Neural circuits for cognition

High capacity codes with modularity; Modular circuits and cognitive flexibility

MIT Course 9.49/9.490

Instructor: Professor Ila Fiete

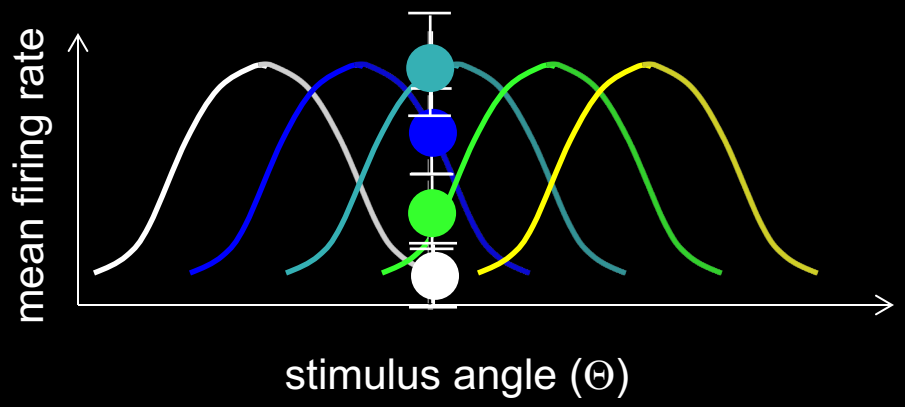
Last time: Coding questions

- Why does the brain have large populations encoding the same variable? E.g. ~2000 HD cells?

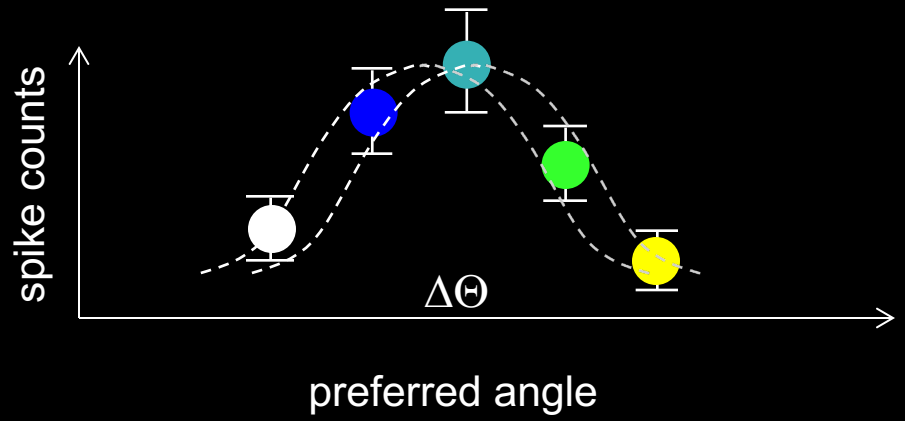
- Given redundancy, can encode information in many possible ways. Why does the brain make a particular choice?

- Given spikes from neurons and an unbiased decoder of an encoded variable, what can I say about the variance/squared error of this decoder?

Population codes enable accurate representation of variables with noisy elements



e.g. V1 orientation tuning.

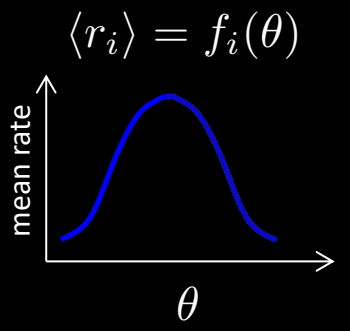


Many neurons firing at each value of variable: smaller total estimation error.

Error control performance

$$\theta \rightarrow \mathbf{r} \rightarrow \hat{\theta}$$

$$MSE \equiv \langle (\hat{\theta} - \theta)^2 \rangle = ?$$



$$r_i = \mathrm{Poiss}(f_i(\theta)\Delta t)$$

Accuracy gain:

$$J \sim N$$

$$MSE \sim \frac{1}{N}$$

Seung & Sompolinsky, 199
Brunel & Nadal, 1998
Abbott & Dayan, 1999
Bethge et al., 2002
Latham et al. 2003

Classical population codes for analog variables:
polynomial gains in accuracy with N.

Examples of classical population codes (CPCs)

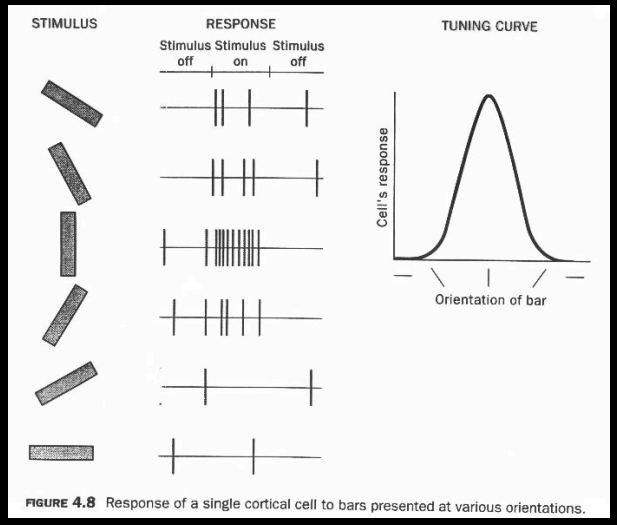
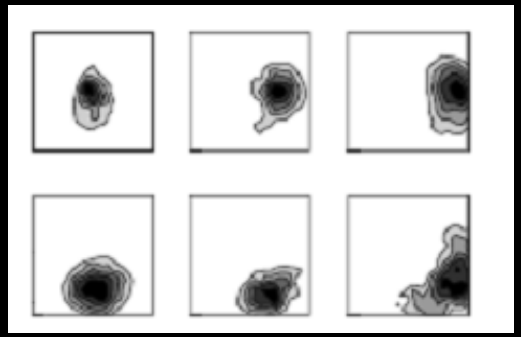


FIGURE 4.8 Response of a single cortical cell to bars presented at various orientations.

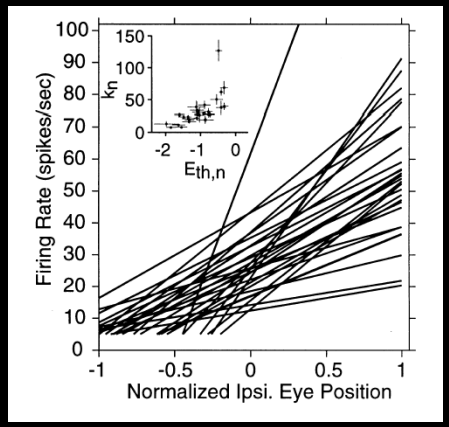
V1 orientation tuning

Hubel & Wiese.



Place cell location tuning

O'Keefe, 1998



Oculomotor eye position tuning

Most codes in sensory, motor peripheries and some cognitive areas are CPCs.

Are neural tuning curves Fisher-optimal? Thalamic HD code

Chaudhuri et al., 2019

Combination of truly 1D coding and finite tuning width suggests that Fisher-optimality is NOT the governing factor in neural tuning.

Is this the best possible use of neural $$?

Information rate: information bits/total bits

$$\rho \sim \frac{I}{N} \sim \frac{\log(N)}{N} \to 0$$

Brunel & Nadal 1998

$I \sim \log(J) \sim \log(N)$

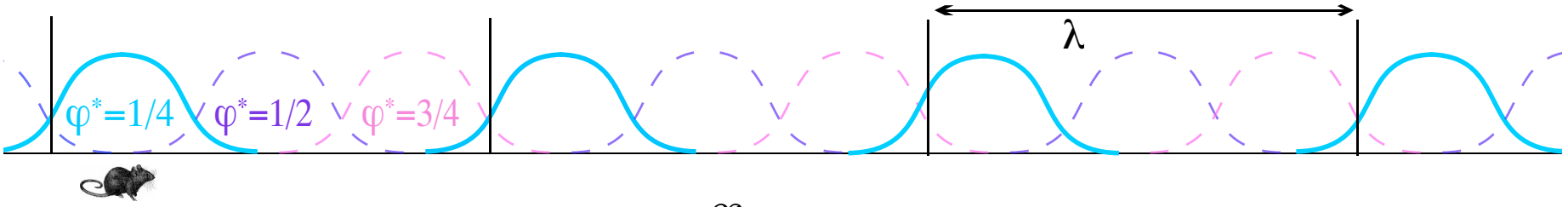
Grid cells: How/what/why?

How are such periodic responses generated from neurons with monotonic nonlinearities?

What are grid cells doing?

Why such a bizzare representation of 2-coordinate position?

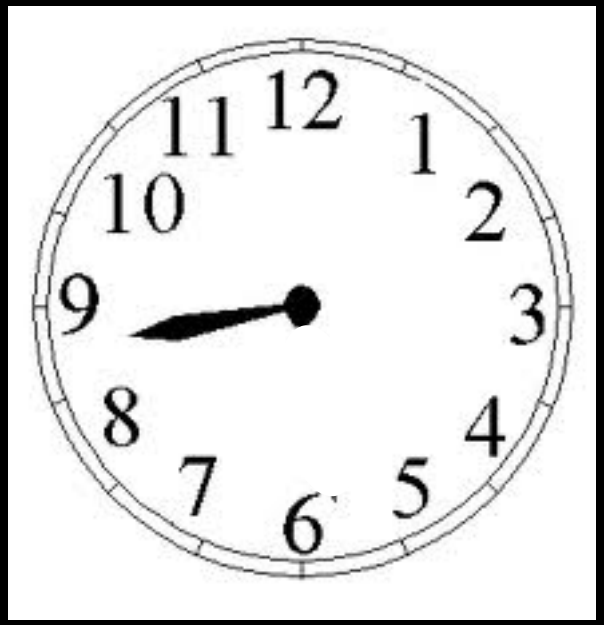
An entire grid cell population encodes location only as a spatial phase

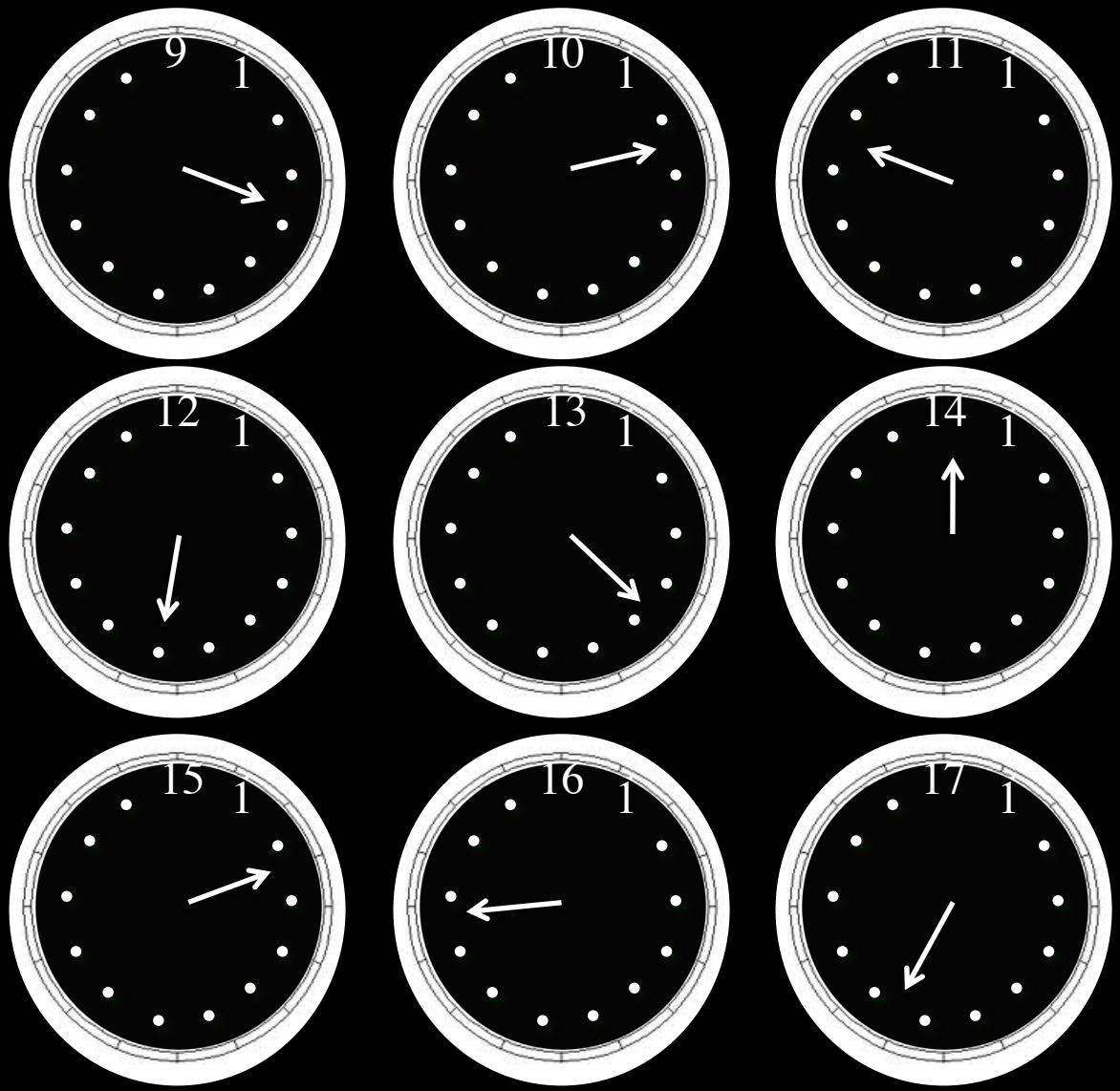


$$\phi = \frac{x \bmod \lambda}{\lambda} = 1/4$$

x

Population code of diverse populations





Counting with modulo remainders

There are certain things whose number is unknown. Repeatedly divided by 3, the remainder is 2; by 5 the remainder is 3; and by 7 the remainder is 2. What will be the number?

Sun Tsu, 4th century AD.

Properties of the code: combinatorial capacity

$$x \rightarrow \vec{\phi} \equiv \{x \bmod \lambda_1, x \bmod \lambda_2, \ldots x \bmod \lambda_N\}$$

idealization: x, λ's whole numbers (dimensionless)

x	$\lambda = (3, 4)$
0	$0, 0$
1	$1, 1$
2	$2, 2$
3	$0, 3$
4	$1, 0$
5	$2, 1$
6	$0, 2$
$\vdots$	$\vdots$
12	$0, 0$
13	$1, 1$

Fiete et al., 2008.

$$R = LCM(\lambda_i) = \prod_i^N \lambda_i \sim \lambda^N$$

Narrow register range

	registers	capacity
decimal	$(10^5, 10^4, 10^3, 10^2, 10^1, 10^0)$	1,000,000
modulo	(1003, 103, 13)	1,343,017
modulo	(18, 17, 16, 15, 14, 13)	1,113,840

decimal		modulo (18, 17, 16, 15, 14, 13)
45	=	(9, 11, 13, 0, 3, 6)
800,000	=	(8, 14, 0, 5, 12, 6)
800,001	=	(9, 15, 1, 6, 13, 7)

Fiete, Burak, Brookings 2008.

Network parameters need not span several orders of magnitude. “Whitened”. Lesions of any lattice equal.

Carry-free arithmetic

decimal $(10^2, 10^1, 10^0)$

```
  1 1
    9 7
  +   4
-------
  1 0 1
```

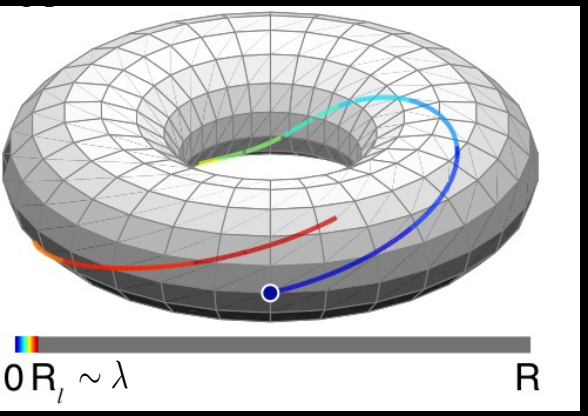
modulo (7,6,5)

```
    6 1 2
+   4 4 4
---------
    3 5 1
```

Fiete, Burak, Brookings 2008.

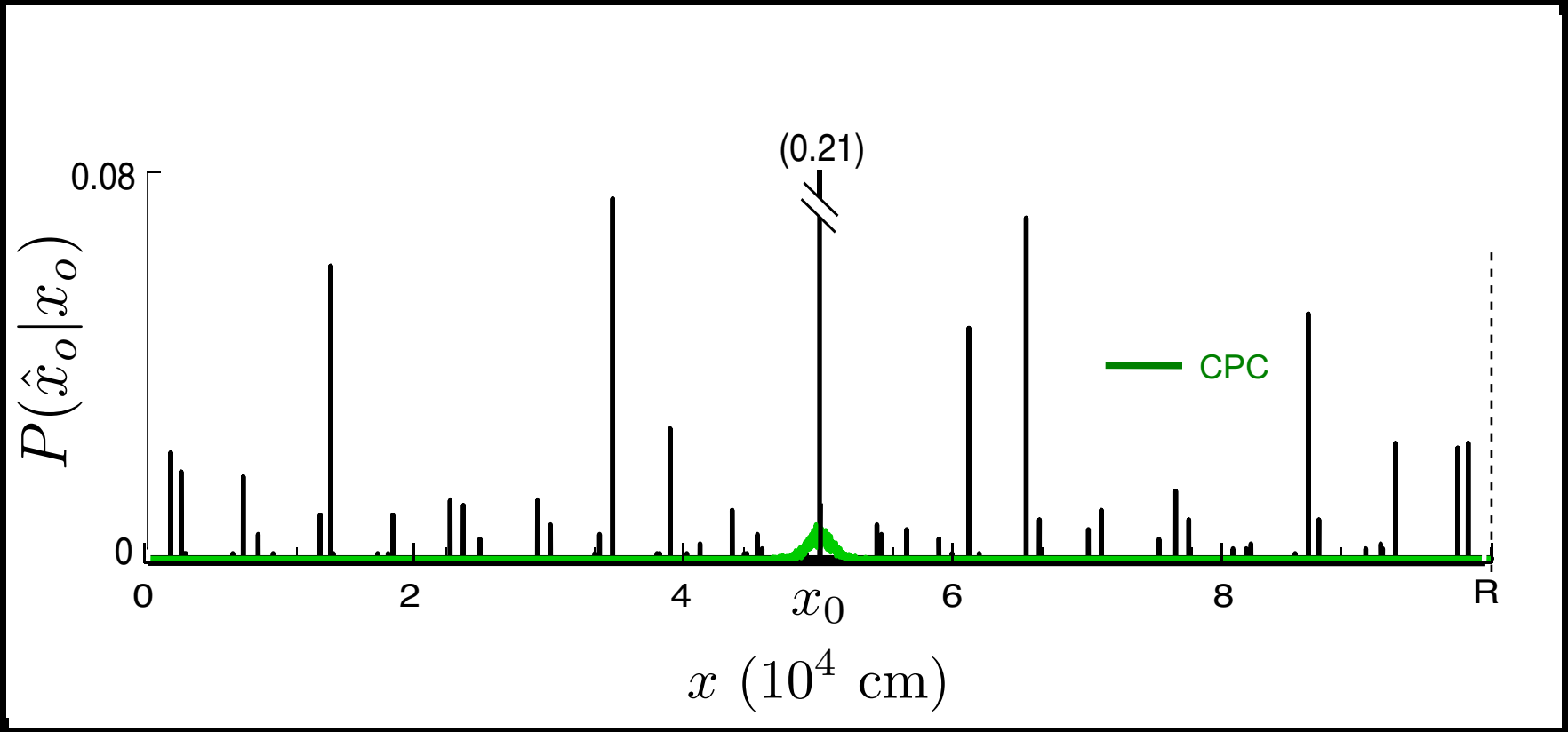
Parallel (non-serial) updating across moduli: speedup.
No information carry across modules (biological advantage).

Unique representation for exponentially large range of locations



Fiete et al. 2008.

Exquisite noise sensitivity over full range R

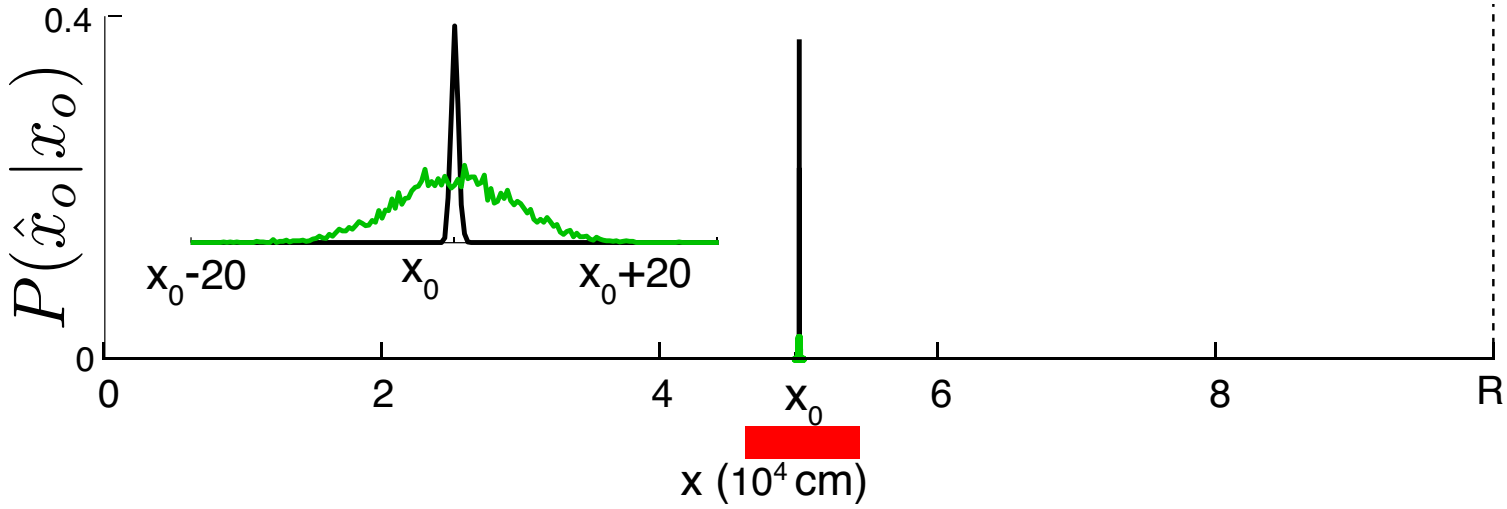


Small independent errors per phase → massive position errors.

Superior error control over range << R

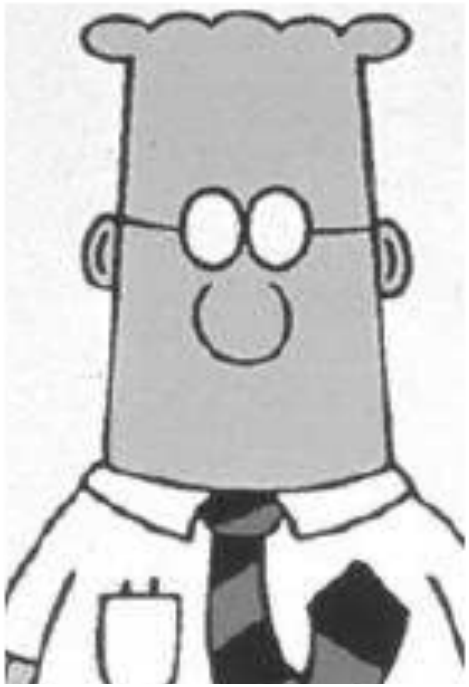
$$R_l \ll R$$

$$(\sim \lambda e^{\rho\beta N}) \quad (\sim \lambda e^{\beta N})$$

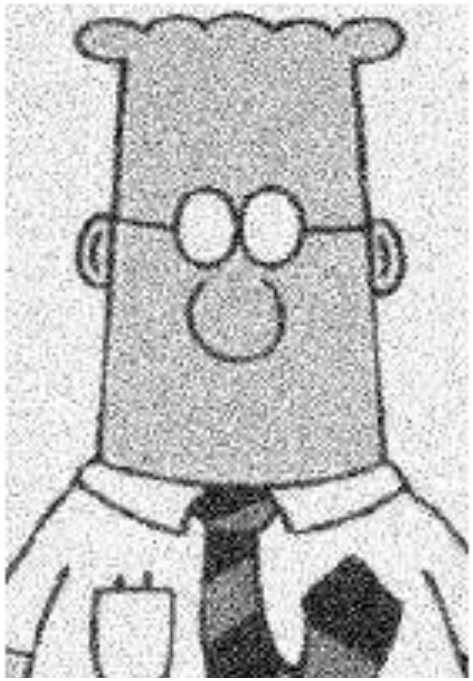


$$\text{information rate} :\sim \frac{\log R_l/\lambda}{\log R/\lambda} = \rho$$

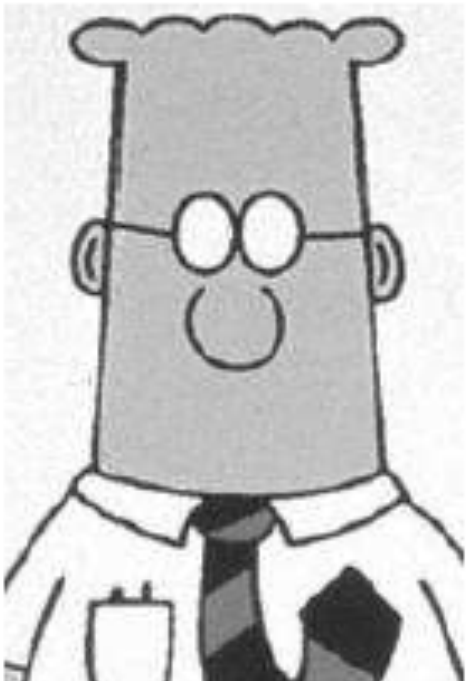
Superior error control over range << R



original



CPC posterior (20 iterations)

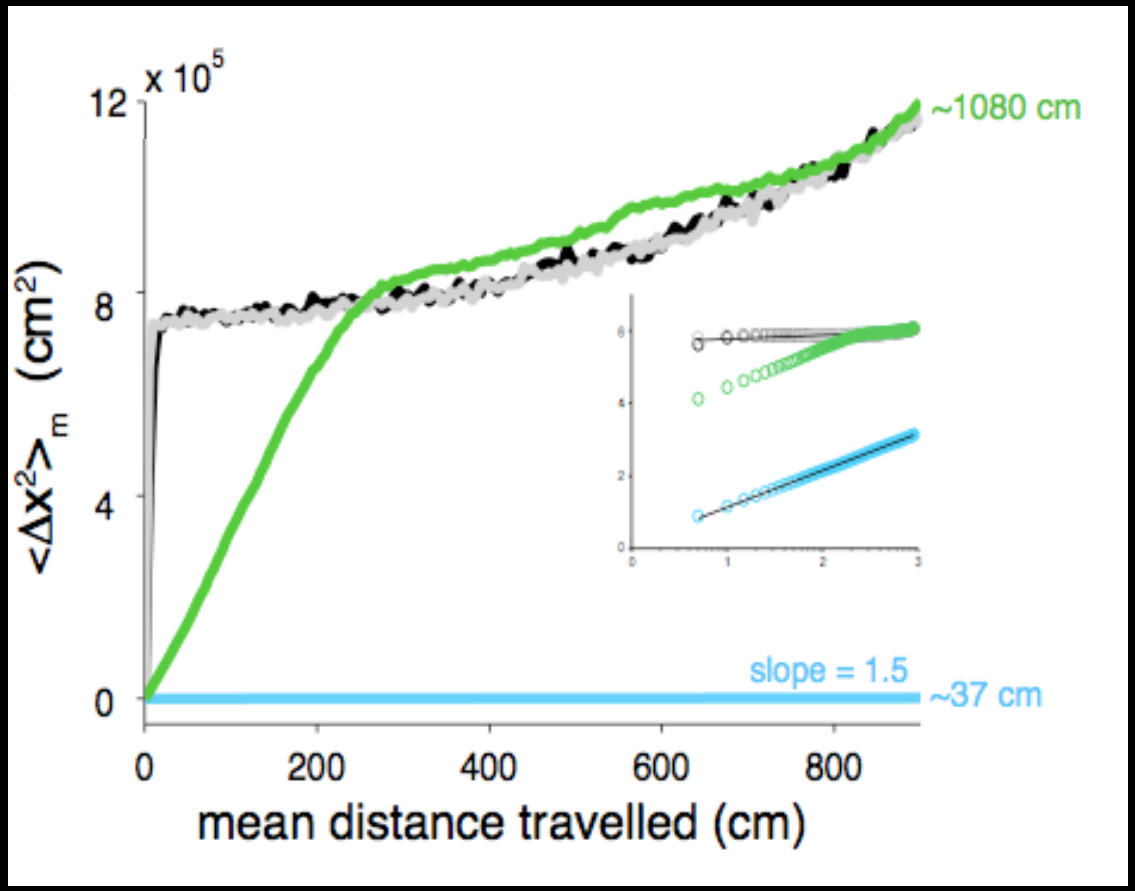


GC posterior (20 iterations)

$$\text{information rate} :\sim \frac{\log R_l/\lambda}{\log R/\lambda} = \rho$$

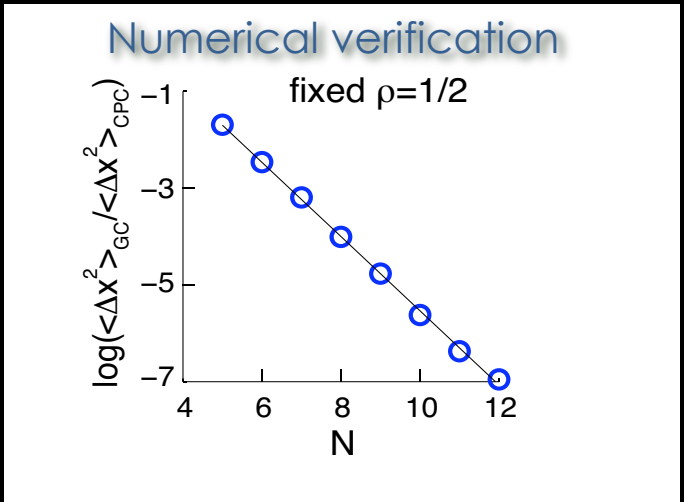
Sreenivasan & Fiete 2011

Iterative error correction during path integration **without external cues**



Very slow error accumulation. >10^4-fold improvement over CPC.

How much better than CPCs?

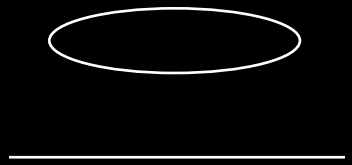


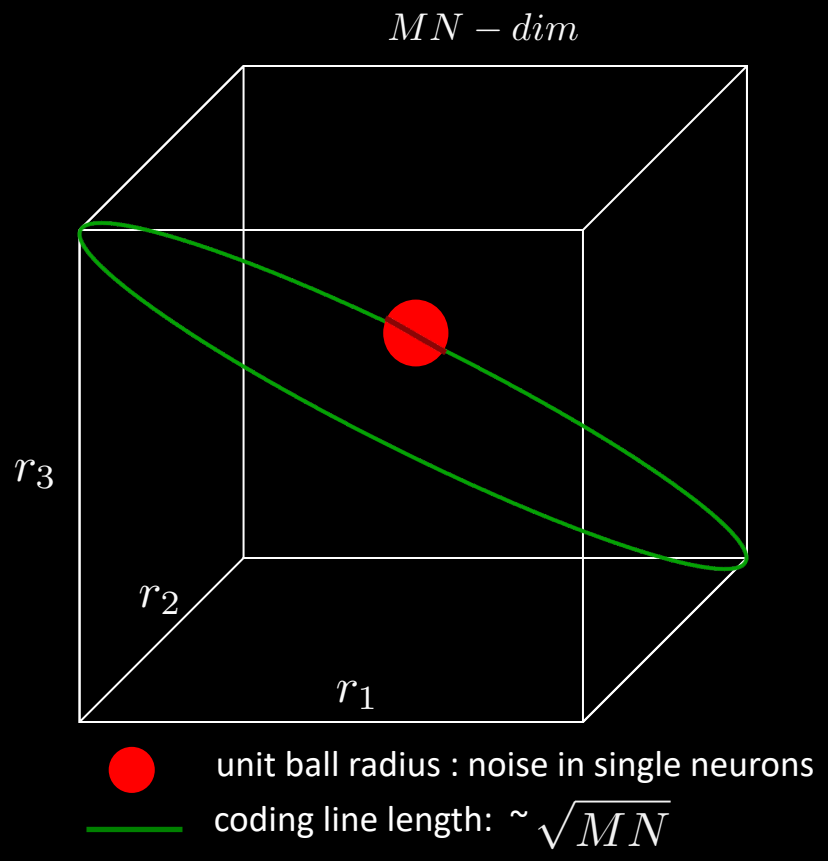
$$\frac{||\Delta x||_{GC}}{||\Delta x||_{CPC}} \sim e^{-\rho\beta N}\sqrt{N}$$

Sreenivasan & Fiete 2011
Mathis et al. 2012

Grid code enables *exponentially better* estimation accuracy than CPCs *at finite information rate,* ρ.

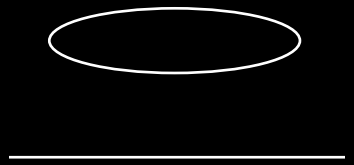
CPC: low-dimensional coding line embedded in high-dimensional state space enables error reduction

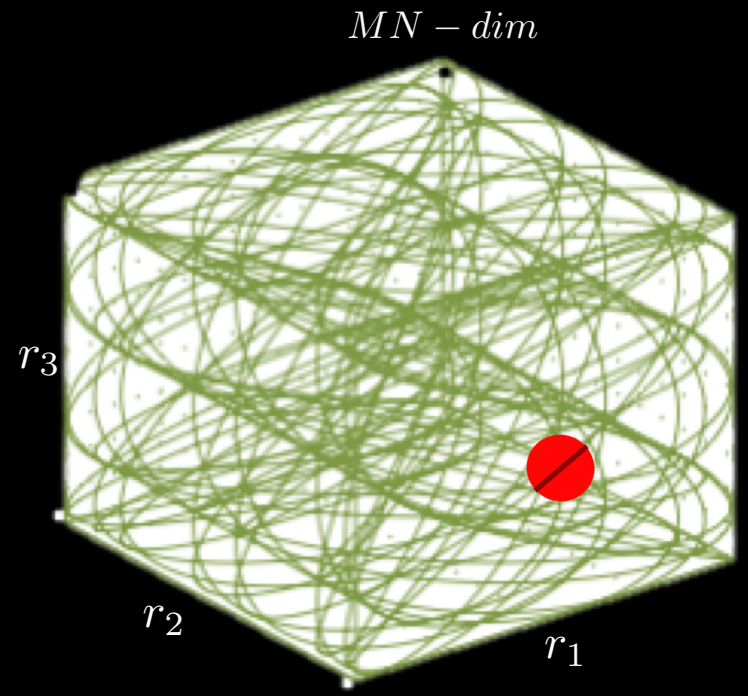




$$fractional \;\; error \sim 1/\sqrt{MN}$$

GC: low-dimensional coding line *stretched out* and embedded in high-dimensional state space *at higher density* enables stronger error-correction



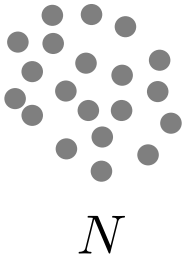


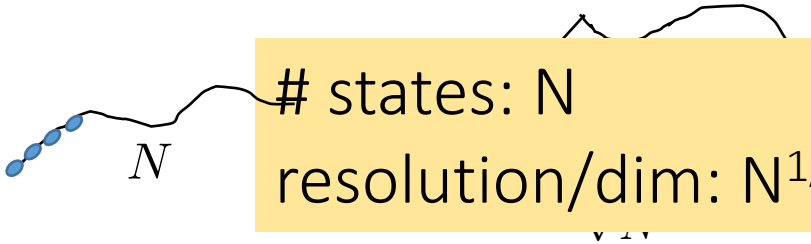
unit ball radius : noise in single neurons

coding line length: ~ $e^{\rho\beta N}$

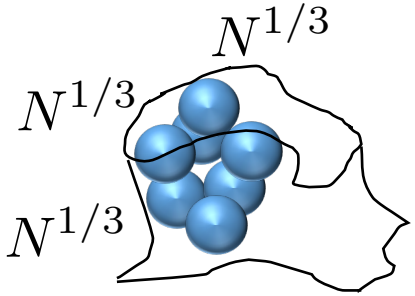
$$fractional\ \ error\ \sim e^{-\rho\beta N}$$

Non-modular: vary dimension (fixed range)

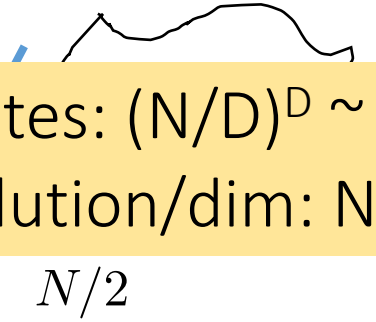




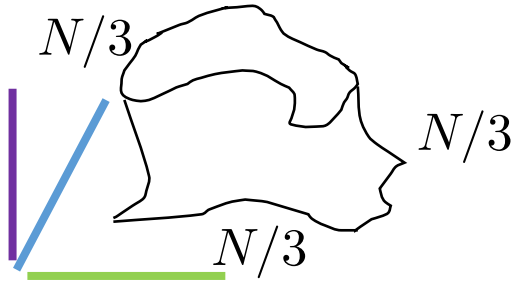
states: N
resolution/dim: $N^{1/D}$



Modular

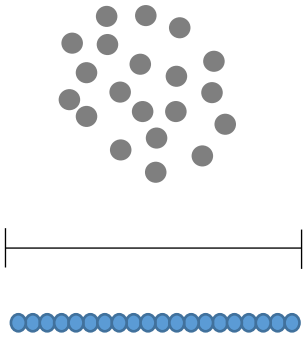


states: $(N/D)^D \sim N^D$
resolution/dim: N/D

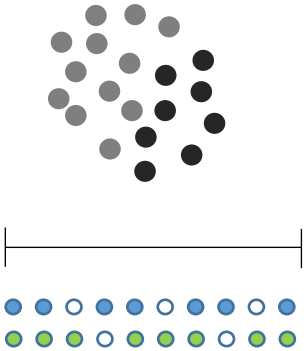


Modularity can solve the “curse of dimensionality” in representation

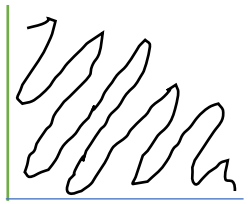
Modularity and range (fixed dimension)



$R \sim N$

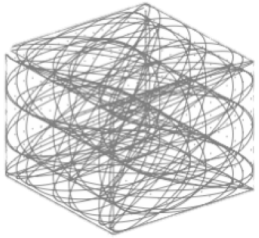


$R \sim (N/2)^2$



$\cdots$

$R \sim (N/M)^M$



range: exponential in # modules

Fiete et al., 2008;
Sreenivasan & Fiete 2011

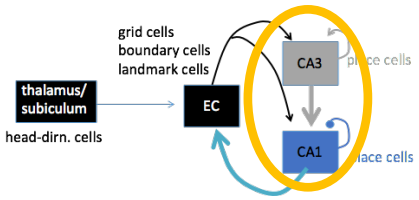
Summary

• Appropriately modular codes enable exponential coding capacity.

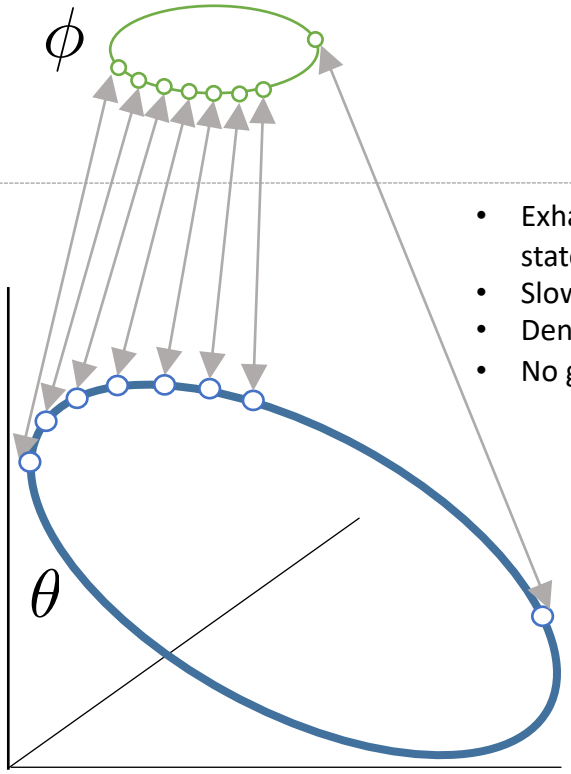
• Grid cell codes are both locally Euclidean, metric-preserving, and globally update based on velocity cues to generate a vast library of consistent new codewords as one moves about the world.

Flexibility beyond spatial and 2D representations

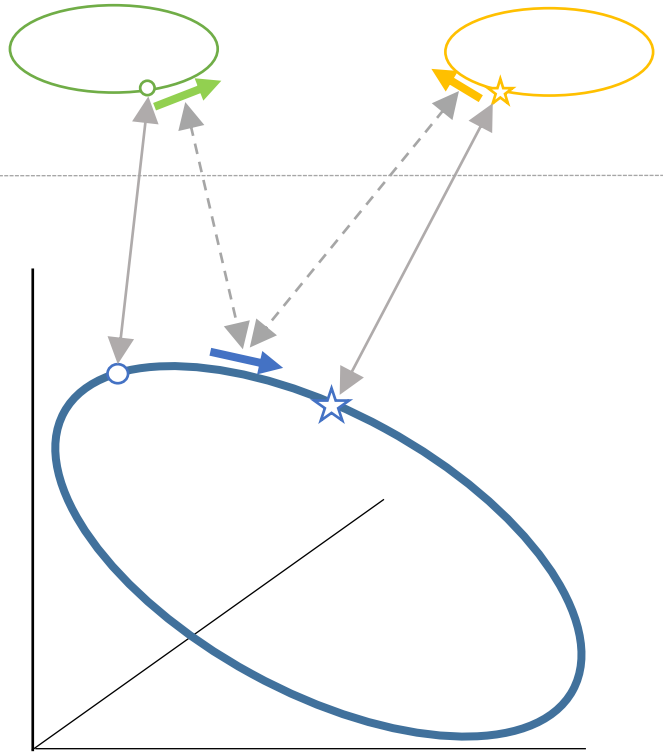
Same brain/circuits, different behaviors



Beyond space: Flexible/fast representation of any metric/locally Euclidean space of matching or smaller dimension



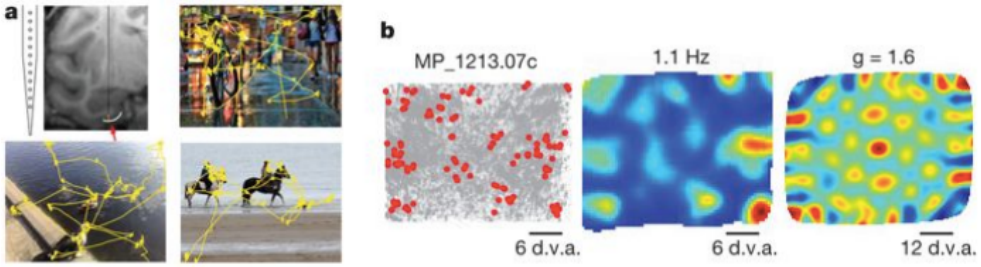
- Exhaustive mapping, state-by-state
- Slow
- Dense ff connectivity
- No generalization



Efficient “zero-shot” process: traverse one piece of ϕ, get self-consistent on-the-fly codes for rest

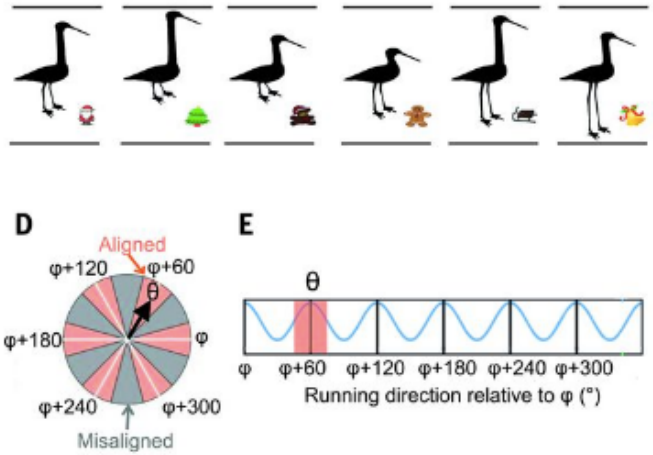
Hypothetical mechanism for demonstrated representation of non-spatial variables

Gaze location (monkey):



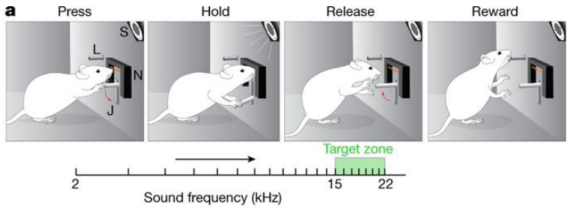
Killian et al. 2012

Abstract 2D parametric space (human):

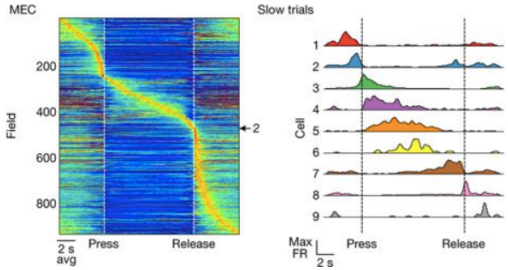


Constantinescu et al. 2016

Auditory tone space (rat):



Aronov et al. 2017



Beyond space: Flexible representation of several metric/locally Euclidean space of equal or smaller dimension

Coupled with velocity inputs *from any metric spaces of equal or smaller dimension*, can represent those metric space by integrating movements or displacements through those spaces.

No reconfiguration of recurrent network.

Spontaneously produces a code for all values of that variable, even if many values of the variable have not been previously visited.

Summary

• Canonical neural circuits with low-dimensional dynamics can perform a range of integration, memory, and decision-making tasks.

• They can be used and reused flexibly.

• Combining multiple such circuits into appropriately modular codes enables high (exponential), large dictionary of on-th—fly coding and memory states.

• These are some of the neural circuit motifs underlying cognitive computation.