Submitted to: The Annual Computational Neuroscience Meeting (CNS*02)
Chicago, Illinois, 21-25 July 2002.
Revision: 9 February 2002.

Biases in white noise analysis due to non-Poisson spike generation

Jonathan W. Pillow

Center for Neural Science
New York University
pillow@cns.nyu.edu

Eero P. Simoncelli

Howard Hughes Medical Center, and
Center for Neural Science
New York University
eero.simoncelli@nyu.edu

Abstract

White noise analysis methods for receptive field characterization typically ignore the dynamics of neural spike generation (i.e., they assume Poisson spike generation). We show that a linear integrate-and-fire spike generation model can induce significant bias in the estimation of the linear kernel, and leads to reverse correlation estimates that depend on stimulus amplitude, as has been observed in experimental data. We develop a modified estimator for linear characterization of such neurons, and demonstrate its effectiveness in simulation. We also apply it to physiological data, demonstrating that the estimated linear kernel is stable with respect to changes in stimulus amplitude.

White noise analysis has become a widely used technique for characterizing response properties of spiking neurons in sensory systems. A sequence of stimuli are drawn randomly from an ensemble and presented in rapid succession, and one examines the subset that elicit action potentials. This "spike-triggered" stimulus ensemble can provide information about the neuron's response characteristics. In the most widely used form of this analysis, one estimates a linear approximation to the receptive field (i.e., first-order Wiener kernel) by computing the spike-triggered average (STA); that is, the average stimulus preceding a spike [e.g., 6, 8]. Under the assumption that spikes are generated by a Poisson process with instantaneous rate determined by linear projection onto a kernel followed by a static nonlinearity, the STA provides an unbiased estimate of the underlying kernel [5].

The white noise approach is considered to have several advantages over traditional characterization approaches, including the the ability to explore a large portion of the input space, insensitivity to the strong adaptive effects of hand-optimized stimuli, and receptive field estimation that is robust to drift or fluctuation in the responsiveness of a neuron. Despite these advantages, it has also become clear that there are drawbacks to the characterizations obtained with white noise methods. One such shortcoming is the well-known phenomenon that the STA measured at different stimulus amplitudes (e.g., low vs. high contrast) often changes substantially [e.g., 13, 4, 9] . This type of change cannot be explained by a linear model followed by a static nonlinearity and Poisson spike generation

• We thank Liam Paninski and Michael Shadlen for helpful discussions.
• We thank E.J. Chichilnisky and Divya Chander for providing electrophysiological data.
• This research was supported by the Sloan-Swartz Theoretical Visual Neuroscience Center at NYU, and by the Howard Hughes Medical Institute.

(we'll refer to this as a L-N-P model), since it implies a change in the linear front end. We have previously shown that nonlinear suppressive interactions such as those found in cortical neurons can explain biases in the STA, that a spike-triggered covariance analysis can be used to characterize these suppressive interactions, and that the resulting corrected model can account for the changes of STA with contrast [10].

Here, we explore another potential source of failure in white noise characterization: the assumption of Poisson spike generation. It is commonly held that the temporal dynamic (i.e., non-Poisson) properties of biological spike generation do not significantly affect the white noise characterization of a neuron.[1] However, we show that in simulated white noise experiments, a linear model which drives an integrate-and-fire spiking mechanism is inaccurately characterized by the STA. Furthermore, we show that the integrative behavior of this model can account for some of the changes in STA estimated at different stimulus amplitudes in real neurons. Finally, we propose a new method for recovering the linear temporal "receptive field" governing neural response. We demonstrate through simulation that this approach can correctly estimate the linear kernel of a model neuron, and we also apply our method to real neural data, demonstrating that the recovered linear kernel is fairly stable with changes in stimulus contrast. We thus conclude that the recovered linear kernel may provide a more fundamental functional description of neural behavior, and might well be more directly related to the mechanisms underlying neural response.

Leaky integrate-and-fire model

Our analysis is based on a leaky integrate-and-fire neural model (LIF), consisting of a fixed linear filter K, whose convolution with the input feeds into a leaky integrative spike generator. When the level of this integrator reaches a threshold value, the neuron fires a spike and the integrator is reset to zero. The time evolution of the model membrane potential $V(t)$ is characterized by a single differential equation:

$$\frac{dV}{dt} = -gV(t) + I(t),$$

where g represents the leak conductance, and $I(t)$ is the input current, generated by convolving the input signal $S(t)$ with the fixed kernel K:

$$I(t) \;=\; K * S(t) \;=\; \int_{-\infty}^{0} K(t')S(t-t')dt'.$$

This model has an analytical solution relative to the time of the most recent spike:

$$V(t) = \int_{t^-}^{t} I(t')e^{g(t'-t)}dt', \tag{1}$$

where t^- is the time of occurrence of the last spike before t. This dependence on the time of the previous spike (and past input to the integrator) represents a fundamental departure from L-N-P model described earlier, where the probability of firing a spike is an instantaneous function of the projection of the stimulus onto K.

[1] Note, however, that Arcas et al. have recently examined the behavior of a Hodgkin-Huxley model under white noise stimulation, and have made a number of interesting observations regarding the spike-triggered stimulus ensemble [2].

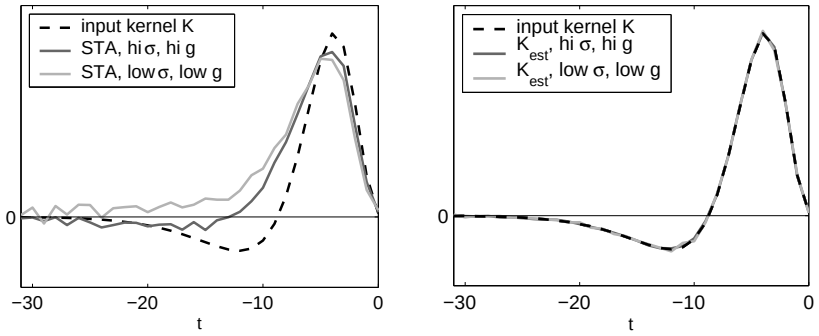

Figure 1: Simulation of integrate-and-fire neuron. **Left:** STA kernels retrieved for two different contrast levels (solid lines), plotted along with the true model kernel (dashed curve). **Right:** Kernels recovered using our algorithm.

Simulation results and comparison

We simulated a white noise analysis experiment with the model described above. In our simulations, the kernel K was chosen to be a 32-sample function whose shape loosely resembles temporal kernels measured in retinal ganglion cells. The leak conductance g was set to two different values, corresponding to exponential decay constants ($1/g$) of 30 or 10 time steps of the simulation. As in classical white noise analysis, we used a random stimulus $S(t)$ that was temporally white (up to the sampling frequency). For our simulations, $S(t_i)$ was drawn from a Gaussian distribution $\mathcal{N}(0, \sigma^2)$ on each time step t_i, where σ reflects the amplitude (or "contrast") of the noise as referred to above. We compute the STA as the average stimulus in the 32 time bins preceding each spike.

Figure 1a shows a plot of the actual kernel K superimposed on the estimated STA for two different values of g. First, note that in both cases, the STA is significantly biased away from K. This bias clearly reflects the integrative spiking mechanism of the LIF model, as the STA should be an unbiased estimator of K if the same input were given to a L-N-P model [5].

Furthermore, the discrepancy between K and the STA depends on the leak conductance g, which is also the reciprocal of the equivalent membrane time constant τ in this model. In particular, large values of g (i.e., small τ) give rise to an STA which is more tightly correlated with K, whereas smaller g leads to loss of high frequencies and larger biases away from K. Physiological evidence suggests that the membrane conductance of neurons increases at higher firing rates, which gives rise to faster dynamics (i.e., smaller values of τ) [e.g., 3, 7, 1]. Thus, bias towards lower temporal frequencies in STAs measured at low contrast is predicted by an integrative spiking model with reduced leak conductance.

Recovering the linear kernel

Assuming that the input to an integrate-and-fire spiking model is determined by projection onto a linear kernel, how can the kernel be recovered from the response to white noise stimuli? Equation (1) provides a deterministic expression for the voltage at any time since

the last spike. The next spike is generated when the membrane potential (equal to the leaky integral of all input since the last spike) reaches the threshold. Expressed in discrete time steps:

$$V(t^+) = V_{th} = \sum_{t_i=t^-}^{t^+} I(t_i) e^{g(t_i - t^+)},$$

where V_{th} is threshold, t^- is the time index of the previous spike and t^+ that of the current spike. If g were known, this expression provides a linear constraint on K, since I is just a convolution of the stimulus with K. Each spike provides such a linear constraint on K, so an estimate of K may be obtained by solving this overconstrained linear inverse problem.

The full problem is not quite so simple because one does not generally know the true value of g. Also, because we are working with a discretized version of the problem, $V(t^+)$ is actually slightly greater than threshold at the time of each spike, rather than exactly equal to it. Finally, there are an additional set of constraints for all time steps within the inter-spike interval: $\{V(t_i) < V_{th}, \quad t^- < t_i < t^+\}$. We define a squared error function for each spike, that enforces the constraint that the voltage threshold is crossed during the time step before the spike occurs:

$$E(K, g) = \left(\lfloor V_{th} - V(t^+;\, K, g) \rfloor^2 + \lfloor V(t^+ - 1;\, K, g) - V_{th} \rfloor^2 \right),$$

where $V(t_i; K, g)$ is the value of V computed at the ith time step assuming particular values for K and g, and $\lfloor \cdot \rfloor$ indicates a floor operation. This error function is summed over all spikes, and the resulting expression is minimized using a gradient descent over K and g.

This algorithm is guaranteed to converge, but the solution may not be unique. In practice, it seems to always converge to the true kernel when the STA is used as the initial estimate for K. Figure 1b shows the kernels recovered from the same simulated spike trains used to estimate the STAs shown in 1a. The model neuron was stimulated for $40,000$ time samples and emitted 1600 and 3100 spikes at low and high contrast, respectively. Model leak conductances were set to the reciprocals of τ = 30 and 10, and the estimates accompanying the recovered kernels were 29.6 and 9.8.

Recovering a kernel from neural data

Our procedure for linear kernel estimation is based on an overly simplistic integrate-and-fire model for neural spike generation. We thus have no strong reason to believe it should be applicable to real neural data. But we note that STA techniques have been used for decades to estimate linear kernels under the assumption of a Poisson spike generator. The integrate-and-fire model, while simplistic, incorporates a dependency on the time of the previous spike which is likely to be relevant in real neurons.

We have applied our procedure directly to data drawn from a monkey retinal ganglion cell [4]. The data were recorded *in vitro*, using a stimulus consisting of 80,000 time samples of full-field 120 Hz flickering binary white noise The stimulus vectors $\vec{s}$ of this sequence are defined over a 25-segment (0.21 sec) time window. Two data sets were recorded, at contrasts of 32% and 64%.

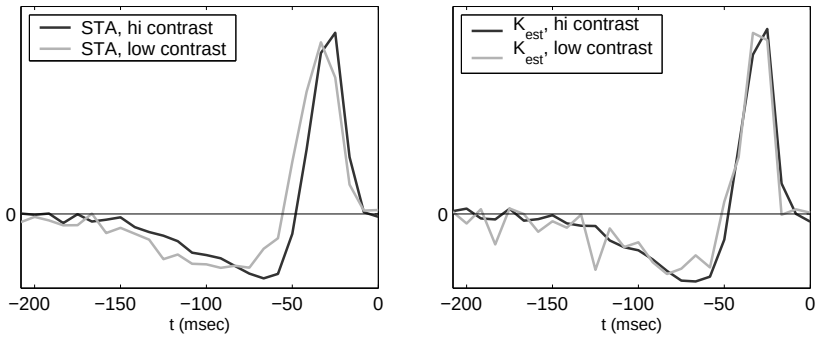

Figure 2: Analysis of *in vitro* ganglion cell data in monkey retina. **Left:** STA estimates based on responses recorded at two different input contrast levels. **Right:** Kernels recovered using our procedure. The associated time constant estimates are 19.1 and 6.5 msec.

Figure 2a shows example STA estimates for both contrast levels. The kernels are quite different, with the STA measured at low contrast containing more low temporal frequencies and having a peak that is shifted earlier in time relative to the high contrast STA. Figure 2b shows the kernels resulting from our estimation procedure. Note that the estimated kernel is now quite stable across different contrasts, a desirable property for a functional description of neural behavior. The recovered time constants of 19.1 msec and 6.5 msec are within ranges considered biologically plausible, although their ratio indicates a greater change with amplitude than is commonly reported for cortical neurons [e.g., 3, 7, 1].

1 Discussion

Our results show that spike generation mechanisms can affect the interpretation of results obtained with white noise analysis. In particular, we have shown that even for a simple integrate-and-fire model, the temporal STA does not accurately recover the temporal linear input kernel. For this model, the magnitude of bias in the STA is influenced by the membrane conductance, which is believed to vary with stimulus strength. This amplitude-dependence of the STA mirrors changes in the STA of real neurons measured at different contrasts, and cannot be captured by an L-N-P model.

Based on this simple LIF model, we have developed a new method for the recovery of the linear kernel integration time constant from responses to white noise stimluli. To our surprise, this kernel estimation procedure recovers a stable linear kernel when applied to data recorded from monkey retinal ganglion cells, and the associated estimates of membrane conductance are within a biologically plausible range.

Our results suggest a mechanistic explanation of the behaviors captured by current functional models of retinal ganglion cells [e.g., 12], in which a nonlinear feedback signal is used to adjust the gain of the neuron. We have also previously shown that nonlinear gain control operations might account for a variety of apparent changes in receptive field properties at different contrast levels [11]. The results presented in this paper suggest that temporal examples of such changes may be due to intracellular mechanisms of spike generation. It would be interesting to test such hypotheses using intracellular measurements.

We are currently exploring the generalization of these results to more realistic models. In particular, we have have found that the incorporation of a voltage floor in the model (corresponding to an ionic reversal potential) produces an STA which is sharper and closer to the true input kernel at high contrast, independent of any changes in membrane conductance. The significance of this phenomenon, along with that of other nonlinearities associated with spike generation, remains to to be analyzed. Finally, while not discussed here, our technique also appears to be quite robust to the presence of noise in the membrane potential.

References

[1] J S Anderson, M Carandini, and D Ferster. Orientation tuning of input conductance, excitation, and inhibition in cat primary visual cortex. *J Neurophysiol*, 84:909–926, 2000.

[2] B A Y Arcas, A L Fairhall, and W Bialek. What can a single neuron compute? In *Advances in Neural Information Processing Systems*, volume 13, pages 75–81, 2000.

[3] L Borg-Graham, C Monier, and Y Fregnac. Visual input evokes transient and strong shunting inhibition in visual cortical neurons. *Nature*, 393:369–373, 1998.

[4] D Chander and E J Chichilnisky. Adaptation to temporal contrast in primate and salamander retina. *J Neurosci*, 21(24):9904–9916, 2001.

[5] E J Chichilnisky. A simple white noise analysis of neuronal light responses. *Network: Computation in Neural Systems*, 12(2):199–213, 2001.

[6] E deBoer and P Kuyper. Triggered correlation. In *IEEE Transact. Biomed. Eng.*, volume 15, pages 169–179, 1968.

[7] J Hirsch, J M Alonso, C R Reid, and L Martinez. Synaptic integration in striate cortical simple cells. *J Neuroscience*, 15:9517–9528, 1998.

[8] J P Jones and L A Palmer. The two-dimensional spatial structure of simple receptive fields in the cat striate cortex. *J Neurophysiology*, 58:1187–11211, 1987.

[9] K J Kim and F Rieke. Temporal contrast adaptation in the input and output signals of salamander retinal ganglion cells. *J. Neurosci.*, 21(1):287–299, 2001.

[10] O Schwartz, E J Chichilnisky, and E P Simoncelli. Characterizing neural gain control using spike-triggered covariance. In T G Dietterich, S Becker, and Z Ghahramani, editors, *Adv. Neural Information Processing Systems*, volume 14, Cambridge, MA, 2002. MIT Press. Honorable mention, best student paper award.

[11] O Schwartz and E P Simoncelli. Natural signal statistics and sensory gain control. *Nature Neuroscience*, 4(8):819–825, August 2001.

[12] R M Shapley and J D Victor. How the contrast gain control modifies the frequency response of cat retinal ganglion cells. *J. Physiol. (Lond)*, 318:161–171, 1981.

[13] S M Smirnakis, M J Berry, David K Warland, W Bialek, and M Meister. Adaptation of retinal processing to image contrast and spatial scale. *Nature*, 386:69–73, March 1997.