

Spectral dimension reduction on parametric models for spike train statistics

Cesar Ravello¹, Ruben Herzog^{a1,2}, Bruno Cessac³, Maria-Jose Escobar⁴, Adrian Palacios¹





1 Centro Interdisciplinario de Neurociencia de Valparaiso, Universidad de Chile. 3 INRIA Sophia Antipolis. 4 Universidad Tecnica Federico Santa Maria cesar.ravello@postgrado.uv.cl; rubenherzog@ug.uchile.cl

INTRODUCTION

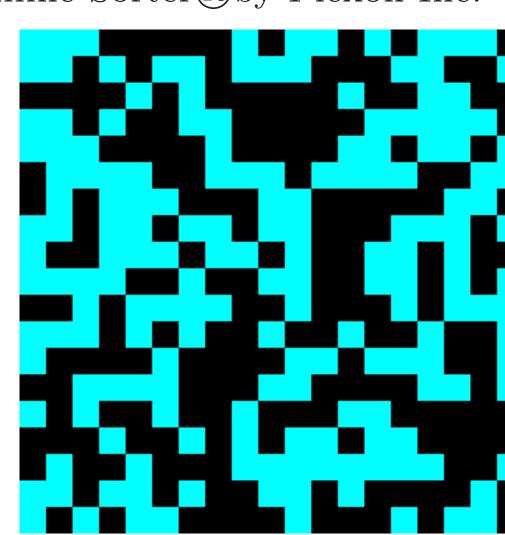
It has been shown that the neurons of visual system present correlated activity in response to different stimuli. The role of these correlations is an unresolved subject. These correlations vary according to the stimulus, specially with natural images. To uncover the role of these correlation and characterize the population code, it is necessary to measure the simultaneous activity of large neural populations. This has been achieved thanks to the advent of Multi-Electrode Array technology, opening up a way to better characterize how the brain encodes information in the concerted activity of neurons. In parallel, powerful statistical tools have been developed to accurately characterize spatio-temporal correlations between neurons. Methods based on Maximum Entropy Principle, where statistical entropy is maximized under a set of constraints corresponding to specific assumptions on the relevant statistical quantities, have been proved successfully, specially when they consider spatiotemporal correlations. [ref] They are although limited by (i) the assumption of stationarity, (ii) the many possible choice of constraints, and (iii) the huge number of free parameters.

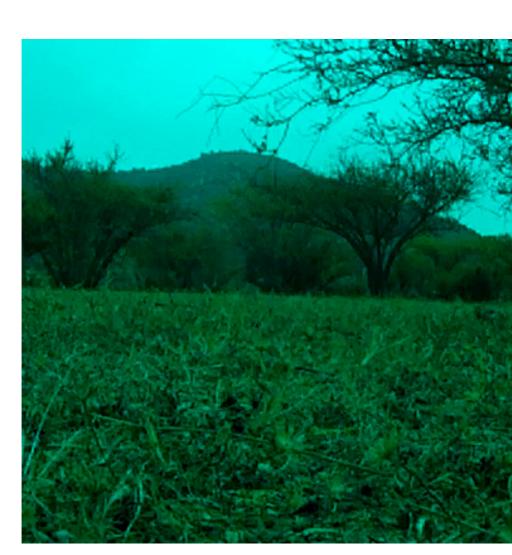
In this context, focusing on (ii), (iii), we propose a method of dimensionality reduction allowing to select a model fitting data with a minimal number of parameters. This method is based on the spectral analysis of a symmetric, positive matrix, summing up all relevant spatial-temporal correlations, closely related to the Fisher metric in statistical analysis and information geometry, but extended here to the spatio-temporal domain. Based on synthetic and real data - RGC responses to different stimuli in a diurnal rodent- we show that the spectrum of this matrix has a cut-off beyond which the corresponding dimensions have a negligible effect on the statistical estimation. This dimensionality reduction reduces the risk of over-fitting. The method is used to characterize differences in response to different classes of visual stimuli (white noise, natural images).

RECORDINGS AND STIMULI

Extracellular recording of the electrical activity of retinal patches from healthy, young individuals of the rodent species Octodon degus were performed ex vivo using a Multielectrode Array (USB-MEA256 from Multichannel Systems, MCS GmbH). Distance between electrodes was $100\mu \text{m}$ in a square array. The retina was kept at $33\,^{\circ}\text{C}$ in AMES medium and a gas mix of $95\% O_2 5\% CO_2$.

Stimuli consisted of white noise (WN, binary checkerboard pattern) and a Natural Movie (WN) recorded in the animal's natural environment. The images were projected using a conventional DLP LED projector at 60Hz, with custom optics to reduce and focus the image through the microscope onto the array. Spontaneous activity (SP) was recorded in total darkness. The voltage traces (sampled at 20KHz) were analyzed to extract the timing of spikes using the software Offline Sorter®by Plexon Inc.





White Noise

Natural Movie

5 different experiments with 3 different conditions (SP, WN, NM) were used and incremental (inc = 4, starting from a network size, N, equal to 16) random subsamples of the whole population were used for the *Dimensionality Reduction* analysis. This procedure was repeated 18 times for each experiment and condition. For comparison, synthetic data was generated using Bernoulli and Pairwise models.

ACKNOWLEDGEMENTS

FONDECYT 1110292, 1140403, KEOPS/ANR-47 CONICYT, Millennium Institute ICM-P09-022-F. NICOP # N62909-14-1-N121 ONRG. CONICYT Fellowship to RH and CR. Programa Doctorado en Neurociencia Universidad de Valparaiso. Instituto de Sistemas Complejos de Valparaiso.

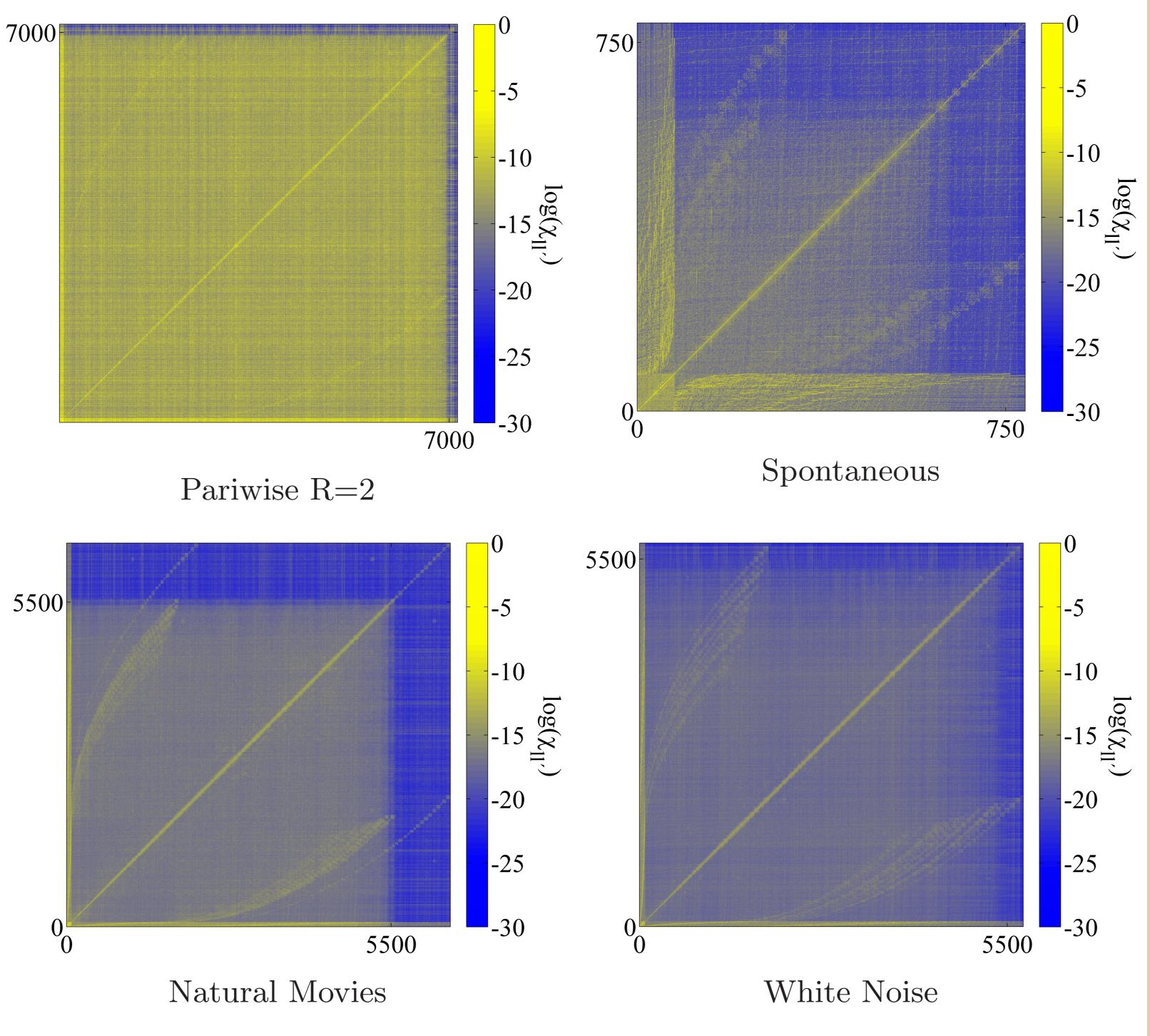
Powered@NLHPC: This research was partially supported by the supercomputing infrastructure of the NLHPC (ECM-02), by Dr. Patricio Orio CNSLab supercomputer adn the ENAS software for the computations https://enas.inria.fr/.

THE SUSCEPTIBILITY MATRIX

The statistical model we are considering (Gibbs Distribution) is characterized by two sets:(i) spatiotemporal events ($neuron_i$ spiking at time t_i and $neuron_i$ spiking at time t_i), denoted by m_l , l = 1...L and (ii) a set of parameters h_l , called parameter conjugated to m_l . This sets defines a probability distribution μ depending parametrically on h_l . Each h_l is tuned so the probability of event m_l predicted by μ is equal to the empirical probability observed on the data. Moreover, μ maximizes the statistical entropy.

A small change on h_l modifies the average value of m_l predicted by μ . The Susceptibility Matrix is the derivative $\delta^{\mu(m_l)}/\delta_{h_l} = \chi_{ll'}$. χ is also called the Fisher metric in information geometry and is the second derivative of the Kullback-Liebler divergence. The structure of χ tells us about the effect of the modification of one parameter of the distribution on the global

We observe (i) χ is essentially dominated by the parameters corresponding to firing rates, however, behind this principal structure appears a more complex fractal structure connected to high order correlations. (ii) this structure strongly depends on the type of stimulus and presumably on the retina structure.



 χ is a symmetric positive matrix, although its numerical estimation can contain vanishing eigenvalues related to unobserved events. As a consequence, χ has real positive eienvalues λ_i . The size of λ_i is directly related to the quality of the empirical finite size statistical estimation in the corresponding eigendirection. The larger the λ_i , the more reliable the statistical estimation. As a corollary, eigenvalues λ_i close to zero have dramatic impact on parameters estimations of the model. Therefore, we expect to have cut-off's in the spectrum beyond which the corresponding degrees of freedom are irrelevant for the statistics.

Additionally, we observe scaling laws (i.e. power-laws) in the spectrum that depends on N, in particular, the number of observed events at a particular network size, N, in a finite size raster dramatically depends on the type of stimulus.

Especially, we observed two cut-off's, the first one at the transition from spike rates eigenvalues to high order correlations eigenvalues, whereas the second cut-off determine the limit beyond which parameters are irrelevant.

Susceptibility Matrix Spectrum and Dimensionality Reduction

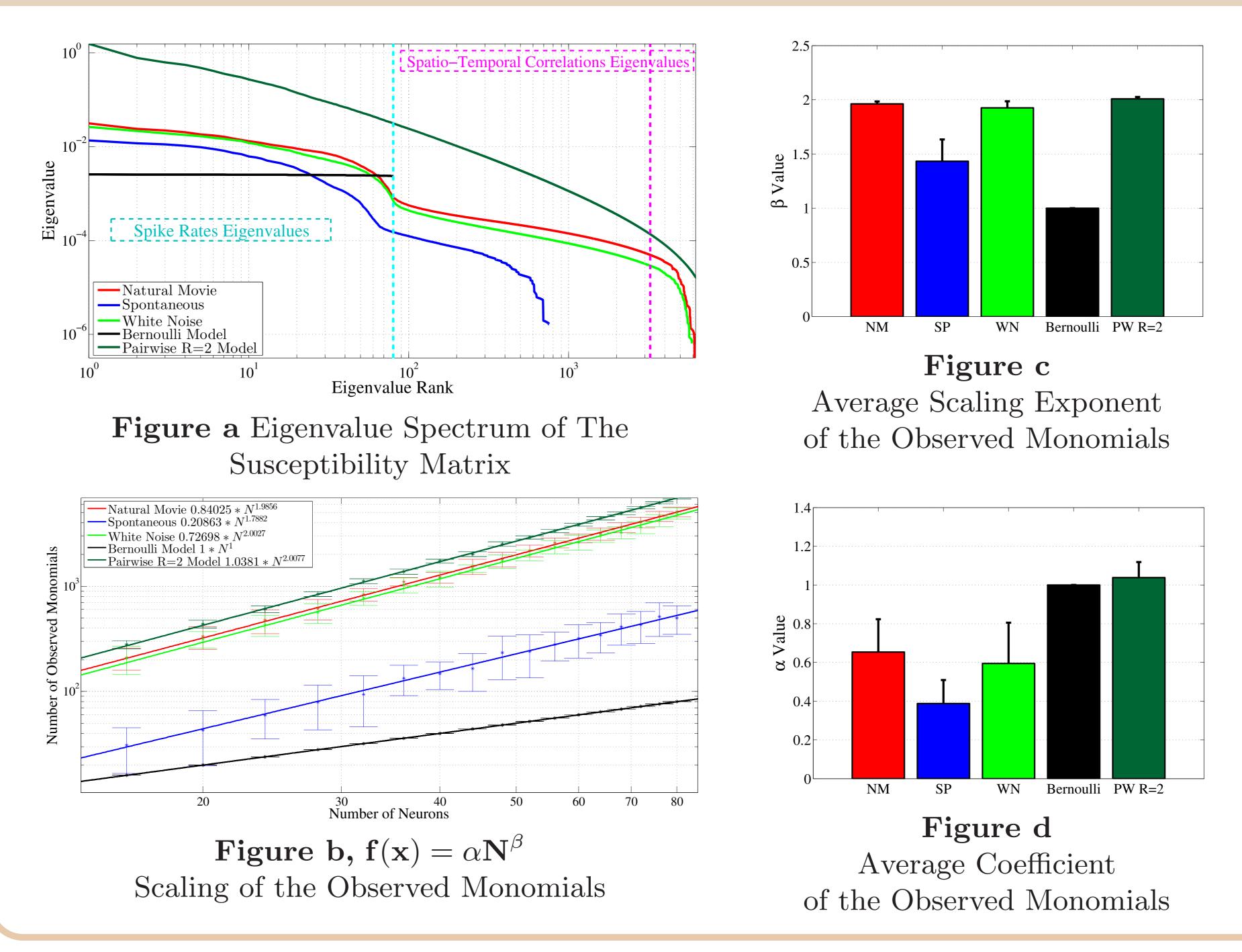
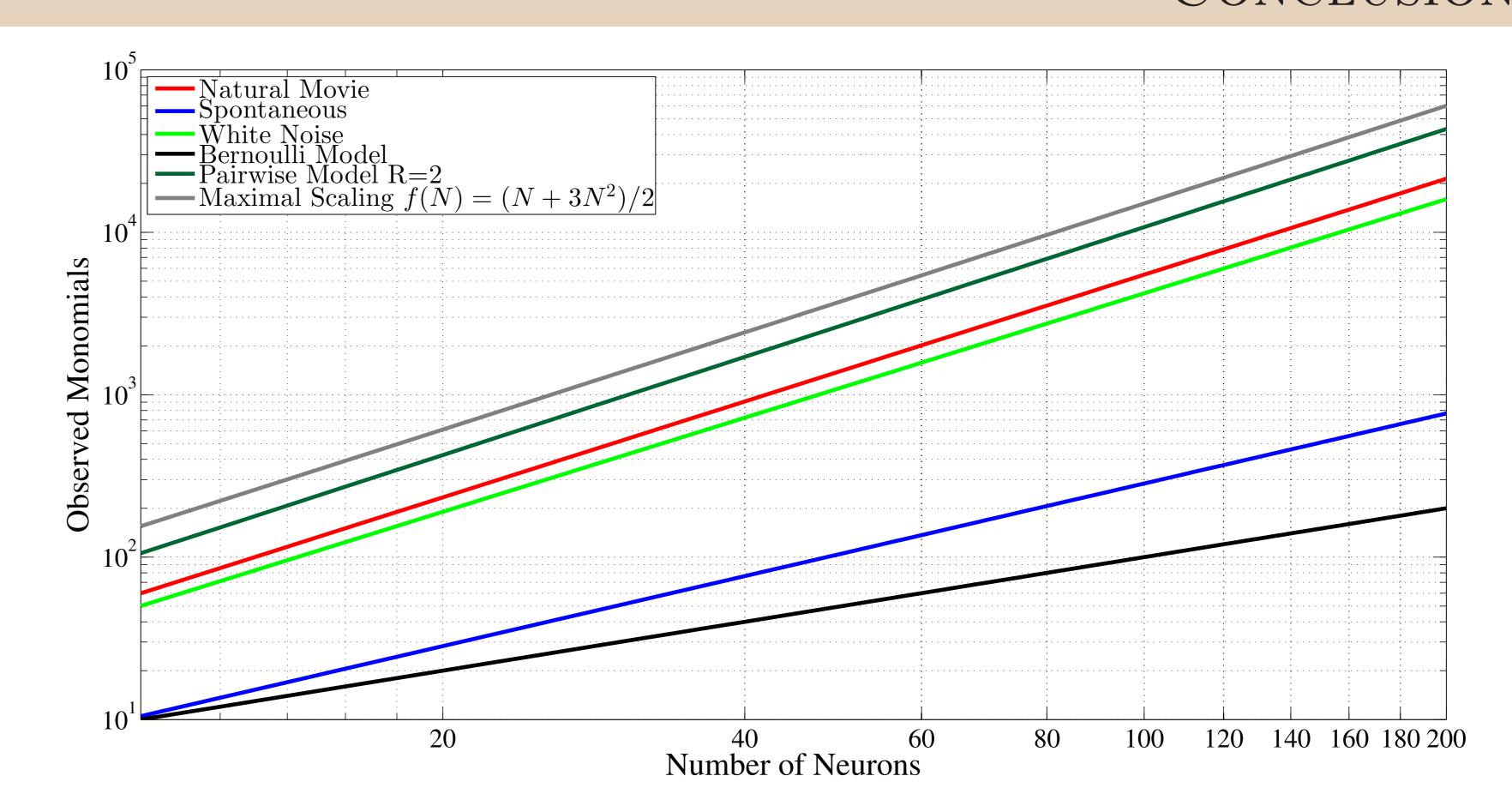
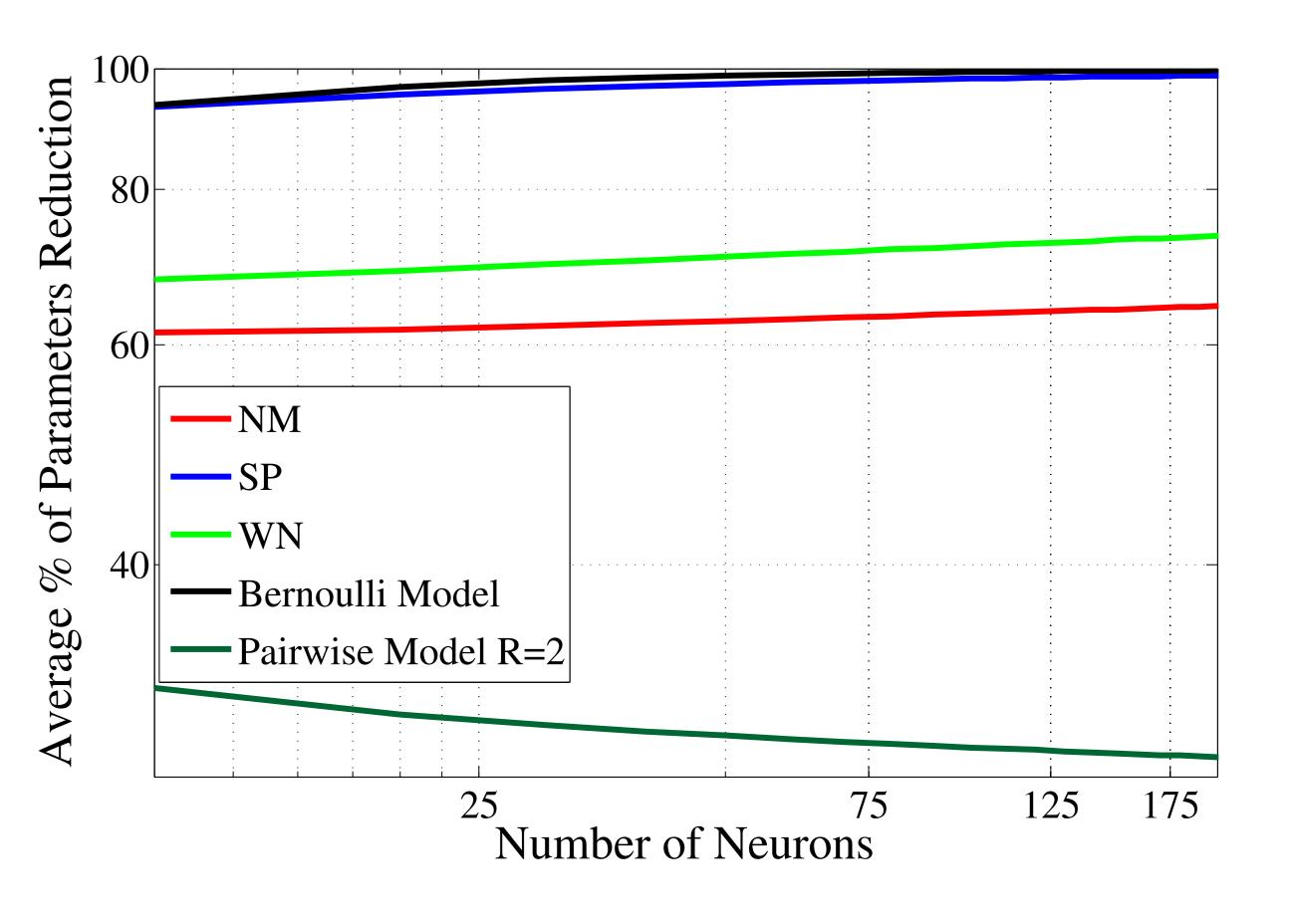


Figure a shows the Spectrum for each raster. For the experimental conditions, there is a clear cut-off near to the end of the rates-related eigenvalues, showing that rates has the highest impact on the raster statistics. PW shows no cut-off, i.e. no significant dimensionality reduction, which is in agreement with the corresponding susceptibility matrix structure. Bernoulli shows only rate-related eigenvalues, as expected. On the opposite, retina data experiments show cut-off. Figure b shows in log-log plot the scaling of the observed monomials for the same rasters than in a.

Figure c and **d** show the average fitted values for α (coefficient) and β (scaling exponent) for the $f(x) = \alpha N^{\beta}$ power-law. It is clear the linear scaling of the Bernoulli model, with α and $\beta = 1$. For NM and WN there is almost the same quadratic scaling ($\beta \sim 2$), with differences within \pm 1SD. PW also shows a quadratic scaling, but with a higher offset (i.e. α) than WN and NM. SP shows a non-linear scaling, but is not quadratic at all. Error bar denotes 1 SD both in the log-log plot and in the bar plot.

Conclusions





- Analyzing the Susceptibility matrix spectrum and its scaling gives a quantitative measure of dimensionality reduction on the estimated Gibbs Distribution.
- Even though the observed monomials in NM, WN and PW are considerably high, they don't reach the maximum expected number of monomials (grey line in left panel).
- This deserves further investigations, especially concerning the interpretation of the cut-off's and the presence of scaling laws.