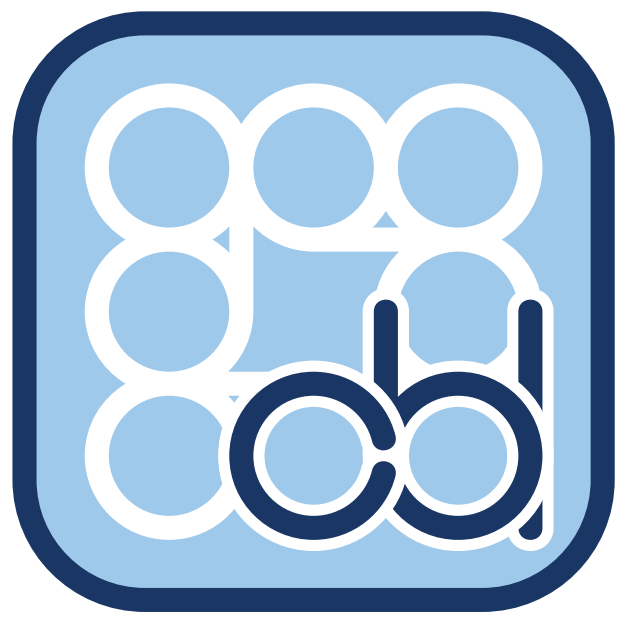# A theory of memory stability in hippocampal area CA3

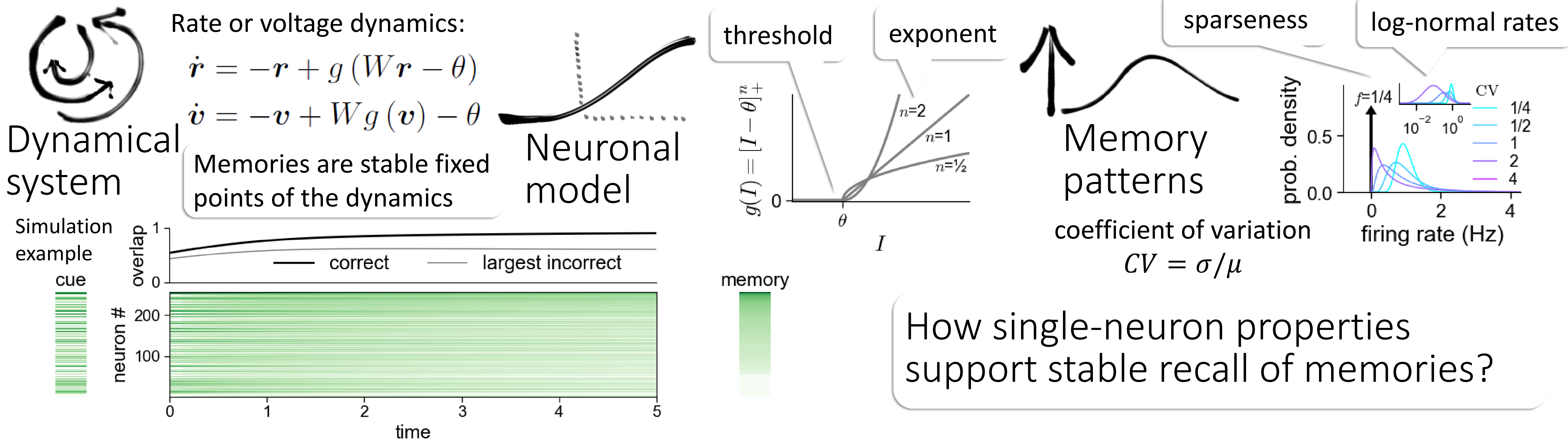Uri Cohen[1], Roland Mason Rodriguez[2], Ole Paulsen[2], Máté Lengyel[1,3]

[1] Computational and Biological Learning Lab, Department of Engineering, University of Cambridge; [2] Department of Physiology, Development and Neuroscience, University of Cambridge; [3] Department of Cognitive Science, Central European University

@uricohen42 𝕏
uc231@cam.ac.uk

Get poster!

Hippocampal area CA3 was suggested to implement an auto-associative memory device [1-6].
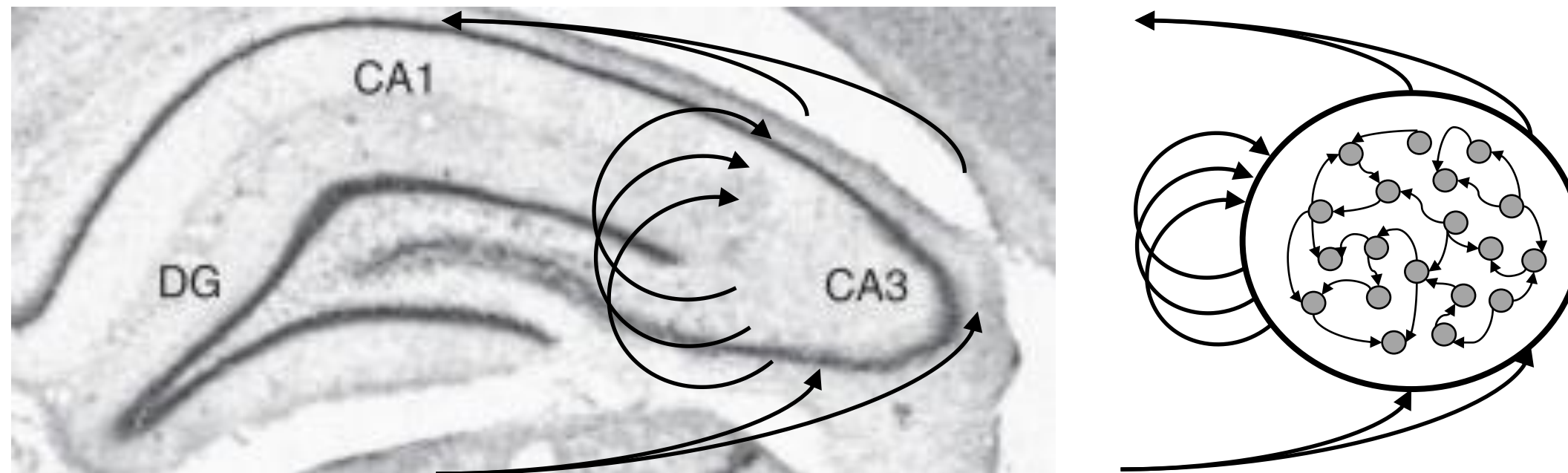
We provide a new analysis of the conditions for high-capacity storage of stable memories using biological-relevant recall dynamics.

Dynamical system

Rate or voltage dynamics:

$$\dot{r} = -r + g\left(Wr - \theta\right)$$
$$\dot{v} = -v + Wg\left(v\right) - \theta$$

Memories are stable fixed points of the dynamics

Neuronal model

threshold   exponent

$$g(I) = [I - \theta]^n_+$$

$n=2$
$n=1$
$n=\frac{1}{2}$

Memory patterns

sparseness   log-normal rates

coefficient of variation
$$CV = \sigma/\mu$$

Simulation example

How single-neuron properties support stable recall of memories?

Mean-field theory

optimising weights for min-norm, thresholds for stability (as in [12])

Mean-field theory (à la [13-15]) fully describe the ability to store patterns as fixed-points.

Storing P memory patterns of N neurons as fixed points: $$r^\mu = g\left(Wr^\mu - \theta\right)$$
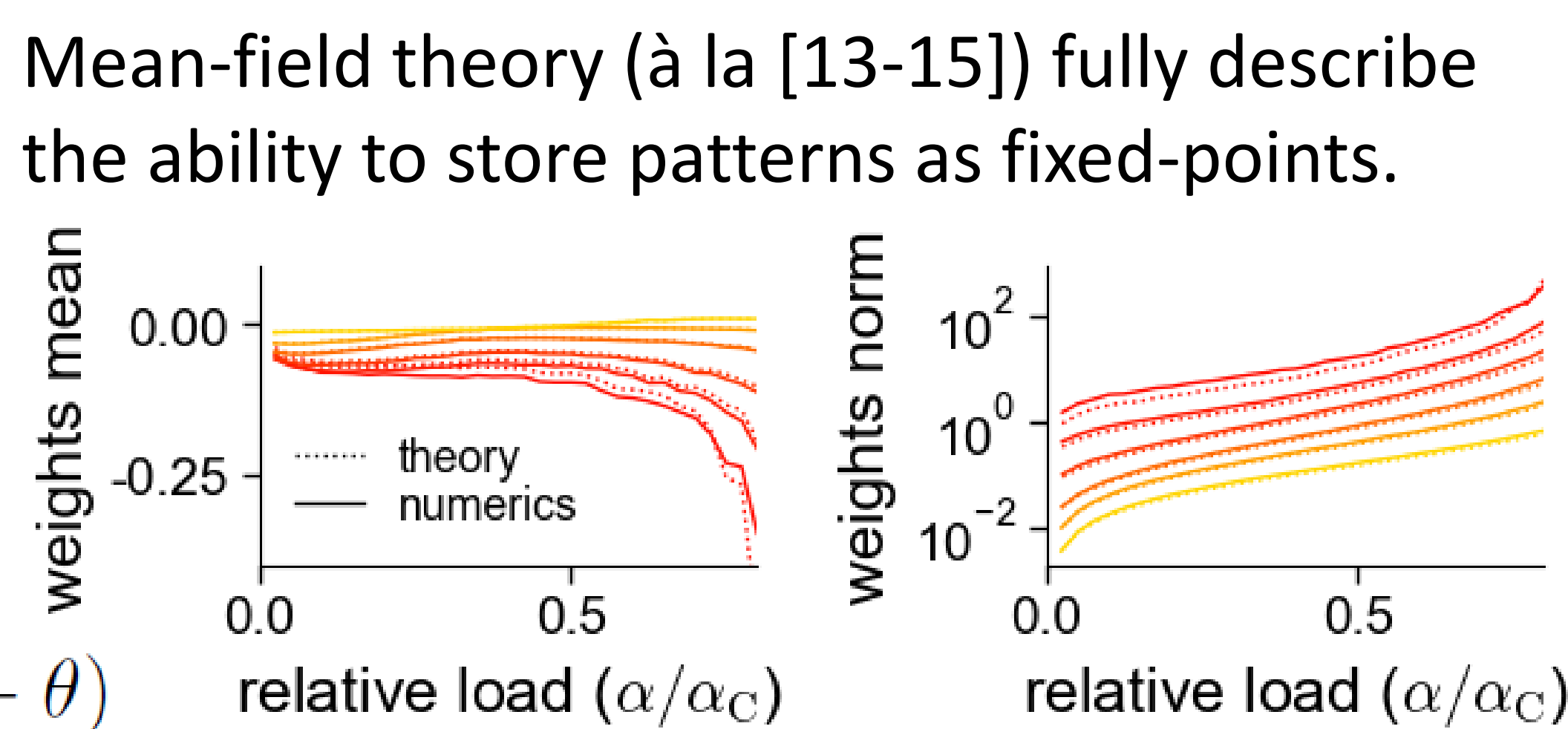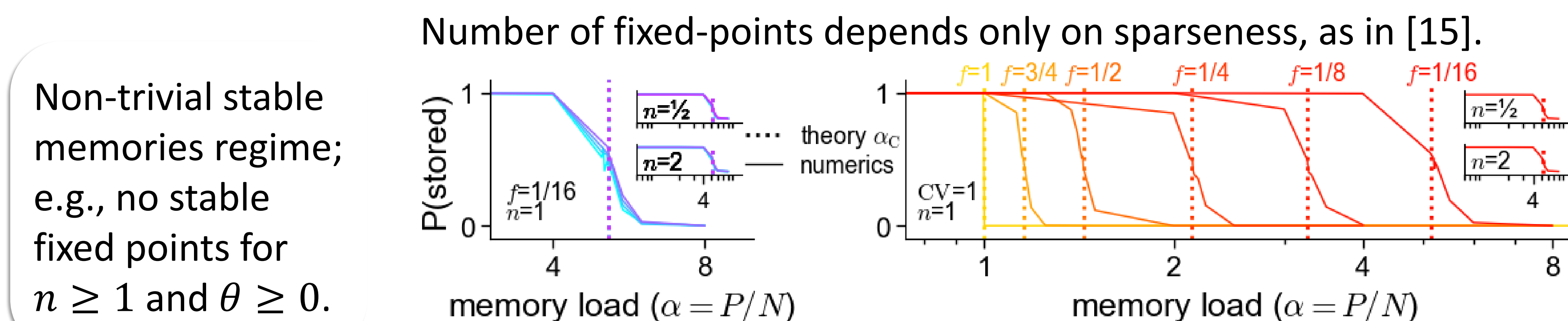
Technical novelty

The Hopfield model [7] and various extensions [8-11] all use *gradient-dynamics* to store stable memories in attractor networks: $\dot{r} \propto -\nabla E$
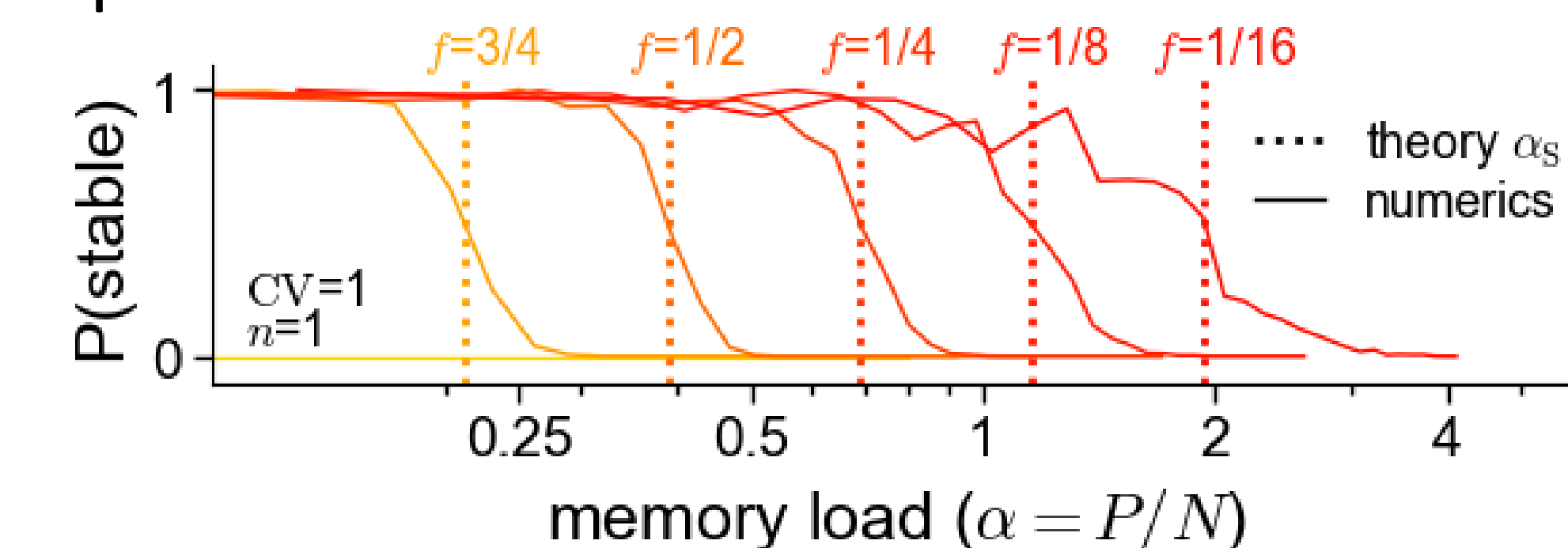✓ stable low energy memories
✗ non-biological dynamics

Our attractor networks has:
✗ possibly unstable memories
✓ biological plausible dynamics experimentally testable predictions are derived from the regime where memories tend to be stable
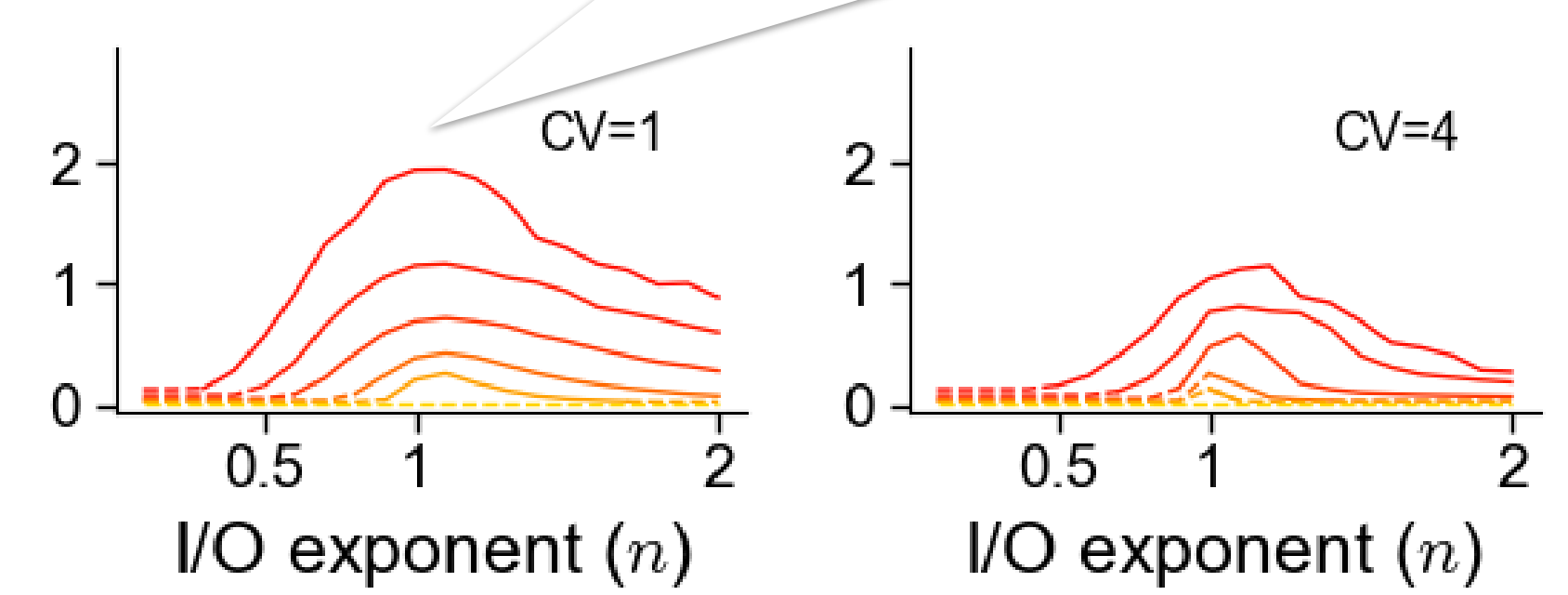
Number of fixed-points depends only on sparseness, as in [15].

Non-trivial stable memories regime; e.g., no stable fixed points for $n \geq 1$ and $\theta \geq 0$.

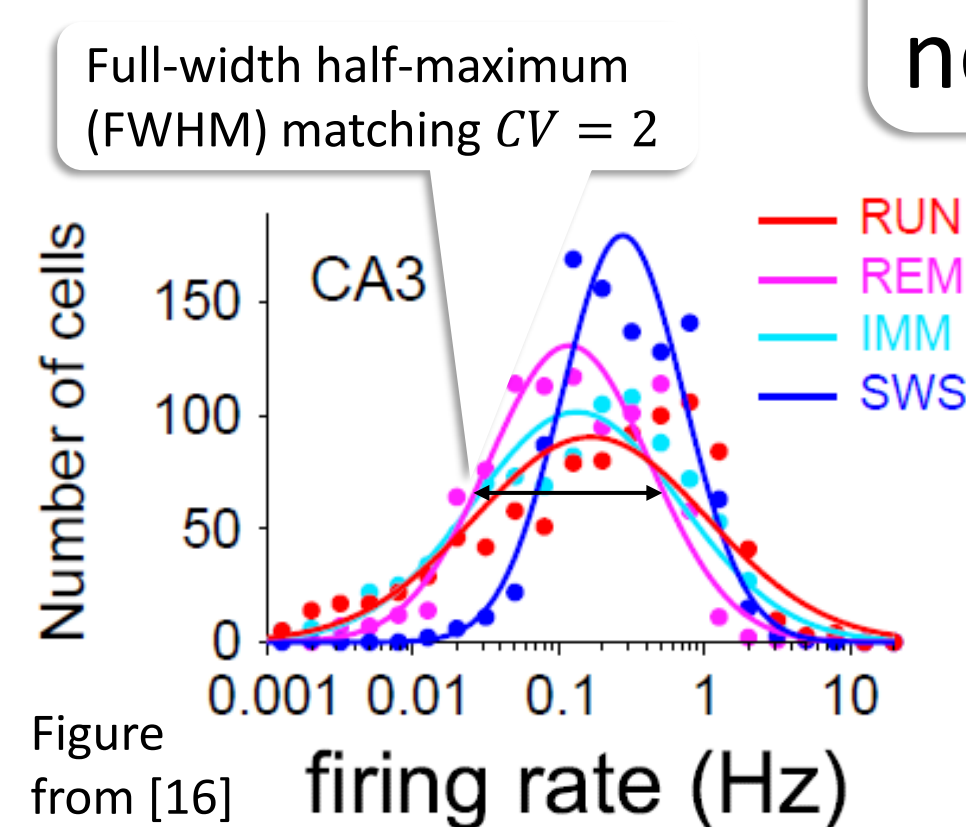In the stable regime, memories are all stable up to a critical load.

At low firing rate CV, a sublinear f-I curve is optimal
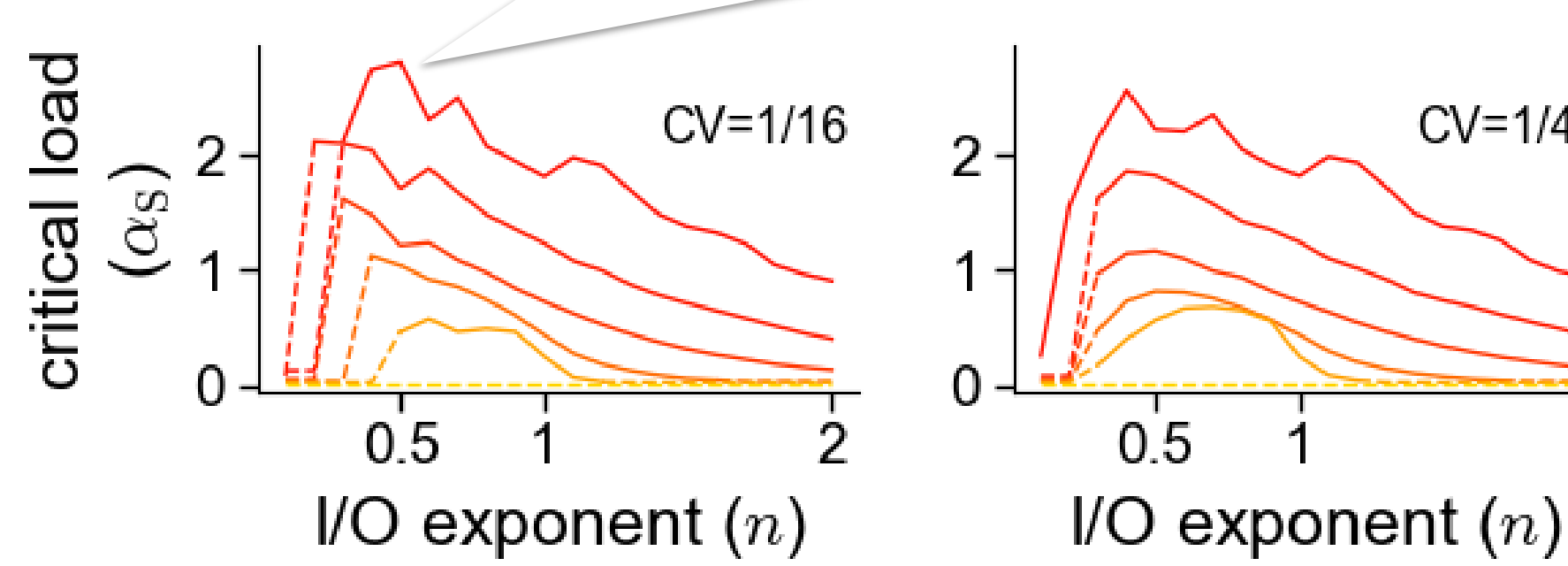
At high firing rate CV, a linear f-I curve is optimal

For high-variation memory patterns, CA3 pyramidal neurons should have near-linear f-I curves, negative threshold (should spontaneously fire if recurrent connections are suppressed)

Analysing existing data [16-17] we estimate $CV \geq 2$ in pyramidal neurons of area CA3.

LFP patch-clamp
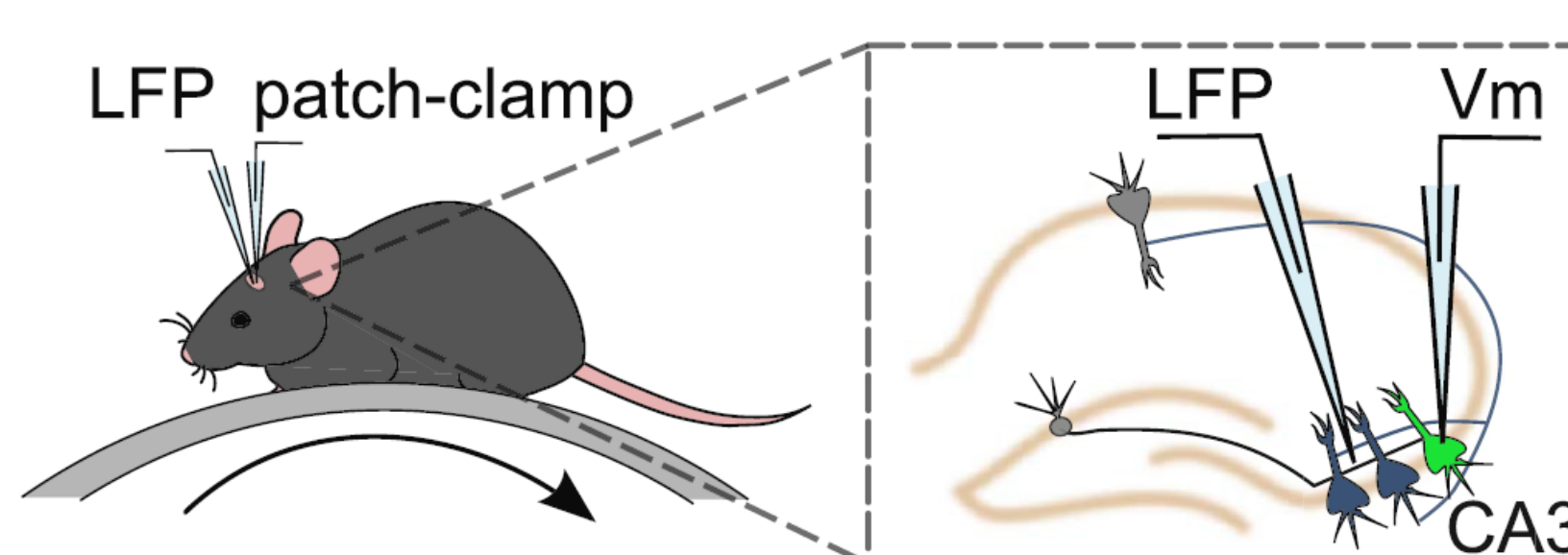
Recording activity from mice CA3 pyramidal neurons in in-vitro whole-cell experiments, we find near-linear f-I curves, and similar results in existing in-vivo data [18].

Firing rate statistics

Firing rate non-linearity
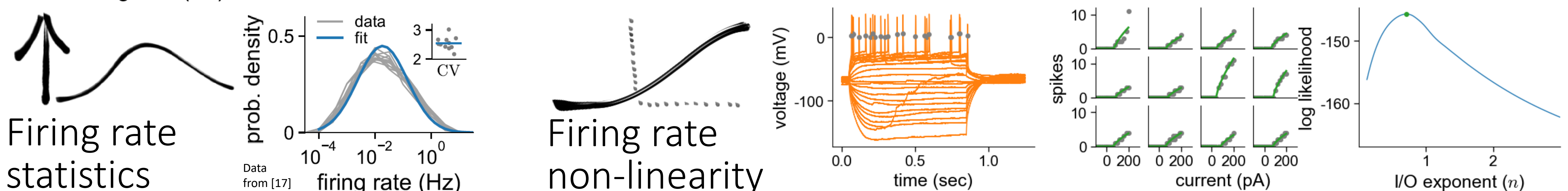
References: [1] Marr D (1971) Simple memory: a theory for archicortex. Philos. Trans. R. Soc. Lond. B Biol. Sci. 262:23-81. [2] McNaughton BL, Morris RG (1987) Hippocampal synaptic enhancement and information storage within a distributed memory system. Trends Neurosci. 10:408-15. [3] Treves A, Rolls ET (1994) Computational analysis of the role of the hippocampus in memory. Hippocampus 4:374-91. [4] Nakazawa K, Quirk MC, Chitwood RA, Watanabe M, Yeckel MF, Sun LD., ... Tonegawa S (2002) Requirement for hippocampal CA3 NMDA receptors in associative memory recall. Science, 297:211-8. [5] Wills TJ, Lever C, Cacucci F, Burgess N, O'Keefe J (2005) Attractor dynamics in the hippocampal representation of the local environment. Science 308:873-6. [6] Khona M, Fiete IR (2022) Attractor and integrator networks in the brain. Nat. Rev. Neurosci. 23:744-66. [7] Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. USA 79:2554-8. [8] Tsodyks MV, Feigel'man MV (1988) The enhanced storage capacity in neural networks with low activity level. EPL 6:101. [9] Treves A (1990) Graded-response neurons and information encodings in autoassociative memories. Phys. Rev. A 42:2418. [10] Krotov D, Hopfield JJ (2016) Dense Associative memory for pattern recognition. Adv. Neural Inf. Process. 29. [11] Lengyel M, Kwag J, Paulsen O, Dayan P (2005) Matching storage and recall: hippocampal spike timing-dependent plasticity and phase response curves. Nat. Neurosci., 8:1677-83. [12] Festa D, Hennequin G, Lengyel M (2014) Analog memories in a balanced rate-based network of EI neurons. Adv. Neural Inf. Process. 27. [13] Gardner E (1988) The space of interactions in neural network models. J. Phys. A Math. Theor. 21:257. [14] Amit DJ, Gutfreund H, Sompolinsky H (1985) Storing infinite numbers of patterns in a spin-glass model of neural networks. Phys. Rev. Lett. 55:1530. [15] Schönsberg F, Roudi Y, Treves, A (2021) Efficiency of local learning rules in threshold-linear associative networks. Phys. Rev. Lett. 126:018301. [16] Mizuseki K, Buzsáki G (2013). Preconfigured, skewed distribution of firing rates in the hippocampus and entorhinal cortex. Cell reports, 4:1010-21. [17] Alme CB, Miao C, Jezek K, Treves A, Moser EI, Moser MB (2014). Place cells in the hippocampus: eleven maps for eleven rooms. Proc. Natl. Acad. Sci. USA, 111:18428-35. [18] Malezieux M, Kees AL, Mulle C (2020). Theta oscillations coincide with sustained hyperpolarization in CA3 pyramidal cells, underlying decreased firing. Cell Reports 32.

BLAVATNIK CAMBRIDGE FELLOWSHIPS
wellcome
BRITISH COUNCIL