King's College London
University of London

Representation and processing of novel stimuli by an associative neural network

Reimer Kühn

Department of Mathematics, King's College London

in collaboration with

Valeria del Prete

http://www.mth.kcl.ac.uk/staff/r_kuehn.html

Overview

- Motivation

- The Model

- Collective Properties

  - Firing Rate Distributions

  - Stimuli, Mutual Information, Correlation

- Perspective:
  - Optimise Representations
  - Dynamics

Motivation

- Representation of stimuli in an associative neural networks

  — neural firing patterns, rates and their distributions

- Of interest: differences in firing rate distributions, if new stimulus is correlated with pre-learnt patterns or not.

- For new stimuli correlated with pre-learnt patterns:
  - mutual information between firing rates and novel stimulus
  - dependence on correlation with pre-learnt patterns
  - dependence on other parameters
    (number of memories, thresholds, neural gain-function)

- Theoretical framework for interpretation of recordings using trained-untrained scenarios?
  - untrained animals represent stimuli in existing cognitive structure
  - coginitive structure changes in response to new stimuli
- Results – at least in principle – experimentally accessible

Context

- A Treves et al (Neural Computation, 1999)
  Recordings from inferior temporal cortex, visual stimuli
  rate-distributions non-exponential, fits to assumed current distributions

- N Brunel (J. Comp. Neurosci., 2000)
  Dynamics of sparsely connected networks of leaky IF neurons with uniform synaptic strengths: study of collective network states

- J Hertz et al (Neurocomputing, 2003); q-bio.NC/0402023
  Dynamics of sparsely connected networks of leaky IF neurons with uniform synaptic strengths: computation of Fano factors

The Model

- Graded response neurons, Kirchhoff equations for coupled leaky integrators

$$C_i \frac{dU_i}{dt} = -\frac{U_i}{R_i} + \sum_{j=1}^{N} J_{ij}\, \nu_j + I_i$$

- Firing rates ν_i via voltage-to-rate transduction-function

$$\nu_i = g(U_i - \vartheta_i)$$

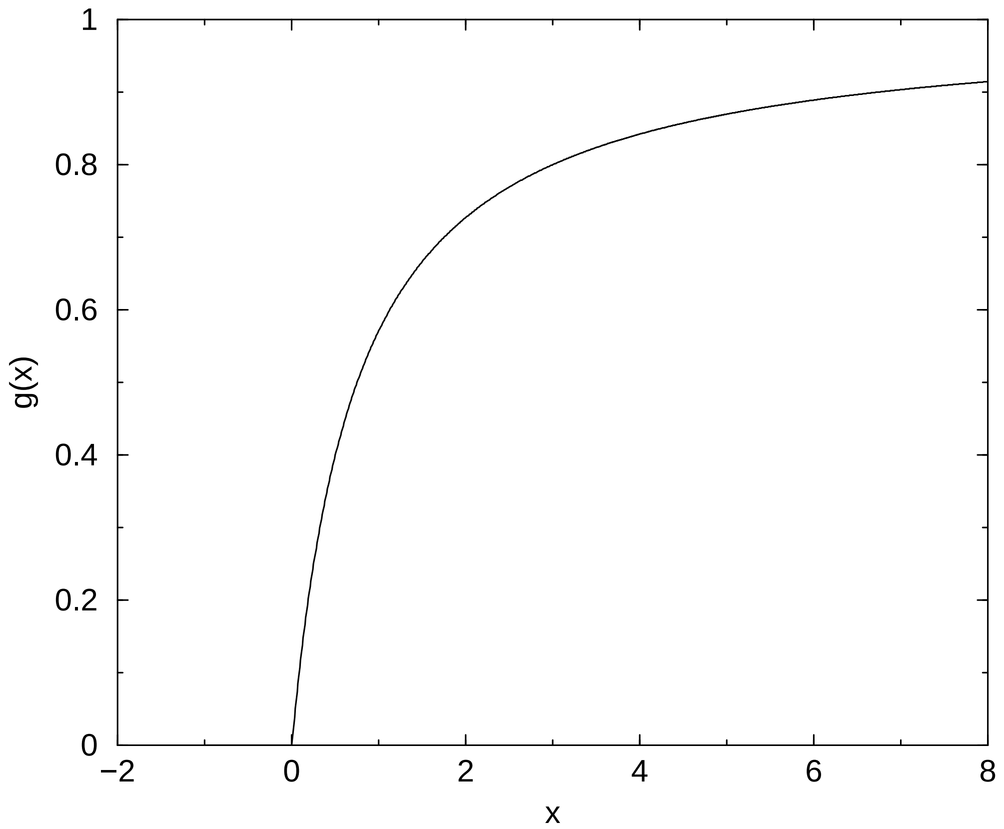
- Synaptic couplings from Hebbian covariance learning rule

$$J_{ij} = \frac{1}{Na(1-a)} \sum_{\mu=1}^{p} (\eta_i^\mu - a)(\eta_j^\mu - a)$$

$$\eta_i^\mu = \begin{cases} 1; & \text{with prob. } a, \\ 0; & \text{with prob. } 1-a, \end{cases} \qquad \text{hence :} \quad \langle \eta_i^\mu \rangle = a$$

- Voltage-to-rate transduction-function for present setup:

$$g(x) = \nu_{\text{max}} \, \frac{x}{U_0 + x} \, \Theta(x)$$



Voltage-to-rate transduction-function, $x = U - \vartheta$, $U_0 = 0.75$, $\nu_{\text{max}} = 1$.

- Present talk: only long time stationary response

- Symmetric couplings: dynamics governed by Lyapunov function

$$H_N(\nu) = -\frac{1}{2} \sum_{i,j=1}^{N} J_{ij}\nu_i\nu_j + \sum_{i=1}^{N} G(\nu_i) - \sum_{i=1}^{N} (I_i - \vartheta_i)\nu_i$$

  with

$$G(\nu) = \int^{\nu} d\nu' g^{-1}(\nu')$$

  $\Longrightarrow$ stationary response from minima of H_N

- Characterisation of attractors (minima of H_N):

  $T \rightarrow 0$-limit of free energy corresponding to H_N

  (RK, S. Bös J. Phys. A, 1993)

Collective Properties

- Stationary limit from equilibrium statistical mechanics.
  - Partition function

$$Z_N = \int \prod_i d\nu_i \,\mathrm{exp}[-\beta H_N(\nu)]$$

  - Free energy

$$f_N(\beta) = -(\beta N)^{-1} \log Z_N$$

  - $T \to 0 \Leftrightarrow \beta \to \infty$-limit: only minima of H_N contribute.
  - A technical point: randomness due to $\{\eta_i^\mu\}$ $\quad \Rightarrow \quad \overline{f_N(\beta)}$
  - Macroscopic characterization of system — order parameters:

$$m^\mu = \frac{1}{N} \sum_i \frac{\eta_i^\mu - a}{a(1-a)} \overline{\langle \nu_i \rangle}\,, \quad q = \frac{1}{N} \sum_i \overline{\langle \nu_i \rangle^2}\,, \quad c = \frac{\beta}{N} \sum_i [\overline{\langle \nu_i^2 \rangle} - \overline{\langle \nu_i \rangle^2}]$$

- Self-consistency equations, $T = 0$-limit:

$$m = \left\langle\!\!\left\langle \frac{\eta-a}{a(1-a)}\,\hat{\nu} \right\rangle\!\!\right\rangle_{\eta,z,I}$$

$$c = \frac{1}{\sqrt{\alpha r}} \left\langle\!\!\left\langle z\hat{\nu} \right\rangle\!\!\right\rangle_{\eta,z,I} \qquad\qquad r = \frac{q}{(1-c)^2}$$

$$q = \left\langle\!\!\left\langle \hat{\nu}^2 \right\rangle\!\!\right\rangle_{\eta,z,I}$$

with

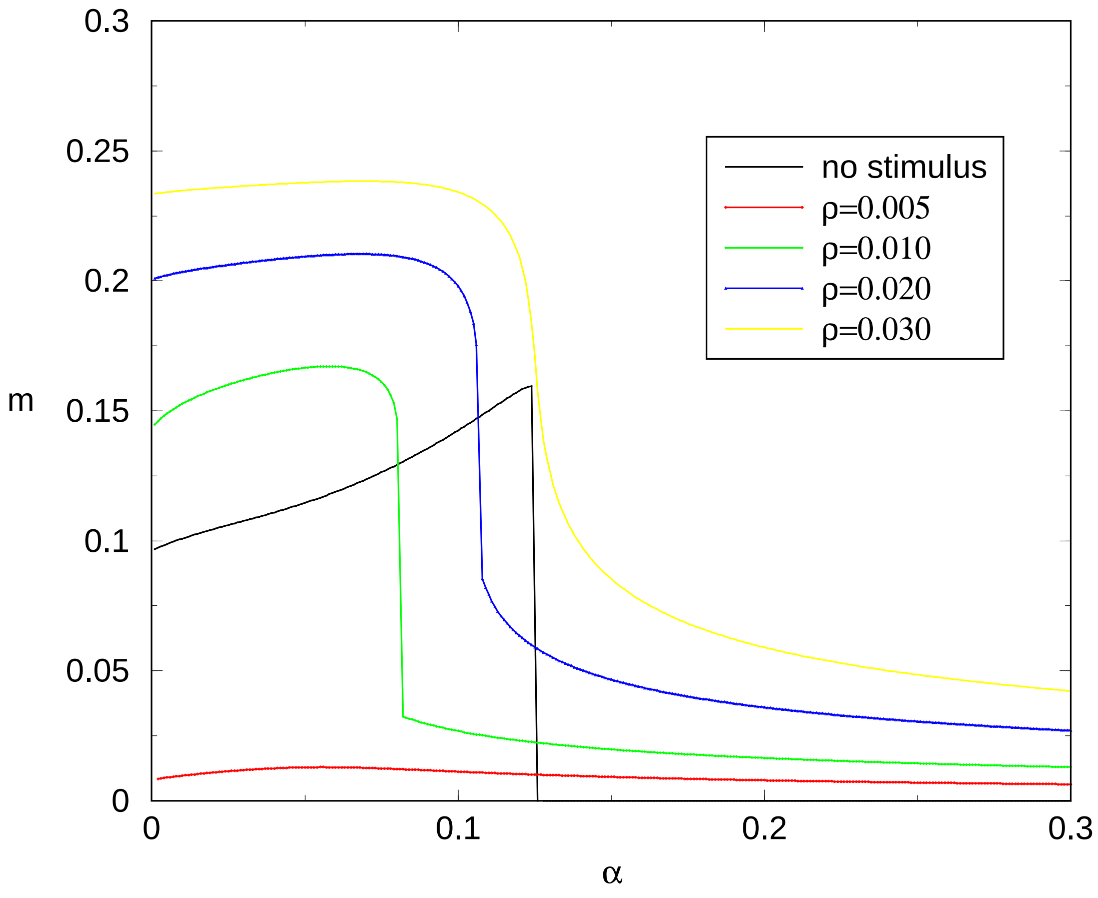
$$\hat{\nu} = \hat{\nu}(\eta,z,I) = g\Big(m(\eta-a) + \sqrt{\alpha r}\; z + \frac{\alpha c}{1-c}\hat{\nu} + I - \vartheta\Big) \qquad (*)$$

- Input currents (stimuli) Gaussian: $I_i = \rho\eta_i + \sigma\xi_i$ with $\xi_i \sim \mathcal{N}(0,1)$

- Firing rate distribution, parameterised by m, c, q; from (*):

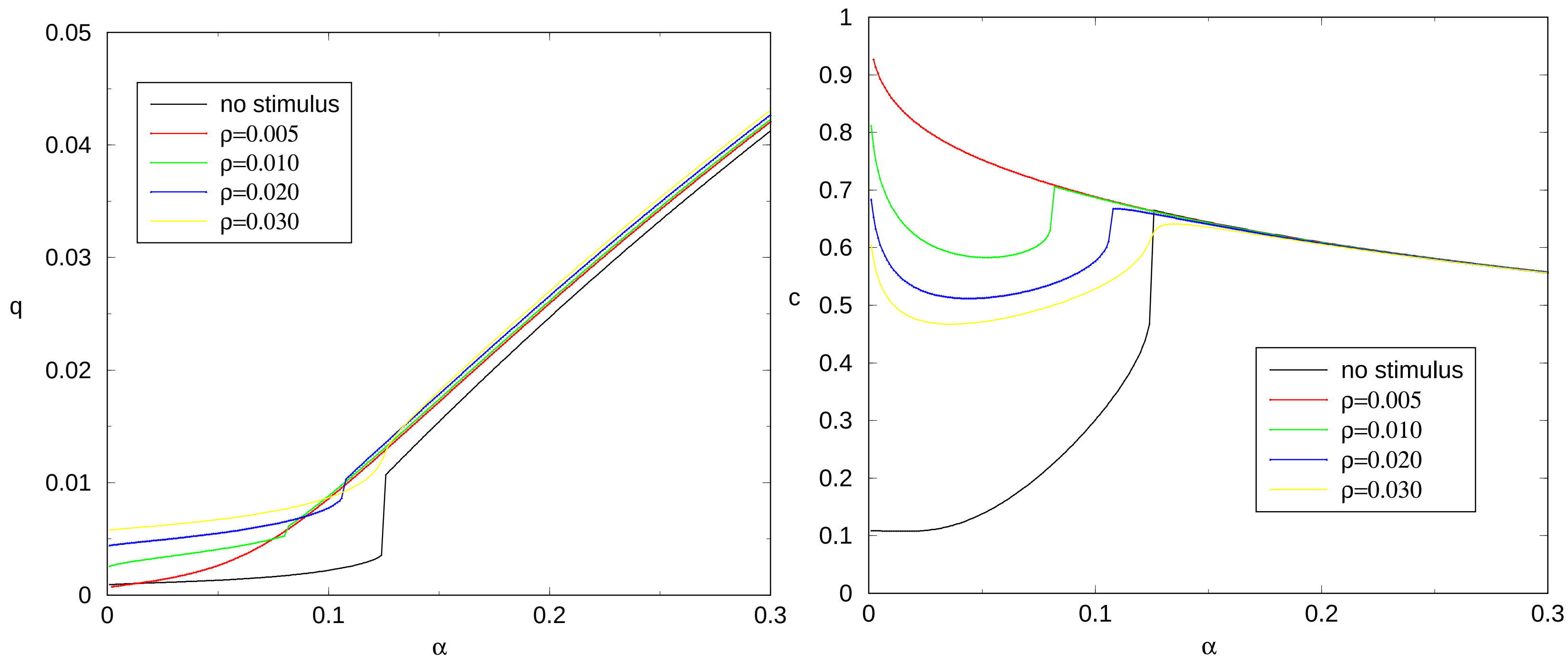
$$p(\nu,I,\eta) = \frac{1}{N}\sum_i \delta_{\eta,\eta_i}\;\delta(I-I_i)\;\Big\langle \delta(\nu-\nu_i)\Big\rangle = \Big\langle \delta(\nu-\hat{\nu})\Big\rangle\Big|_{I,\eta}$$

Order Parameters



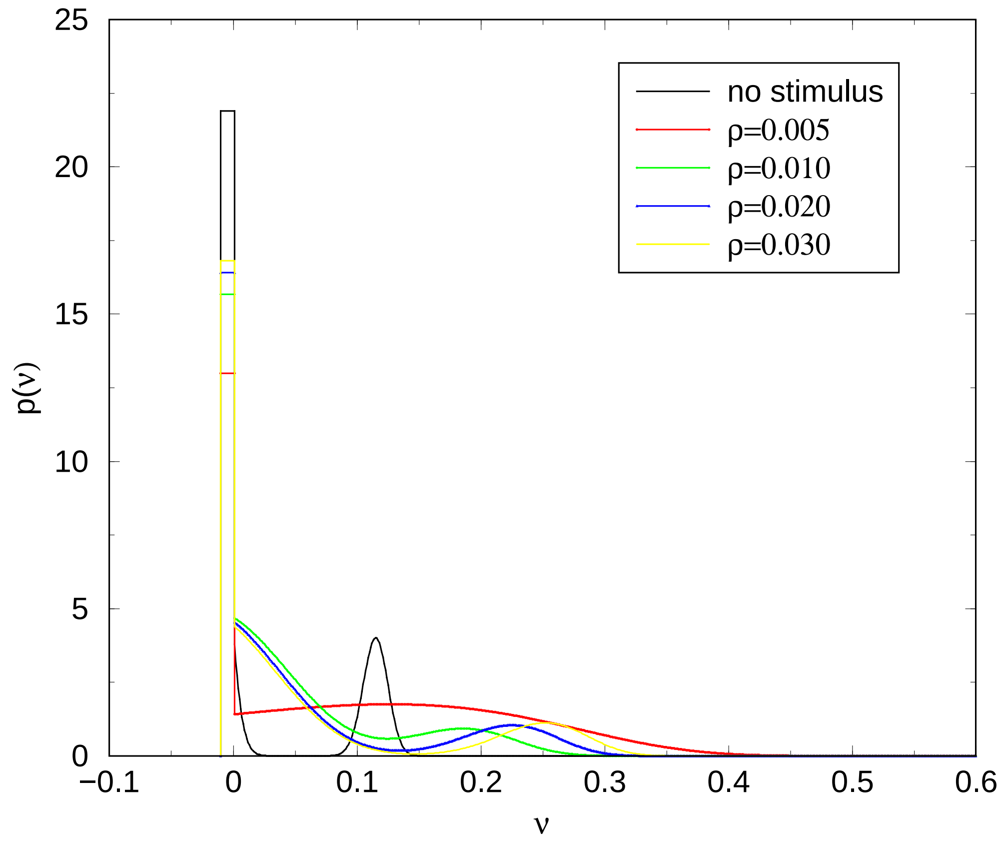
Overlap as function of loading level, without stimulus, and for $\rho = 0.005$, $\rho = 0.01$, $\rho = 0.02$, $\rho = 0.03$, and $\sigma = 0.03$.

Order Parameters



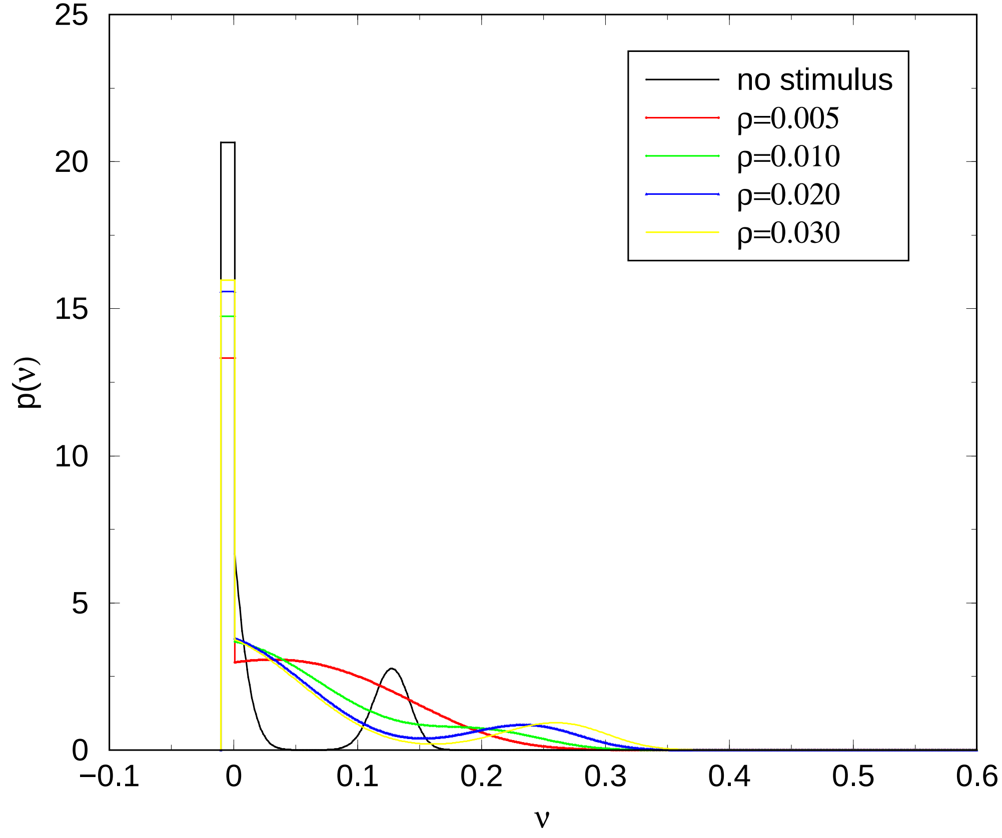
Spin-glass order parameter q and susceptibility c as functions of loading level, without stimulus, and for $\rho = 0.005$, $\rho = 0.01$, $\rho = 0.02$, $\rho = 0.03$, and $\sigma = 0.03$.

Firing Rate Distributions



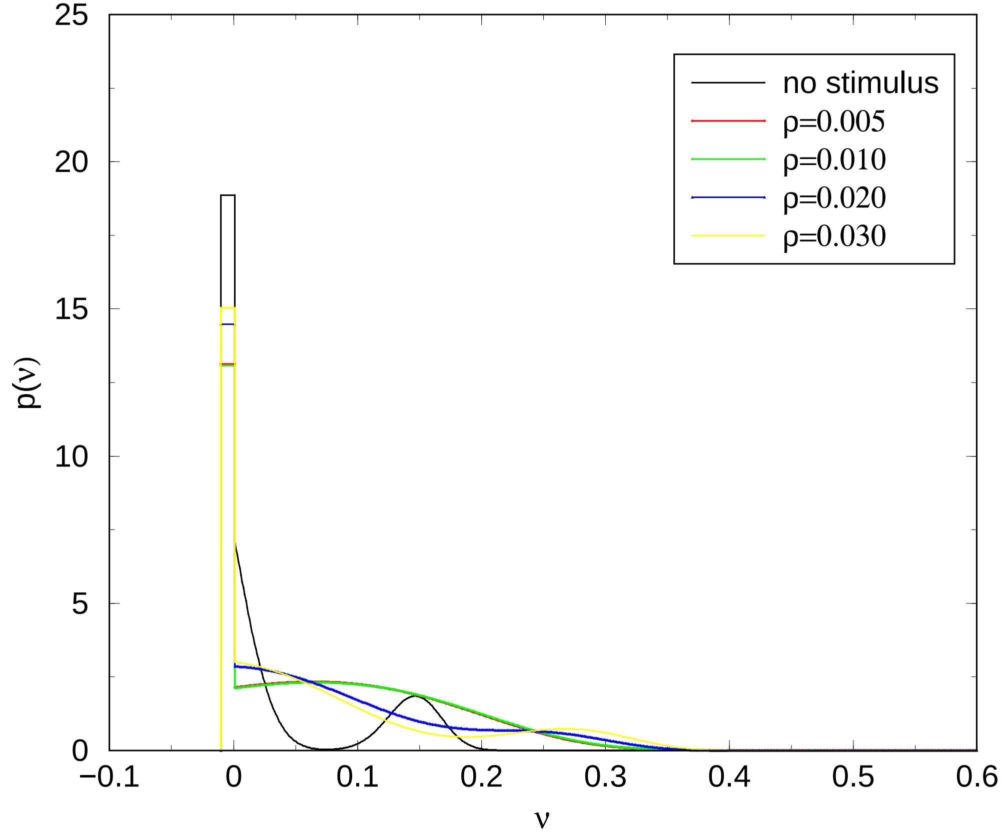
Firing rate distribution at $\alpha = 0.005$, without stimulus, and for $\rho = 0.005$, $\rho = 0.01$, $\rho = 0.02$, $\rho = 0.03$, and $\sigma = 0.03$.

Firing Rate Distributions



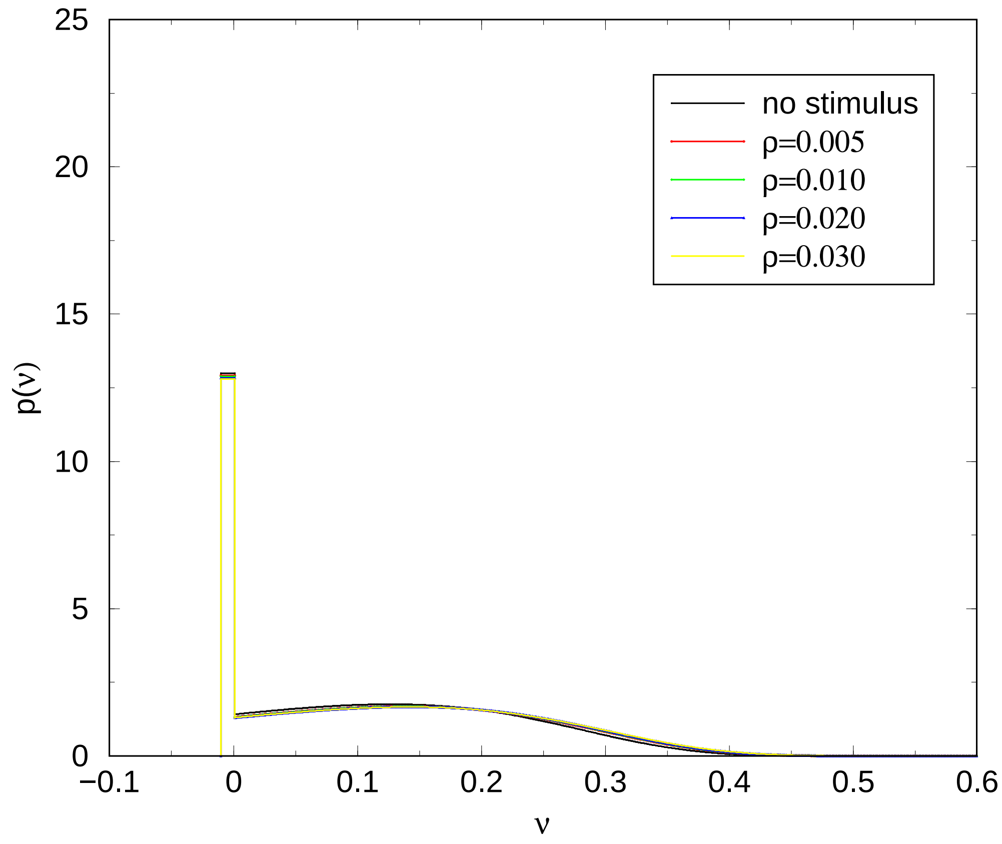
Firing rate distribution at $\alpha = 0.075$, without stimulus, and for $\rho = 0.005$, $\rho = 0.01$, $\rho = 0.02$, $\rho = 0.03$, and $\sigma = 0.03$.

Firing Rate Distributions



Firing rate distribution at $\alpha = 0.10$, without stimulus, and for $\rho = 0.005$, $\rho = 0.01$, $\rho = 0.02$, $\rho = 0.03$, and $\sigma = 0.03$.

Firing Rate Distributions



Firing rate distribution at $\alpha = 0.15$, without stimulus, and for $\rho = 0.005$, $\rho = 0.01$, $\rho = 0.02$, $\rho = 0.03$, and $\sigma = 0.03$.

Performance Measures

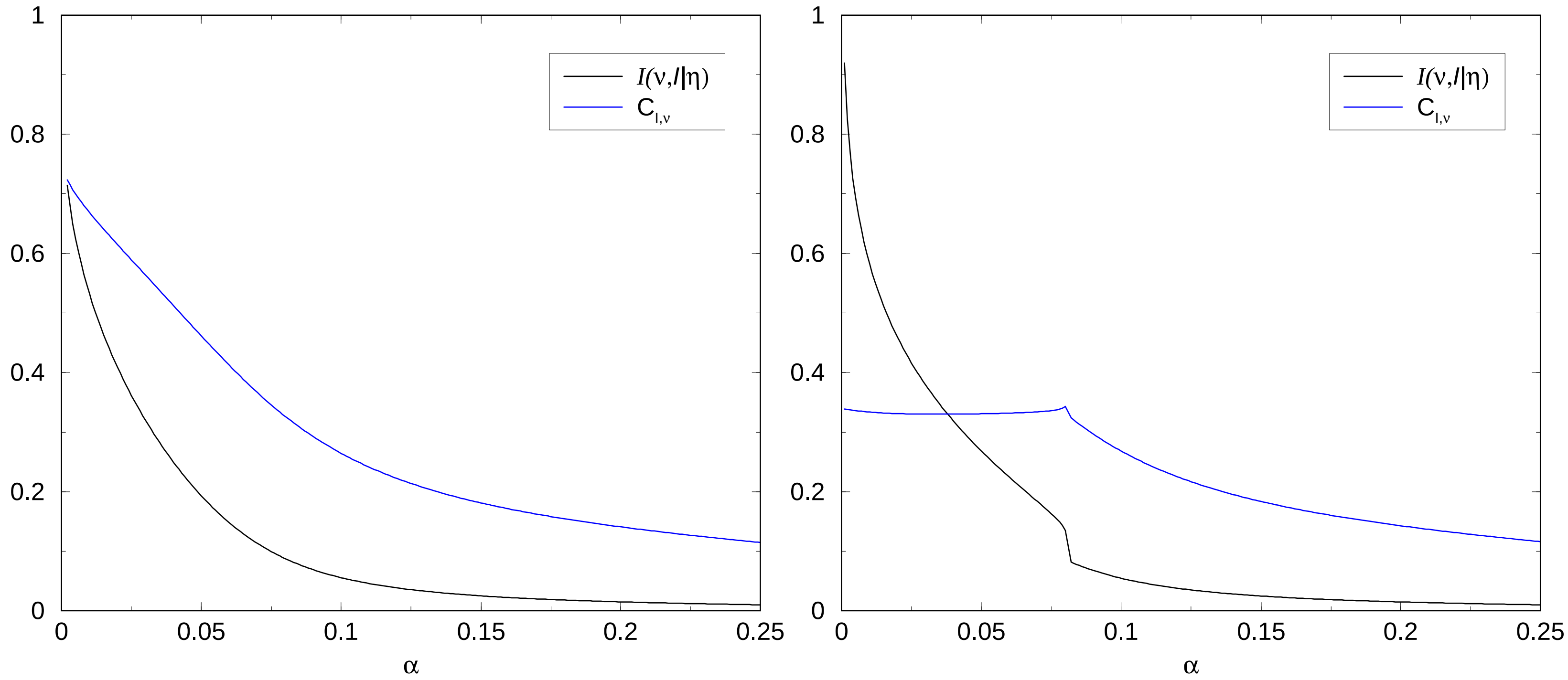
- Mutual conditional information of firing rate distribution and input currents

$$\tilde{I}(\nu, I|\eta) = \sum_{\eta} \int \mathrm{d}\nu \mathrm{d}I \; p(\nu, I, \eta) \; \log_2 \frac{p(\nu, I|\eta)}{p(\nu|\eta)p(I|\eta)}$$
$$= \sum_{\eta} p(\eta) \int \mathrm{d}\nu \mathrm{d}I \; p(\nu|I\eta)p(I|\eta) \; \log_2 \frac{p(\nu|I,\eta)}{p(\nu|\eta)}$$

- Normalised correlation between firing rates and currents

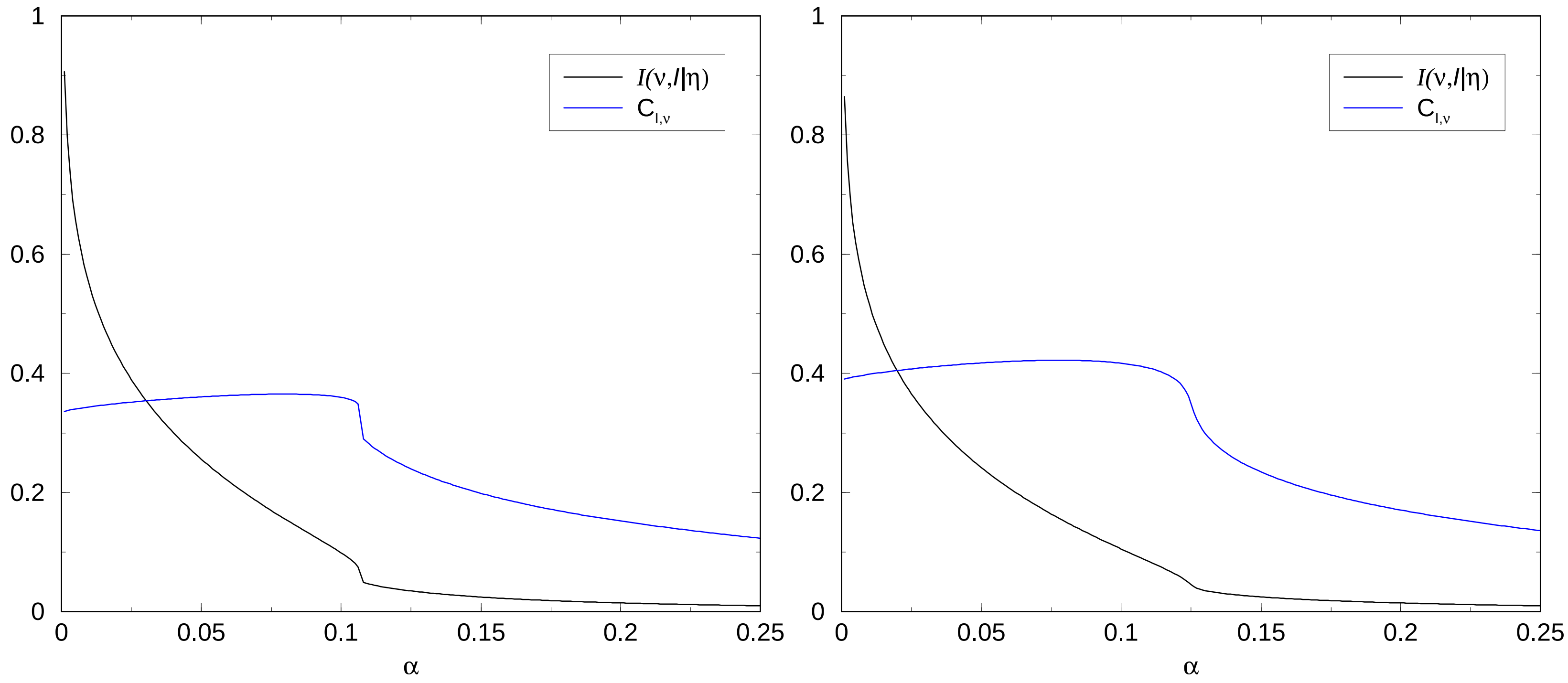
$$C_{I,\nu} = \frac{1}{\sqrt{\sigma_\nu^2 \sigma_I^2}} \Big[\langle \nu \; I \rangle - \langle \nu \rangle \langle I \rangle \Big]$$

Mutual Information and Correlation



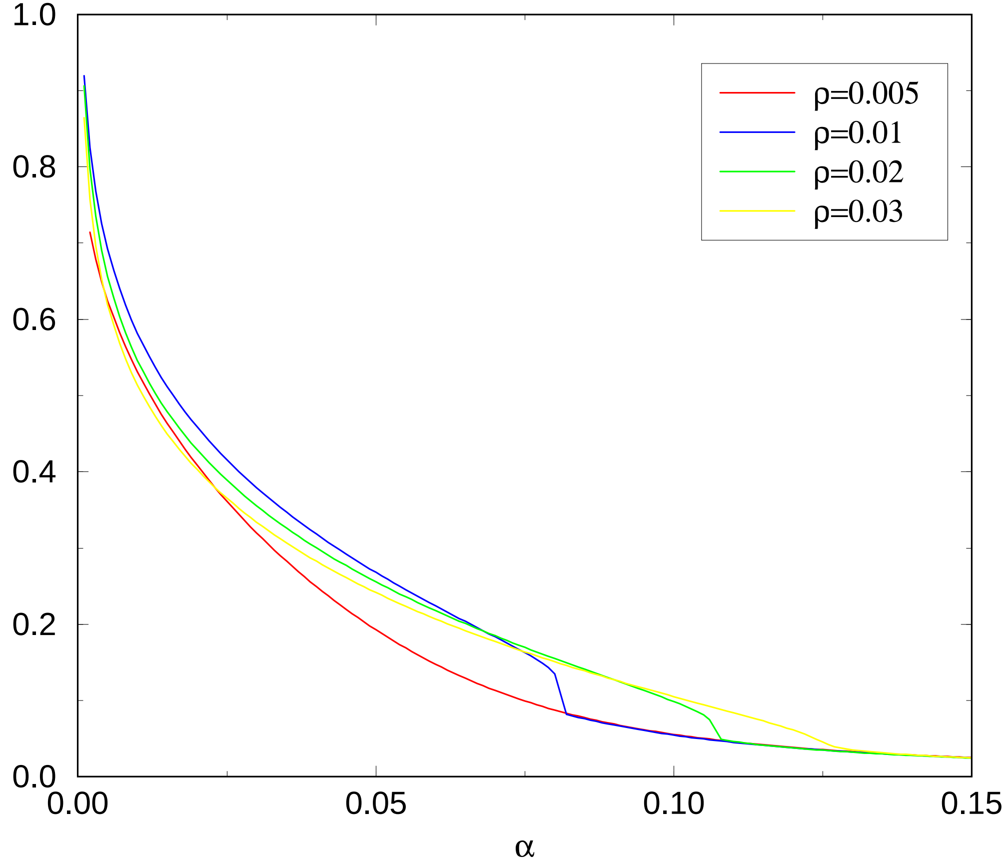
Mutual conditional information and current-rate correlation at $\sigma = 0.03$ and $\rho = 0.005$ (left) and $\rho = 0.01$ (right) as functions of α.

Mutual Information and Correlation



Mutual conditional information and current-rate correlation at $\sigma = 0.03$ and $\rho = 0.02$ (left) and $\rho = 0.03$ (right) as functions of α.

Mutual Information



Mutual conditional information as a function of α for various degrees of correlation.

Summary and Outlook

- Computed firing rate distributions in analogue neuron systems

- Studied dependence on degree of correlation between stimulus and pre-learnt patterns

- Started systematic evaluation of information theoretic performance measures

- Results – at least in principle – experimentally accessible

- Useful as theoretical framework for interpretation of recordings using trained-untrained scenarios?

- Within reach and/or to be done
  - Graded patterns in the learning rule
  - Pattern distribution that maximises mutual information?
  - Asymmetric couplings
  - Non-stationary effects, dynamics $\Rightarrow$ J. Hatchett (KCL)
  - More realistic neuron models