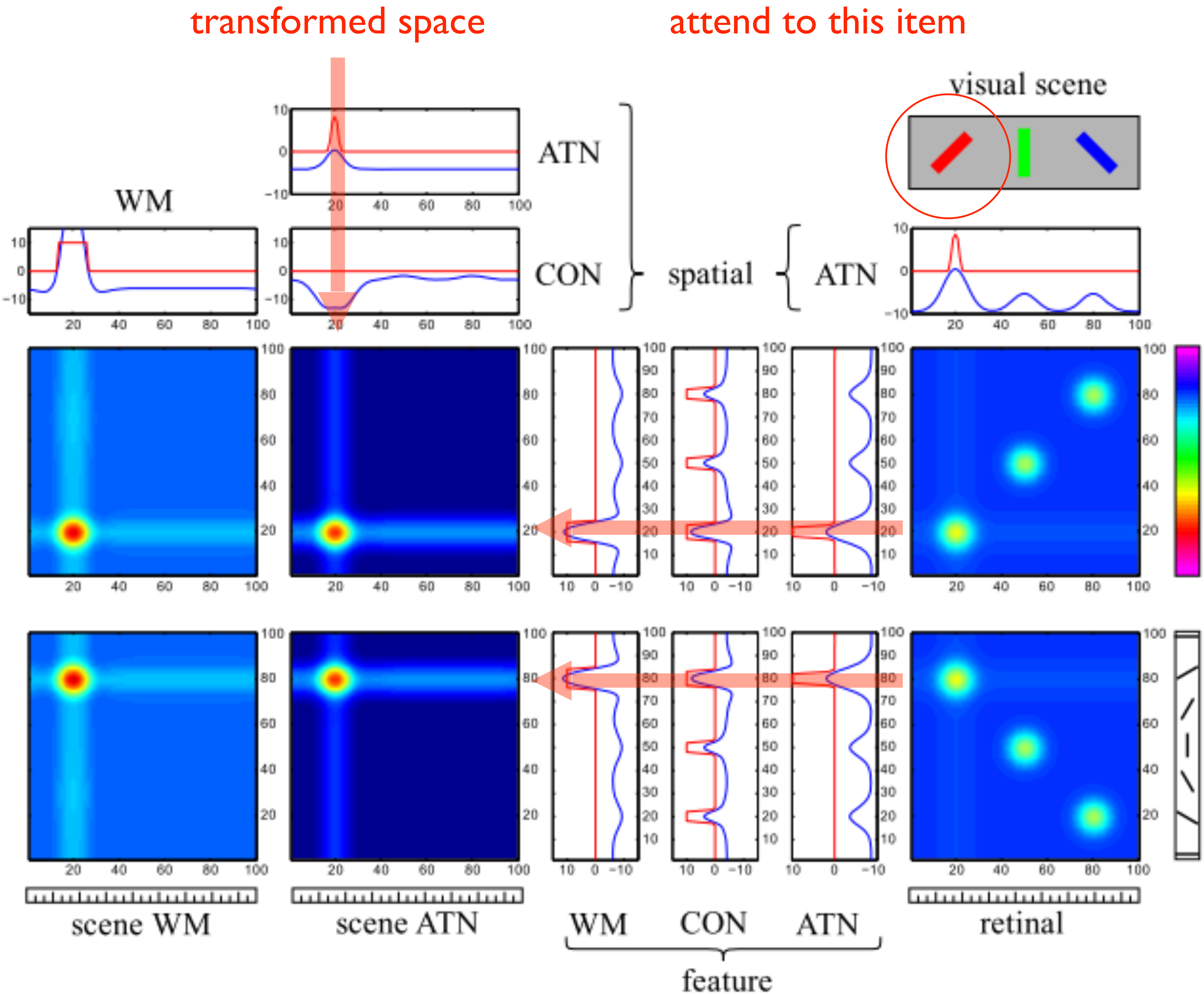




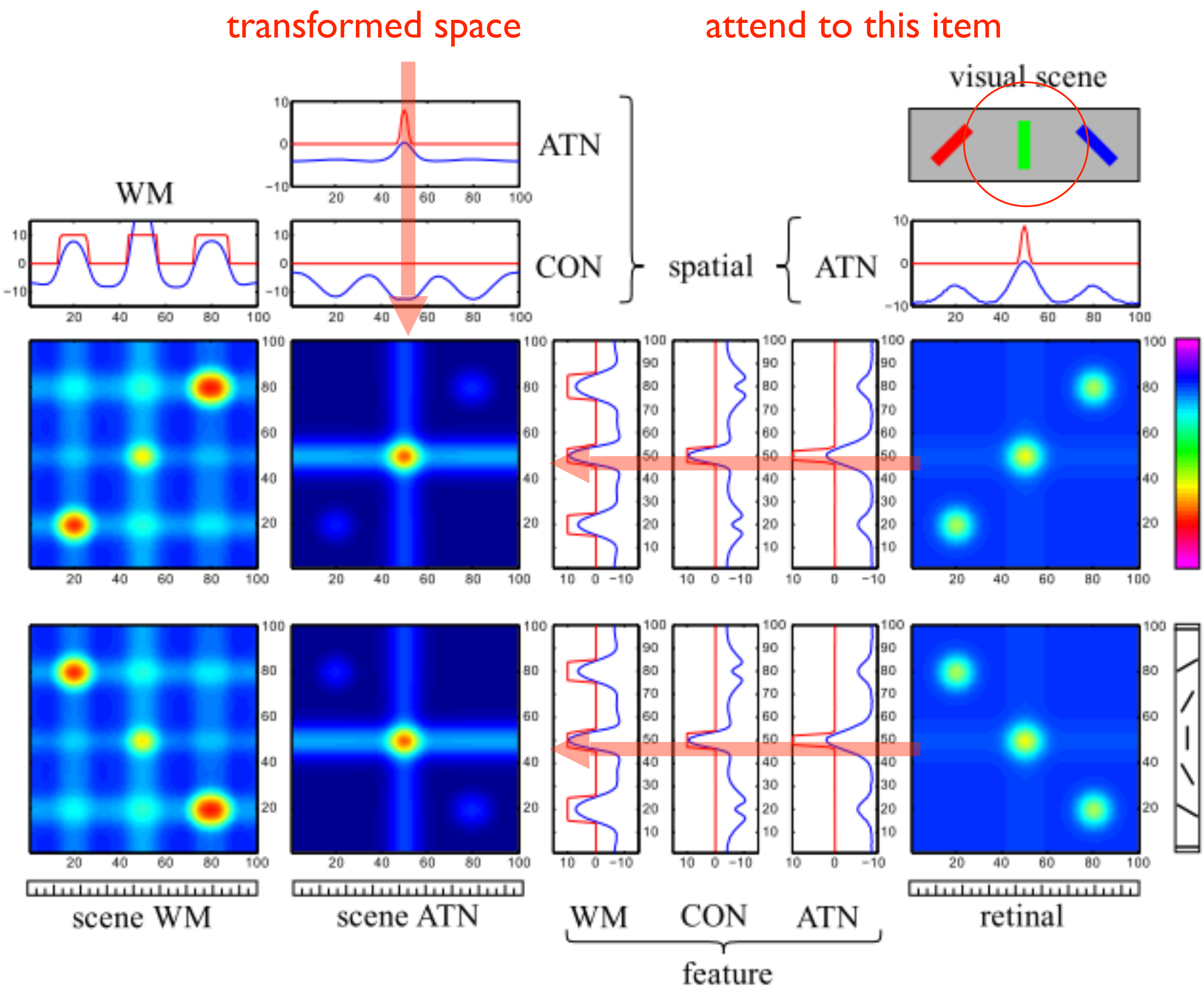
DFT architecture of scene memory



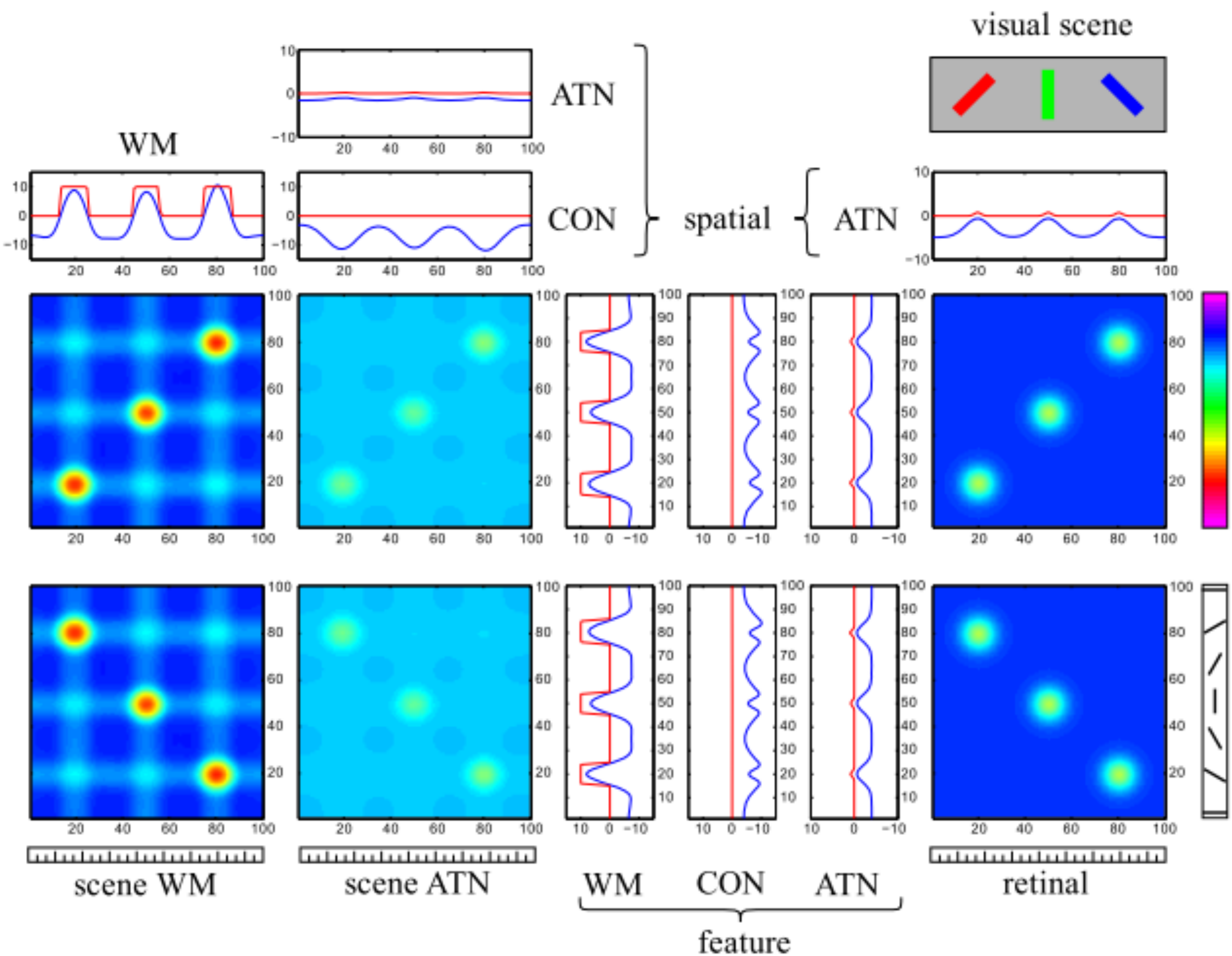
[Schneegans et al.,Ch 8 of *DFT Primer*, 2016]



[Schneegans et al.,Ch 8 of *DFT Primer,* 2016]



[Schneegans et al.,Ch 8 of *DFT Primer,* 2016]



[Schneegans et al.,Ch 8 of *DFT Primer,* 2016]

Dynamic Field Theory (DFT)

- 1 Space: inner states are localized activation patterns in low-dimensional features spaces
- 2 Time: autonomy derives from neural dynamics, attractors and dynamic instabilities
- 3 Coupling: cognition emerges from dynamic coupling across low-dimensional features spaces
- 4 Integration: in DFT architectures

A word on integration

- integration takes the perspective of the modeler: component models that are integrated
- in the brain, everything is always integrated/coupled/connected… the issue is rather:
- decoupling, functional differentiation, flexibility, task-specific configuration
- DFT does address that …

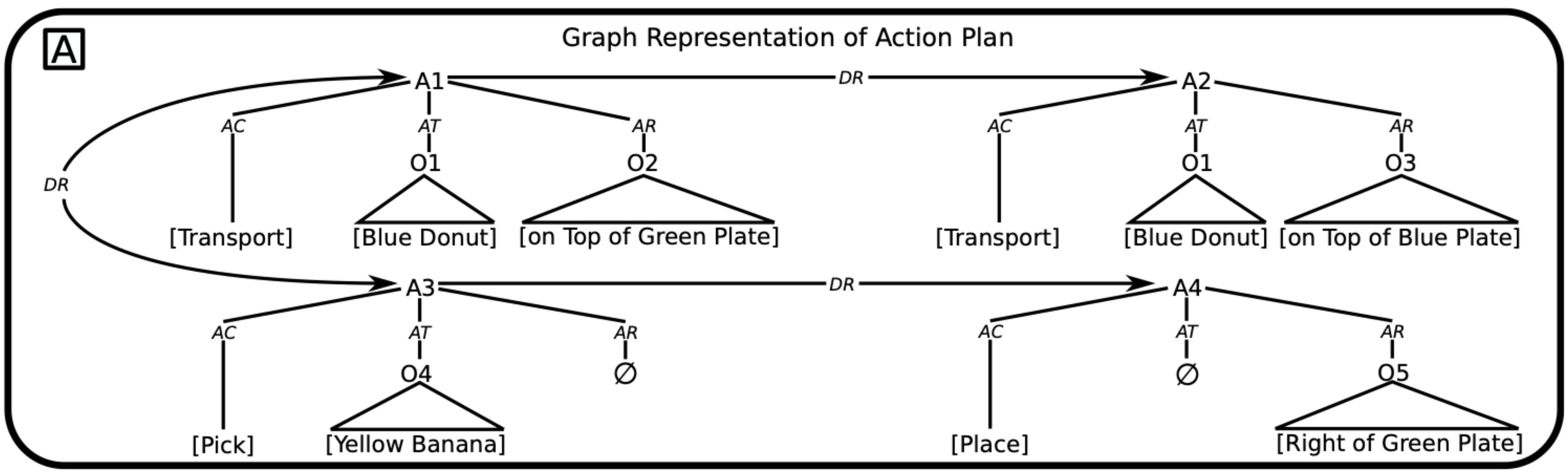
Summary: Conceptual commitments of DFT

- connections determine function [~connectionism]
- localist representations
- in low-dimensional spaces
- interaction dominated dynamics => attractor states
- instabilities generate decisions, sequences: autonomy
- coupling ~ binding as path to higher cognition

… toward higher cognition

- Example: grounding nested imperative phrases
- conceptual structure

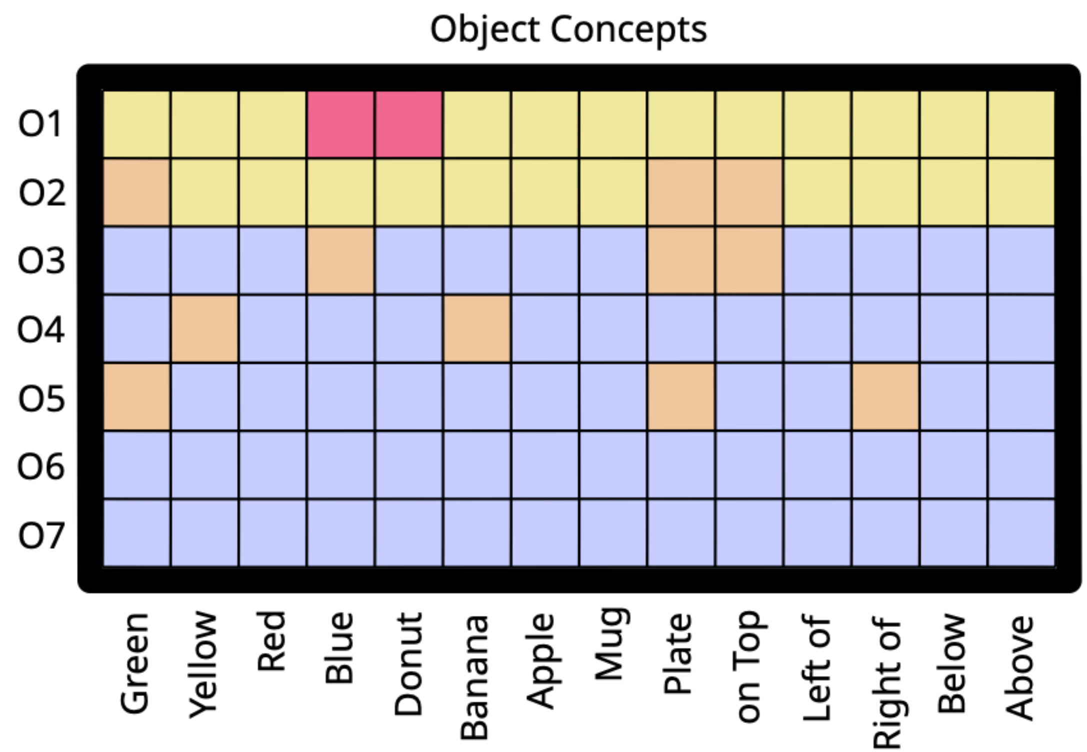
"transport blue donut onto green plate
then transport blue donut to blue plate
or pick yellow banana and place right of the green plate""



[Sehring et al. *CogSci* 2024]

Neural representation of conceptual structure

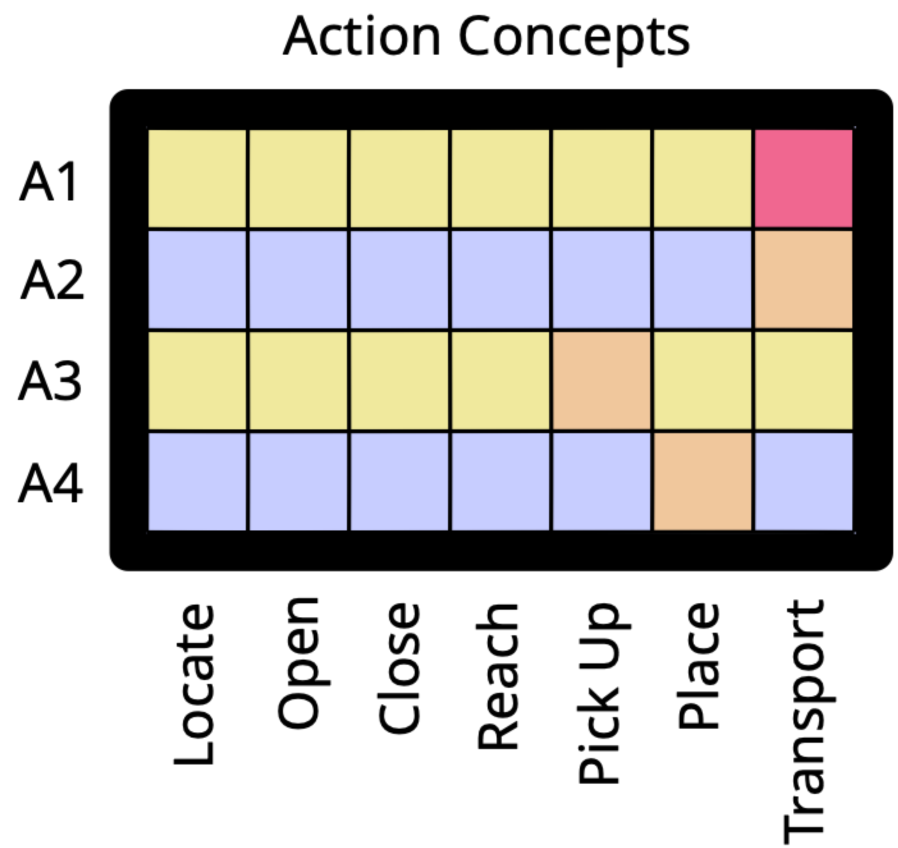
- ordinal index given to each instance (token) of an object concept
- enables representing multiple instances of an object concept
- serves as a binding dimension



[Sabinasz, Richter, Schöner *Cog Neurodyn* 2023; Sehring et al. *CogSci* 2024]

Neural representation of conceptual structure

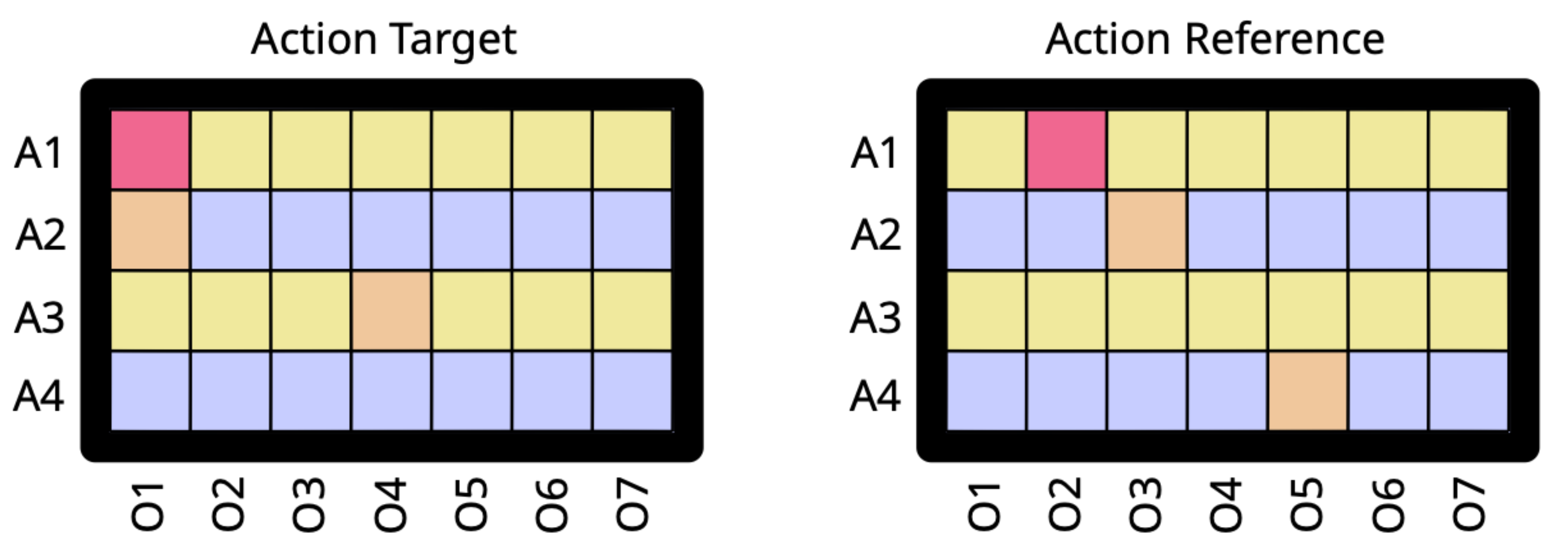
- ordinal index given to each instance of a relation/action concept
- enables representing multiple instances of same relation in a nested phrase

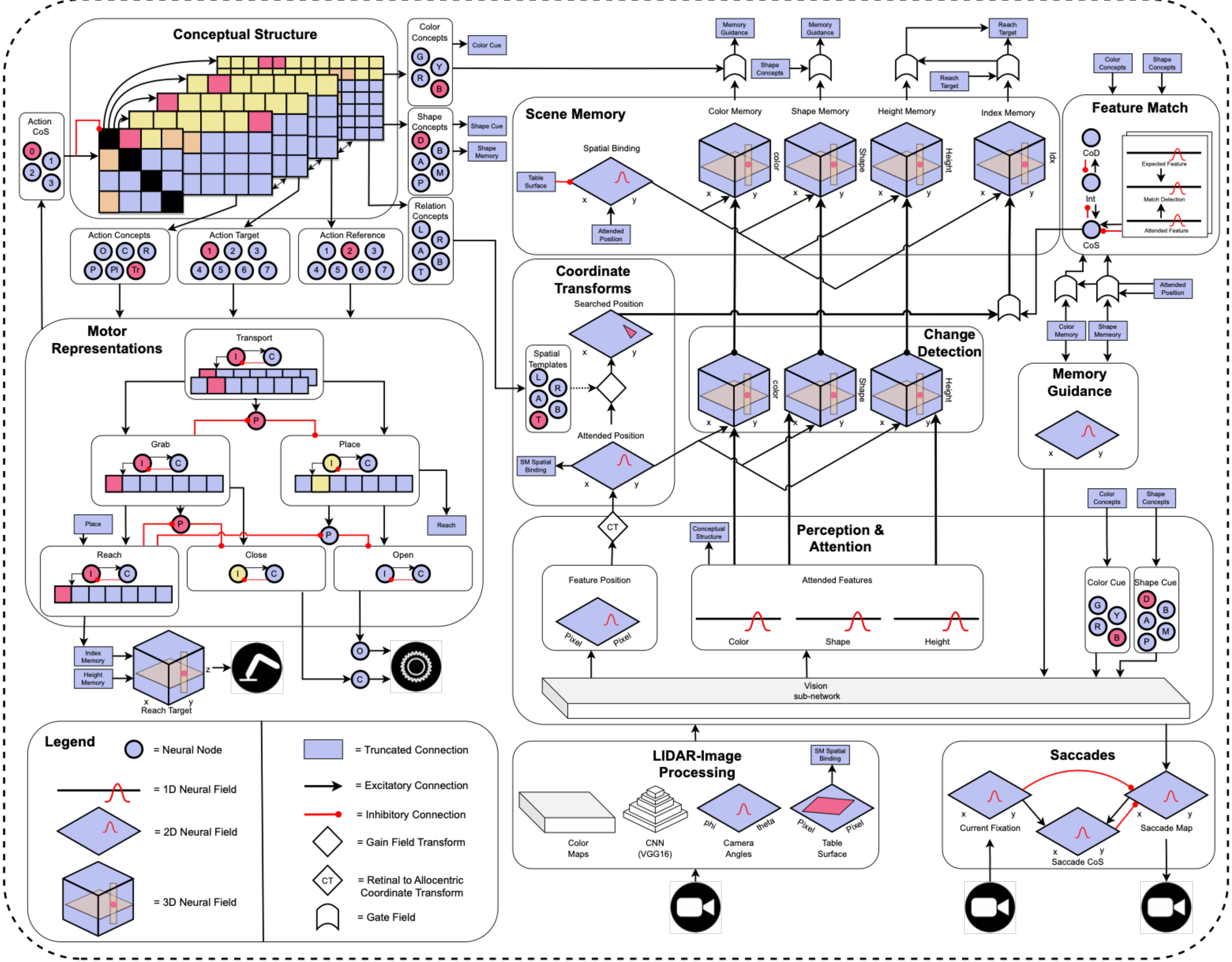


[Sabinasz, Richter, Schöner *Cog Neurodyn* 2023; Sehring et al. *CogSci* 2024]

Neural representation of conceptual structure

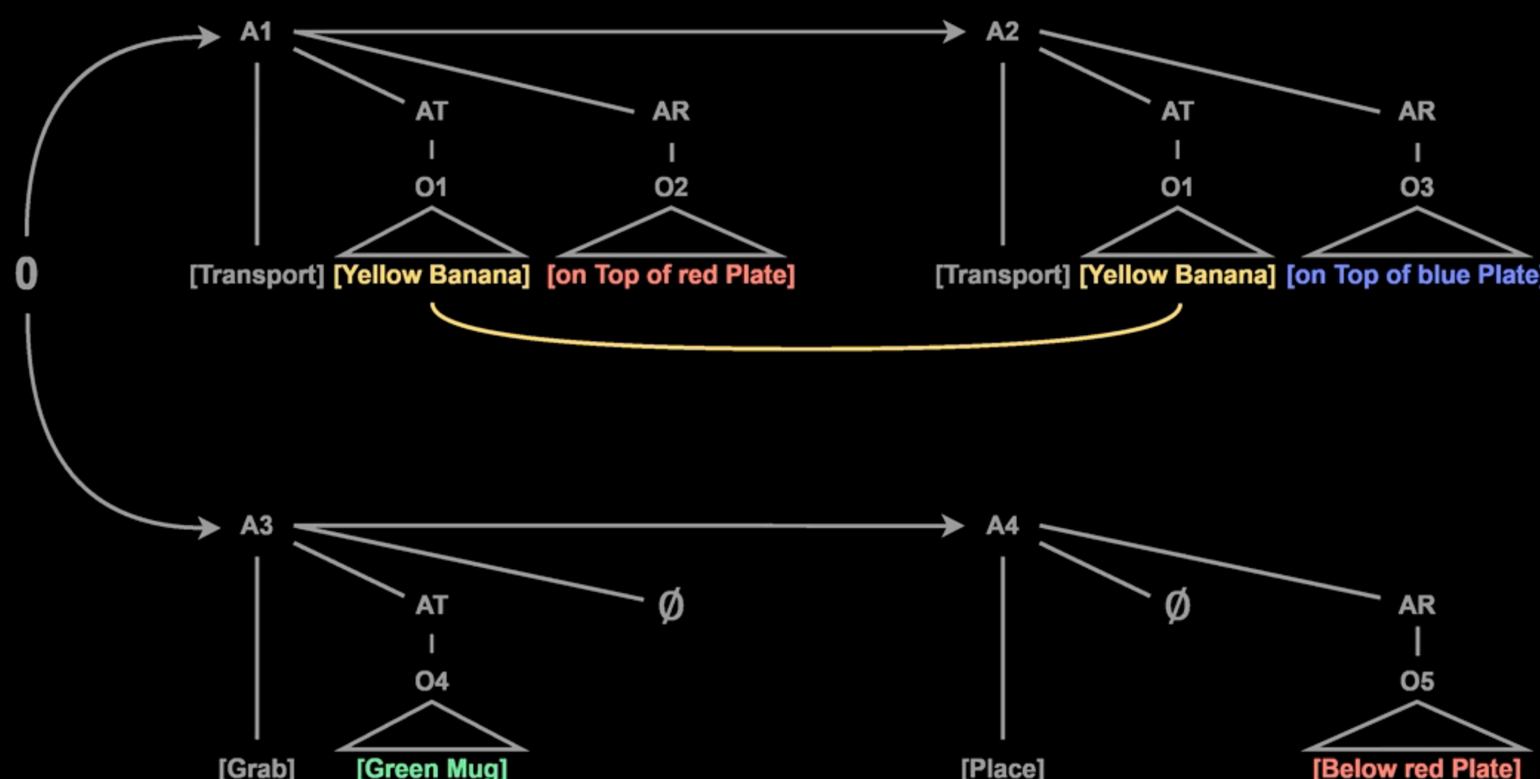
- bind action concepts to object concepts in given roles through in a joint representation of
  - ordinal object concept index
  - ordinal action concept index
  - roles





Conceptual Structure
Action CoS
Color Concepts
Color Cue
Shape Concepts
Shape Cue
Shape Memory
Relation Concepts
Action Concepts
Action Target
Action Reference
Motor Representations
Transport
Grab
Place
Reach
Close
Open
Index Memory
Height Memory
Reach Target
Memory Guidance
Reach Target
Scene Memory
Spatial Binding
Table Surface
Attended Position
Color Memory
Shape Memory
Height Memory
Index Memory
Feature Match
Expected Feature
Match Detection
Attended Feature
CoD
Int
CoS
Coordinate Transforms
Searched Position
Spatial Templates
Attended Position
SM Spatial Binding
Change Detection
Memory Guidance
Perception & Attention
Conceptual Structure
Feature Position
Attended Features
Color
Shape
Height
Color Cue
Shape Cue
Vision sub-network
LIDAR-Image Processing
Color Maps
CNN (VGG16)
Camera Angles
Table Surface
Saccades
Current Fixation
Saccade CoS
Saccade Map
Legend
= Neural Node
= 1D Neural Field
= 2D Neural Field
= 3D Neural Field
= Truncated Connection
= Excitatory Connection
= Inhibitory Connection
= Gain Field Transform
= Retinal to Allocentric Coordinate Transform
= Gate Field

Demo Action Plan



Summary: Conceptual commitments of DFT

- connections determine function [~connectionism]
- localist representations
- in low-dimensional spaces
- interaction dominated dynamics => attractor states
- instabilities generate decisions, sequences: autonomy
- coupling ~ binding as path to higher cognition

- resources, people
- dynamicfieldtheory.org