34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Spline- and Wavelet-based Models of Neural Activity in Response to Natural Visual Stimulation*

Felipe Gerhard[1] and Luca Szegletes[2]

Abstract— **We present a comparative study of the performance of different basis functions for the nonparametric modeling of neural activity in response to natural stimuli. Based on naturalistic video sequences, a generative model of neural activity was created using a stochastic linear-nonlinear-spiking cascade. The temporal dynamics of the spiking response is well captured with cubic splines with equidistant knot spacings. Whereas a sym4-wavelet decomposition performs competitively or only slightly worse than the spline basis, Haar wavelets (or histogram-based models) seem unsuitable for faithfully describing the temporal dynamics of the sensory neurons. This tendency was confirmed with an application to a real data set of spike trains recorded from visual cortex of the awake monkey.**

I. Introduction

A majority of neurons in the visual areas of the brain respond to specific features in the visual input by firing a temporal sequence of action potentials (or *spikes*). A popular statistical description consists of a cascade of a linear filtering stage, followed by a nonlinear transfer function to obtain an instantaneous firing rate from which spikes are stochastically generated [1]. While this model can explain the neural responses to simple stimuli such as light dots or moving bars, it captures only a small variance of the response to natural stimulation [2]. Consequently, instead of directly estimating a tentative linear receptive field and the nonlinearity, researchers have directly estimated the firing intensity over time using semi- or non-parametric methods [3], [4].

Traditionally, the neural firing intensity is estimated from data using binning approaches (a piece-wise constant model of the firing intensity) that are known in the neuroscientific literature as PSTH estimators (peri-stimulus time histograms) [5]. More recently, nonparametric models using regularly spaced cubic splines were employed to model stimulus dependencies of the firing rate [6]. Modern signal processing techniques utilize wavelet decompositions which have been shown to be efficient representations for many naturally occurring signals [7]. However, to our knowledge, despite an early proposal [8], wavelet analysis has not found its way into modeling neural intensity functions. Because time series of visual signals under natural stimulation are highly non-stationary and consist of frequent transients [1], we hypothesized that wavelets might allow efficient signal representations.

It is important to find a good statistical model of the temporal spiking activity because it disentangles two confounding factors in measured correlations of the ensemble of simultaneously recorded neurons: stimulus-induced correlations (neurons share similar receptive fields or respond to correlated features in the input) and network-induced correlations (arising from direct synaptic or more general effective inter-neuron couplings). To study the graph of the effective coupling structure between neurons, one therefore has to carefully account for the modulation of neural activity induced by the stimulus.

The statistics of the signal dictate which basis function representation is most suitable. In the first part of our analysis presented here, we derive a generative model of the expected intensity modulations. To this end, we extract temporal dynamics of a natural movie and apply the cascade neuron model to obtain realistic time series and their statistics. On these time series, we can then evaluate the relative fidelities of spline- and wavelet-based approaches. As a second step, we apply the different models to a data set of recorded neurons in the visual cortex of the awake monkey that is passively and repeatedly viewing excerpts from natural movies. Using a cross-validation procedure, we can directly evaluate the predictive power to explain the observed spiking patterns.

II. Methods

A. Toy data

1) Signal characterization and generative model: We used a movie recorded by a CCD camera mounted on the head of a freely moving cat as the basis for the natural stimulus. The video was recorded by the group of Peter Koenig (University of Osnabrueck, Germany[1]). Full details can be found elsewhere [9]. Briefly, the video is of length 60 seconds (50 Hz) at a resampled resolution of 128 by 128 pixels. Pixel intensities are encoded in 8-bit gray scale (see Figs. 2A and B for example image frames).

We filtered a subregion of the natural movie at a random location with a two-dimensional Gabor function of size 32-by-32 pixels and random orientation and phase to simulate the filtering characteristics of simple cells of the visual cortex (see Fig. 2C for an example).

*Felipe Gerhard is supported by the Swiss National Science Foundation (SNSF) under the grant number 200020_132871. Luca Szegletes was supported by an Erasmus grant.

[1]F. Gerhard is with the Brain Mind Institute, Ecole Polytechnique Federale de Lausanne, 1015 Lausanne EPFL, Switzerland felipe.gerhard at epfl.ch

[2]L. Szegletes is with the Brain Mind Institute, Ecole Polytechnique Federale de Lausanne, 1015 Lausanne EPFL, Switzerland lucaszegletes at gmail.com

[1]Source: http://crcns.org/data-sets/vc/pvc-3/movie. Last access: 02/15/2012.

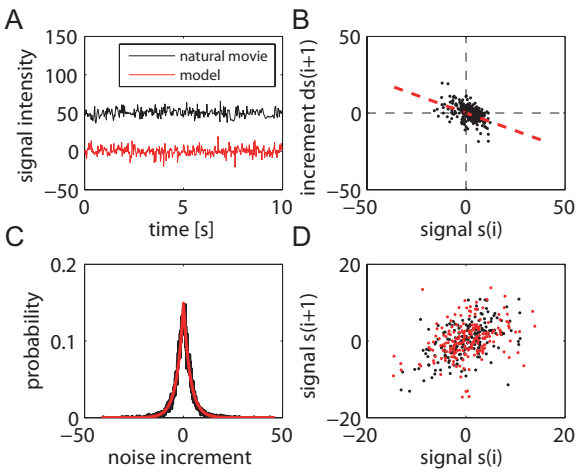

Fig. 1. Characterization of temporal dynamics of natural visual stimuli. (A) Sample trace of a linear Gabor filter applied to a video sequence with natural statistics (black, arbitrarily shifted). The synthetic model of the time series is an AR(1) process with Laplacian-distributed noise increments: $s(i+1) = s(i)(1+\gamma) + \epsilon(i+1)$ (red). (B) Scatter plot of finite differences of signal values $ds(i+1) = \gamma s(i) + \epsilon(i+1)$ versus $s(i)$. The linear slope determines $\gamma = -0.56$. (C) Histogram of noise increments (black). The distribution is symmetric around the origin and well-fitted with a Laplacian with decay constant $\mu = 3.08$ (a.u.) (red). (D) Scatter plot of $s(i+1)$ versus $s(i)$ for both the extracted time series (black) and the model (red). For clarity, only 200 sample pairs of each type are displayed. Model and real time series share similar auto-correlation structure.

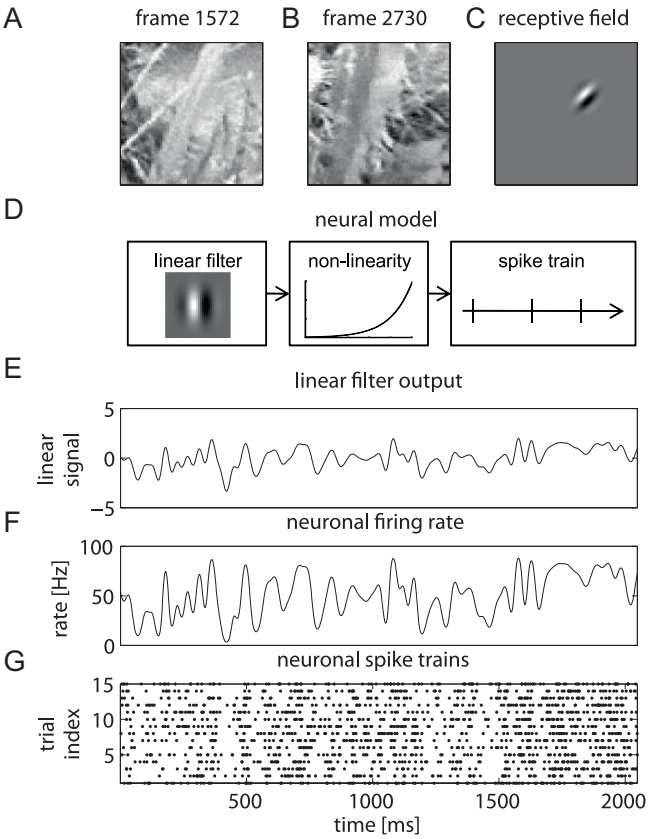

Fig. 2. Generative model of neural responses to natural stimuli. (A-B) Two example frames from the natural stimulus movie. (C) Example receptive field. (D) Scheme of the cascade neuron model. The output of a linear filter stage is transformed into a nonnegative firing intensity from which spikes (point process events) can be stochastically drawn. (E) Signal obtained by convolving the receptive field with the stimulus and normalizing. (F) Firing rate of the simulated neuron is obtained by a non-linear transform of the linear signal. (G) Example spike trains generated as stochastic realizations of the firing intensity.

The statistical features of the resulting one-dimensional time series $s(i)$ suggest to model it as an AR(1) process with non-Gaussian noise increments. Specifically, $s(i+1) = s(i) + ds(i+1)$ with $ds(i+1) = \gamma s(i) + \epsilon(i)$ and $\epsilon(i) \propto Laplace(\mu)$ is a good approximation to the observed time series (see Fig. 1 for details). γ was obtained by the slope of the linear fit of $ds(i+1)$ against $s(i)$ (see Fig. 1B) and μ was obtained by a fit to the distribution of noise increments $\epsilon(i)$ (Fig. 1C). Synthetically generated time series are statistically similar to time series directly extracted from the natural movie (Figs. 1A and D) so that we used the synthetic time series for the remaining analysis. The time series were further resampled by a factor of 20 (using a polyphase filter implementation) to obtain a sampling rate of 1000 Hz in order to match conditions of typical electrophysiological recordings (see Fig. 2E for an example). 100 such synthetically generated time series of length 2.048 seconds (corresponding to 2048 data points) were used for fitting.

For modeling neurons in early visual cortex, we followed the classical cascade neuron model (Linear - Nonlinear - Spiking model) [10]. The linear part is given by the previously described Gabor filtering, respective the synthetic model. This filter output is transformed using a sigmoidal nonlinearity to yield an instantaneous probability of firing an action potential (or spike) within the sampling period $\Delta = 1$ ms: $p(i) = \frac{1}{1+\exp(-s(i))}$, where $s(i)$ is the (resampled) linear signal at time i. Spike trains (represented by the binary sequence $Y(i)$) can be generated from the instantaneous firing probabilities as a realization of a Bernoulli process with given spiking probabilities $p(i)$. We generated $N = 10$ independent realizations ('trials') for each time series to mimick experimental conditions in which the same stimulus is repeatedly displayed during electrophysiological recordings.

2) Model specification and fitting: In practice, neither the input $s(i)$ that is driving the neural activity nor its nonlinear transformation $p(i)$ can be directly observed. Instead, we have to infer an estimate $\hat{s}(i)$ of the time course of $s(i)$ using the spike count observations $Y(i)$. We can represent $\hat{s}(i)$ through a set of basis functions: $\hat{s}(i) = \sum_j W_{ij}\beta_j$ where W_{ij} is the jth basis function evaluated at time i and β_j are the estimated coefficients. We obtain the coefficients as the maximum-likelihood solution of the nonlinear logistic regression problem: $Y(i) \sim Bernoulli(p(i))$ with $p(i) = \frac{1}{1+exp(-\sum_j W_{ij}\beta_j)}$.

We compare the reconstruction performances of three different basis function representations: a) a multi-level wavelet decomposition using 'sym4' wavelets, b) a multi-level wavelet decomposition using Haar wavelets and c) cubic splines with equidistant knot placing. The number of coefficients used for fitting can be varied by taking all

wavelet functions (including the scaling functions) of level k or higher (here, up to level 8). This yields representations with the number of coefficients ranging from 16 to 1024. The number of knot points of the spline representation were matched to the same number of coefficients.

Performance is evaluated using the mean squared error (MSE) between the original signal $s(i)$ and its predicted value $\hat{s}(i)$.

B. Real data set

1) Data acquisition: Full experimental details are published elsewhere [6]. All experimental procedures were approved by local authorities (Regierungspraesidium Hessen, Darmstadt, Germany) and were in accord with the guidelines of the European Community for the care and use of laboratory animals (European Union directive 86/609/EEC). Briefly, electrophysiological recordings were performed in the visual cortex of the awake monkey. Single-unit activity was retrieved from the signals of several extracellularly-placed electrodes and a subsequent spike-sorting procedure. The monkey was fixating and passively viewing short excerpts from natural movies. Data was recorded in 20 sessions, with an average of 24 reconstructed single-unit activities. We selected the first 2.048 seconds of each trial and binned the recorded spike train for each neuron in time bins of width $\Delta = 1$ ms. In each recording session, a specific movie sequence was repeated between 50 and 100 times. For the analysis, we pooled spike trains over all trials into a single time series of spike counts.

2) Nonlinear spiking model and evaluation: We aim to estimate the neural firing intensity as a function of time since trial onset with a nonlinear parametric model. Specifically, we assume that the observations $Y(i)$ (spike counts per time bin) are samples from a Poisson distribution with mean $\mu(i)$ that is a function of model parameters, i. e. the basis function coefficients: $\mu(i) = \exp(\sum_j W_{ij}\beta_j)$ and $Y(i) \sim Poisson(\mu(i))$. We used different covariate matrices W for the three different basis functions (see Sect. II-A). We restricted the number of possible parameters to be between 16 and 512. Parameter estimates were obtained using standard maximum-likelihood estimation. We evaluate the performance of the different models using a 10-fold cross-validation approach: The average deviance (proportional to the negative log-likelihood) on the test sets determines relative performance. Because the absolute scaling of the deviance is arbitrary, we offset it to the smallest value found. When a fit of any model complexity or basis function for a particular neuron was non-convergent (e. g. due to quasi-separability), the corresponding iteration of the cross-validation was excluded from the analysis. In total, 625 fits (12.8%) were used.

III. Results

A. Toy data

We estimated the hidden input to the neuron $s(i)$ from the observed spike trains using models with different basis functions (sym4-wavelets, Haar wavelets and cubic splines)

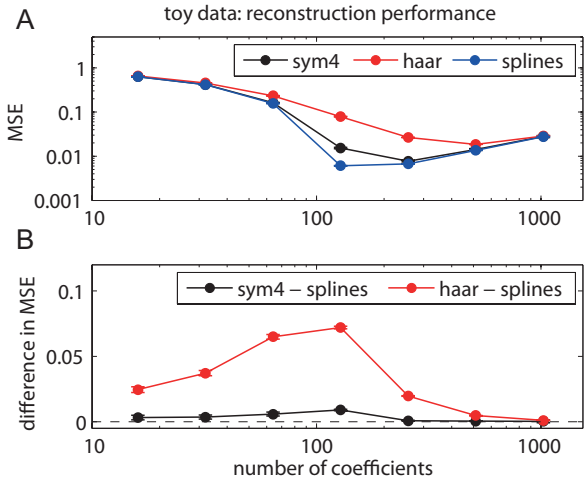

Fig. 3. Results on toy data. (A) An estimate of the signal was obtained from the spiking observations using a nonparametric model with varying number of coefficients. Mean squared errors are shown for three different basis representations: sym4-wavelets (black dots), Haar wavelets (red dots) and cubic splines (blue dots). Error bars denote s.e.m. ($N = 100$ signals) and lie within the marker shapes. Note the logarithmic scales. (B) The difference of the MSE between sym4-wavelets and splines (black dots) and Haar wavelets and splines (red dots) are shown for varying number of model parameters. Values larger 0 indicate a superior performance of the spline basis representation.

and complexity (number of estimated basis functions). First, we present results for the toy data (Fig. 3). Since for this case the ground truth is available, the mean squared error (MSE) can be used to determine the error between the estimated and original signal. We find that performance generally improves for all three sets of basis functions with increasing number of model parameters and starts to decrease for complex models (large number of coefficients) due to the increased variance in parameter estimates (Fig. 3A).

Given the same number of model parameters, the cubic spline representations performs best. It also achieves the overall lowest MSE. The sym4 basis consistently shows an almost equivalent performance; its absolute difference to the spline basis performance is always lower than 0.02 (Fig. 3B). The Haar wavelet (or PSTH) performs significantly worse in the relevant range of model complexities (Fig. 3B). Performance gets almost independent of the actual shape of the basis functions for very low and very high number of parameters.

B. Real data

We show results for fitting models of different complexities and choice of basis functions on a real data set in Fig. 4. The cross-validated deviance first decreases with increasing number of coefficients and then increases again, for all types of basis functions. As for the simulated data, cubic splines show an overall superior performance. The sym4-wavelet basis performs slightly worse for low number of model parameters (16-32), but comparable to the splines for larger values. The Haar wavelet basis is significantly worse than splines for any choice of parameters. The average difference in deviances can reach values up to 8 which corresponds to

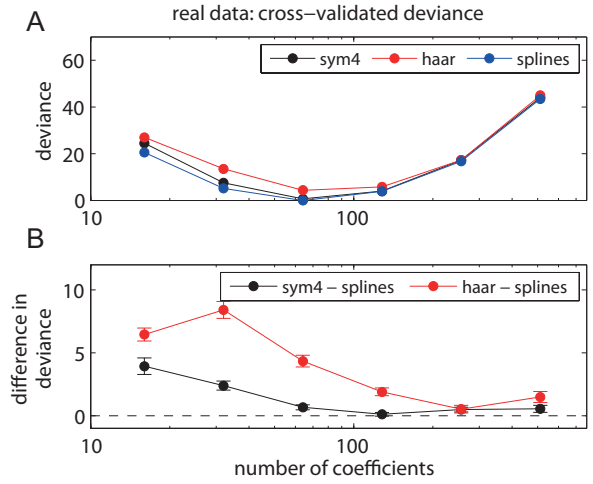

Fig. 4. Results on real data. (A) Cross-validated deviance as a measure of model performance. An estimate of underlying neural firing intensity was obtained from the spiking observations using a nonparametric Poisson regression model with varying number of coefficients. Predictive performance is measured as the cross-validated deviance, obtained from 10-fold cross-validation procedure. A constant term was subtracted to align the minimal value with 0. A lower value indicates better predictive performance. (B) Same curves as in (A), but relative to the performance of the cubic spline basis. Positive values indicate a superior performance of the splines compared to the wavelet basis functions.

a mean likelihood ratio of around $\exp(\frac{8}{2}) \approx 55$ in favor of the cubic spline model.

IV. Discussion

Our first contribution was to establish a generative model of the spiking activity of (simple) cells in the visual cortex in response to natural visual stimulation. Because the linear part of the generative model is well-described by a mean-reverting auto-regressive process with non-Gaussian jumps, we expect our analysis to be general applicable to this broader class of stochastic processes.

Based on simulations and analysis of real data from the awake monkey, we could show that cubic splines and wavelet decompositions are a suitable basis for representing neural activity in time. We showed that the Haar wavelet decomposition cannot account for the temporal dynamics of neural responses. In the context of the analysis of neurophysiological signals, the fixed Haar basis is more commonly known as a peri-stimulus time histogram (PSTH) estimate. We therefore discourage the use of simple histogram-based estimators because they are likely to lead to spurious correlations when used in the context of neural population models. It remains an open question how the detailed modeling of the stimulus using wavelets affects the estimates of their functional coupling.

Beside the choice of the basis representation, the proper model complexity, i. e. the number of model parameters to be estimated, plays an important role. For the application to the real data set, the observed optimal cross-validated performance is in the region of 25 parameters per second of recorded data. This value is close to the one used in a previous study [6] (74 parameters for a data of length 2.8 s). The results presented here therefore provide a confirmation of the their choice of basis function type and number of coefficients that was solely based on visual inspection and computational considerations.

For our analysis, we restricted ourselves to static choices of the basis functions. Specifically, for the cubic splines, we used equidistant knot spacings that are fixed independent of the actual data. We expect that adaptive knot placement strategies (such as BARS [11]) will yield even better results with the addition of increased computational complexity. Finally, for any wavelet decomposition, shrinkage methods have been shown to be very powerful in representing temporal signals with a limited number of parameters [12]. In the future, we will investigate how we can reconcile these modern signal processing techniques with classical neurophysiological models based on point processes that also include history- and cross-coupling effects.

V. Acknowledgment

The authors thank Sergio Neuenschwander, Bruss Lima and Gordon Pipa for supplying the experimental data set and Cedric Vonesch for a useful discussion.

References

[1] E. P. Simoncelli and B. A. Olshausen, "Natural image statistics and neural representation," *Annual Review of Neuroscience*, vol. 24, no. 1, pp. 1193–1216, 2001.

[2] Matteo Carandini, Jonathan B. Demb, Valerio Mante, David J. Tolhurst, Yang Dan, Bruno A. Olshausen, Jack L. Gallant, and Nicole C. Rust, "Do We Know What the Early Visual System Does?," *Journal of Neuroscience*, vol. 25, no. 46, pp. 10577–10597, Nov. 2005.

[3] Robert E. Kass, Valérie Ventura, and Emery N. Brown, "Statistical Issues in the Analysis of Neuronal Data," *Journal of Neurophysiology*, vol. 94, no. 1, pp. 8–25, July 2005.

[4] John P. Cunningham, Vikash Gilja, Stephen I. Ryu, and Krishna V. Shenoy, "Methods for estimating neural firing rates, and their application to brain-machine interfaces.," *Neural Networks*, vol. 22, no. 9, pp. 1235–1246, Nov. 2009.

[5] D. H. Perkel, G. L. Gerstein, and G. P. Moore, "Neuronal spike trains and stochastic point processes. I. The single spike train.," *Biophys J*, vol. 7, no. 4, pp. 391–418, July 1967.

[6] Felipe Gerhard, Gordon Pipa, Bruss Lima, Sergio Neuenschwander, and Wulfram Gerstner, "Extraction of Network Topology From Multi-Electrode Recordings: Is there a Small-World Effect?," *Frontiers in Computational Neuroscience*, vol. 5, no. 4, 2011.

[7] O. Rioul and M. Vetterli, "Wavelets and signal processing," *Signal Processing Magazine, IEEE*, vol. 8, no. 4, pp. 14–38, Oct. 1991.

[8] D. R. Brillinger, "Some wavelet analyses of point process data," in *Signals, Systems & Computers, 1997. Conference Record of the Thirty-First Asilomar Conference on*. Nov. 1997, vol. 2, pp. 1087–1091 vol.2, IEEE.

[9] Christoph Kayser, Wolfgang Einhäuser, and Peter König, "Temporal correlations of orientations in natural scenes," *Neurocomputing*, vol. 52-54, pp. 117–123, June 2003.

[10] L. Paninski, "Maximum likelihood estimation of cascade point-process neural encoding models.," *Network*, vol. 15, no. 4, pp. 243–262, Nov. 2004.

[11] I. Dimatteo, C. R. Genovese, and R. E. Kass, "Bayesian curve-fitting with free-knot splines," *Biometrika*, vol. 88, no. 4, pp. 1055–1071, 2001.

[12] Panagiotis Besbeas, Italia de Feis, and Theofanis Sapatinas, "A Comparative Simulation Study of Wavelet Shrinkage Estimators for Poisson Counts," *International Statistical Review*, vol. 72, no. 2, 2004.