Supplemental Material for Functional Identification of Spike-Processing Neural Circuits

Aurel A. Lazar[1], **Yevgeniy B. Slutskiy**[1*]

[1]Department of Electrical Engineering, Columbia University, New York, NY 10027.

For all GLM comparisons we used the MATLAB code written and distributed by the Pillow group at the University of Texas at Austin (Pillow, 2010), with only minor changes made to the code in order to compare to our methods.

1 Choice of Bases

In order to properly compare identification results between our methodology and the GLM framework, we used the same filter kernels and fixed the input signals used in identification. We also fixed the basis functions with respect to which both methods reconstruct the kernels. The latter was needed since the GLM typically employs basis functions that favor very particular kernels: those that oscillate rapidly close to the origin and have a very coarse structure further away from the origin (although the method can work with any adequately chosen basis (Pillow et al., 2008)). These kernels are shown in Fig. S1.

While such an assumption about the filters might hold for some neural circuits, it is best to use a basis that can represent an arbitrary function on a given time

[*]The authors' names are alphabetically ordered.

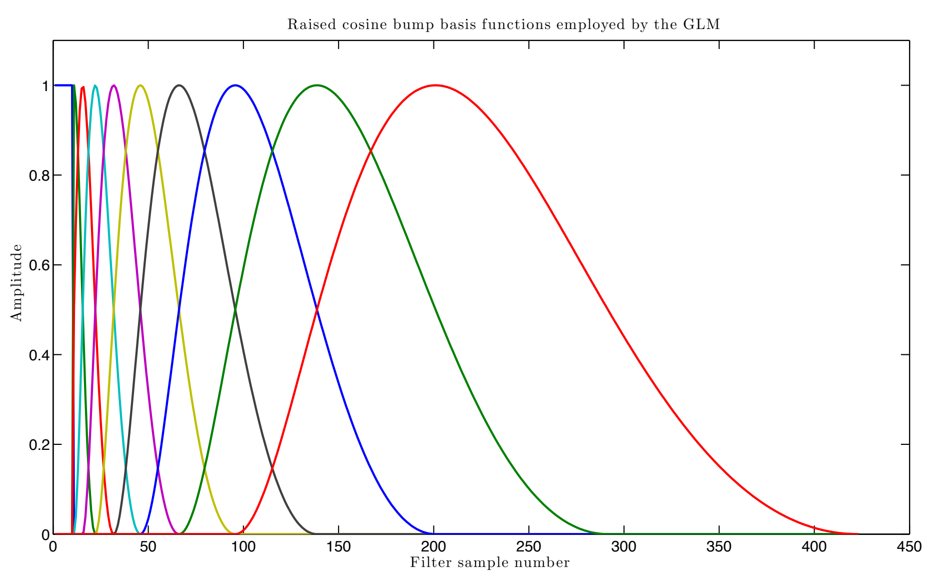

Figure S1: **Basis functions employed by the GLM favor kernels that oscillate rapidly close to the origin and have a very coarse structure further away from the origin**

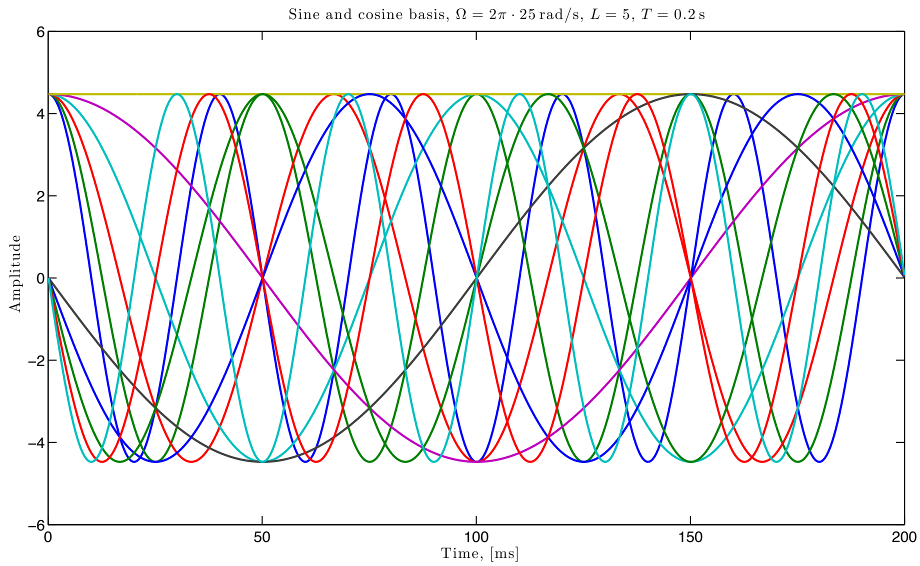

Figure S2: **Basis functions employed by the CIM can represent an arbitrary function $u \in \mathcal{H}$**

interval, when studying an unknown system. The basis functions employed in our methods are orthonormal functions as determined by the reproducing kernel Hilbert space (RKHS) $\mathcal{H}$ employed in identification. For temporal functions in the space of trigonometric polynomials, this set of basis is given by

$$e_l(t) = \frac{1}{\sqrt{T}} \exp\left(\frac{jl\Omega t}{L}\right), \qquad l = -L, -L+1, ..., L \quad , \quad t \in [0, T]$$

where Ω is the bandwidth, L is the number of basis (the order of the space). These functions are defined on the interval $[0, T]$, where $T = 2\pi L/\Omega$, and can describe an arbitrary function $u(t)$, $u \in \mathcal{H}$ (see also Definition 1 in the paper).

Clearly, $e_l(t)$, l=-L,-L+1,..., L, are complex functions. However, for real signals $u \in \mathcal{H}$, we have $e_{-l} = \overline{e_l}$ and an equivalent basis is given by a combination of sine and cosine functions, that can also be used in the GLM framework. One example of such basis for $\Omega = 2\pi \cdot 25\,\mathrm{rad/s}$, $L = 5$ and $T = 200\,\mathrm{ms}$ is shown in Fig. S2.

2 Comparison between GLM and CIM

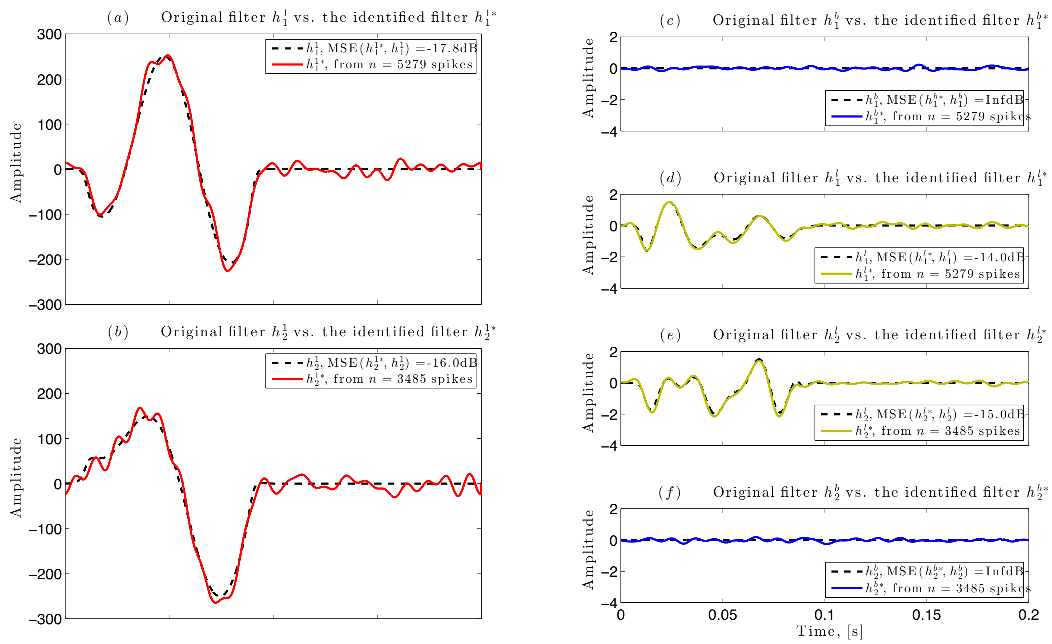

Figure S3: **GLM identification results for the case when the underlying spike generator is described by a nonlinearity and a Poisson spike generator. Note that the kernels are not normalized and the temporal structure of each kernel is recovered well.**

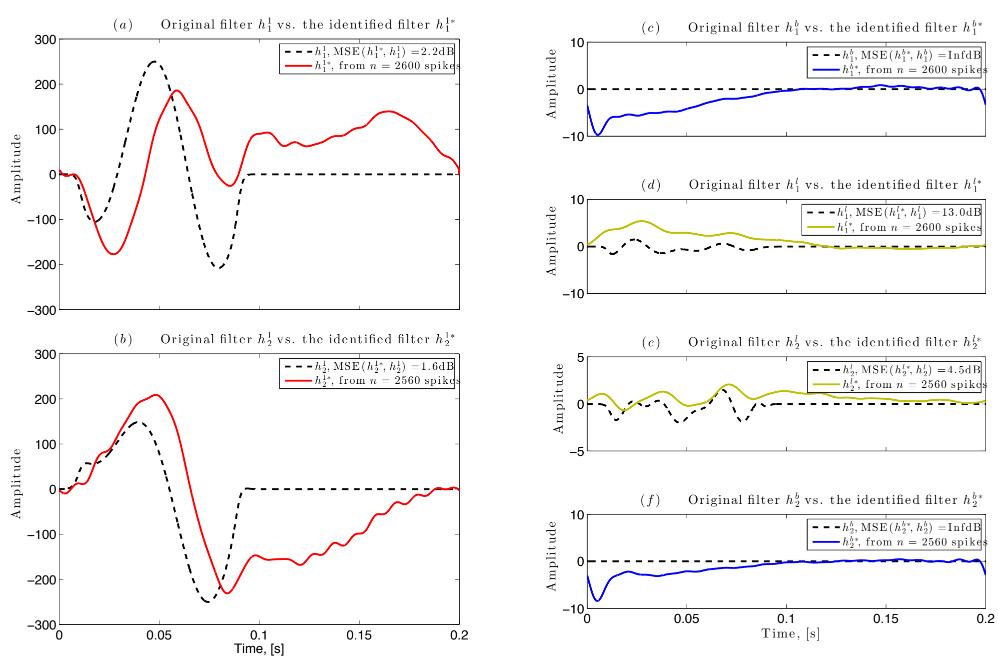

Figure S4: **GLM identification results when the underlying spike generator is an IAF neuron and the nonlinearity is assumed to be exponential.**

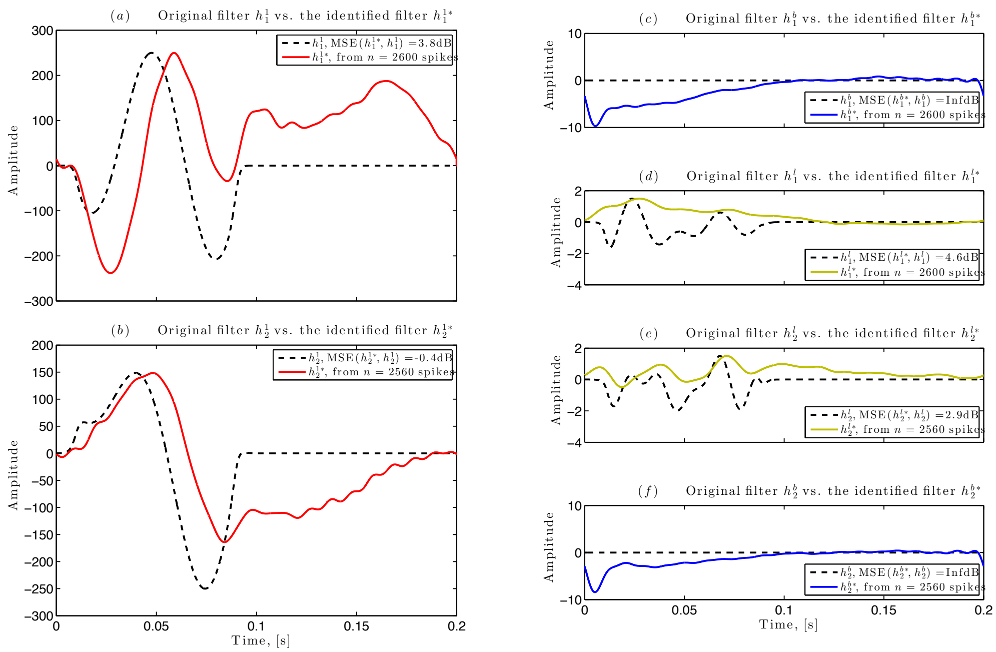

Figure S5: **Same as above, with the kernels normalized.**

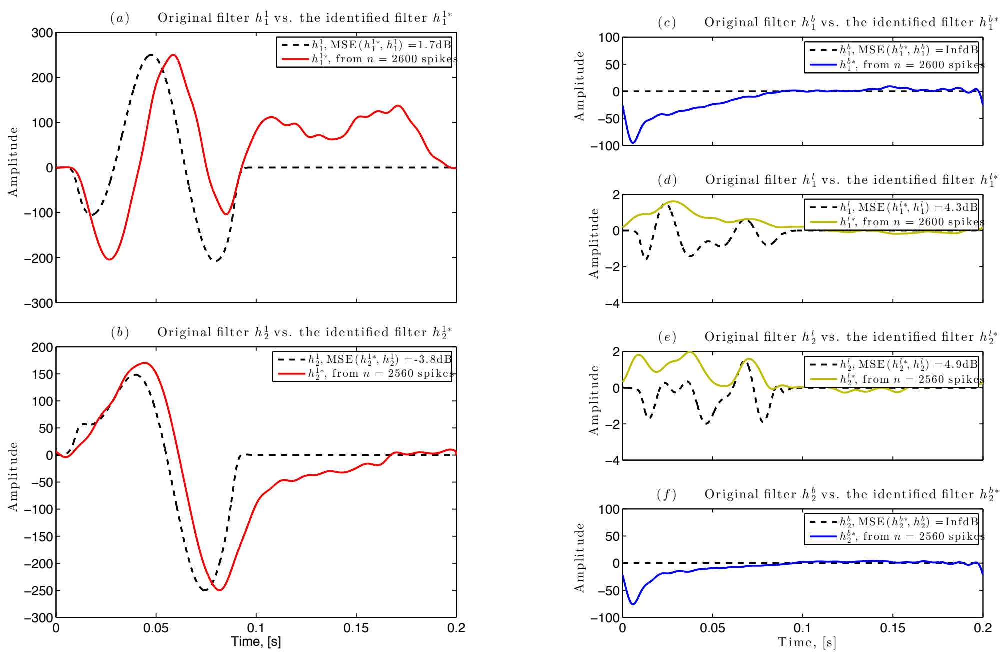

Figure S6: **GLM identification results when the underlying spike generator is an IAF neuron and the nonlinearity is assumed to be log-exponential. The kernels are normalized.**

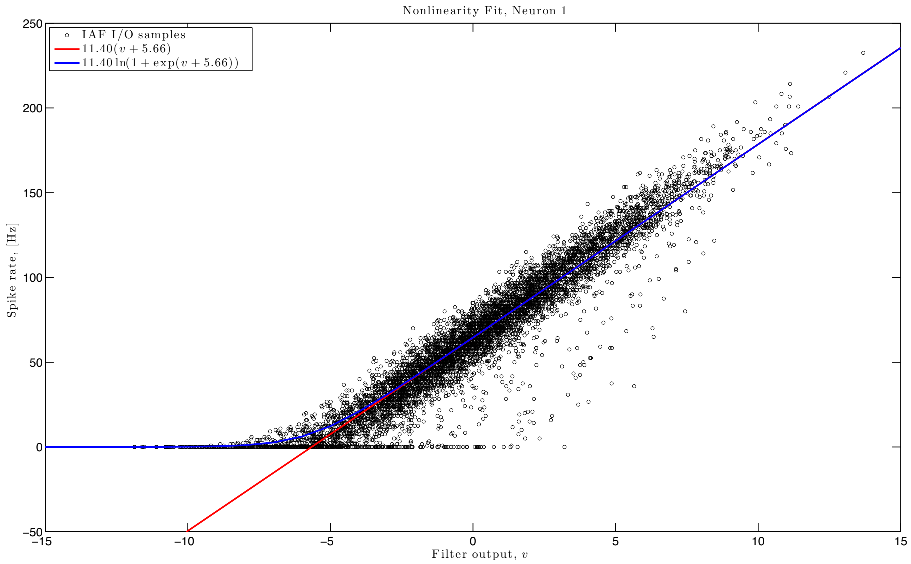

Figure S7: **For GLM simulations with a log-exponential nonlinearity, the nonlinearity of both neurons was fit using a function of the form $f(x) = \ln(1 + \exp(x))$. Shown here is the nonlinearity of the first neuron. We would like to point out however, that while it is certainly possible to fit the nonlinearity in simulations (since one has access to the actual filter output), it may be hard or even impossible to do so in a real biological system. This is because the aggregate output of all filters is not known a priori (since the filters are not known). The GLM algorithm typically uses the spike-triggerred average (STA) as an initial guess for the feedforward filter. While the STA can be certainly computed for continuous signals, it presents problems when working with spikes.**

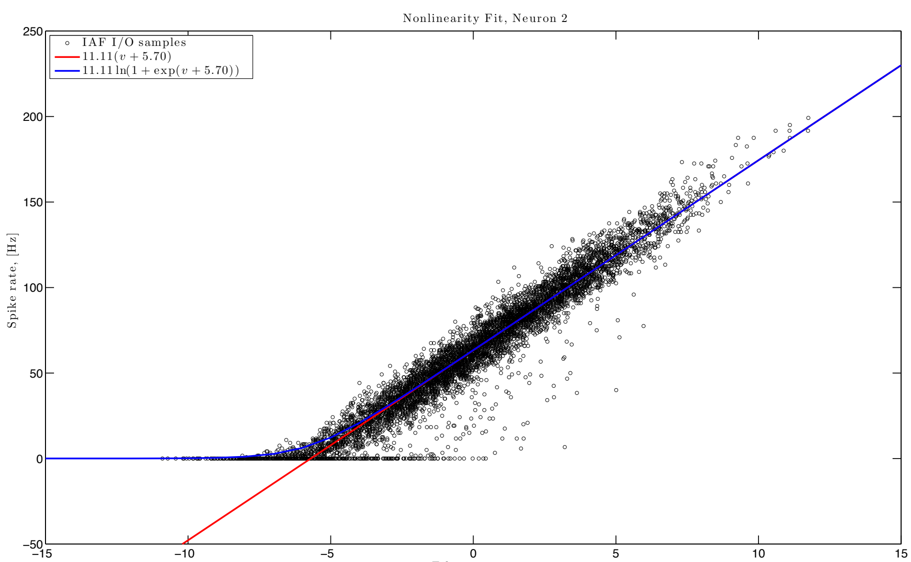

Figure S8: **Same as above for the second neuron.**

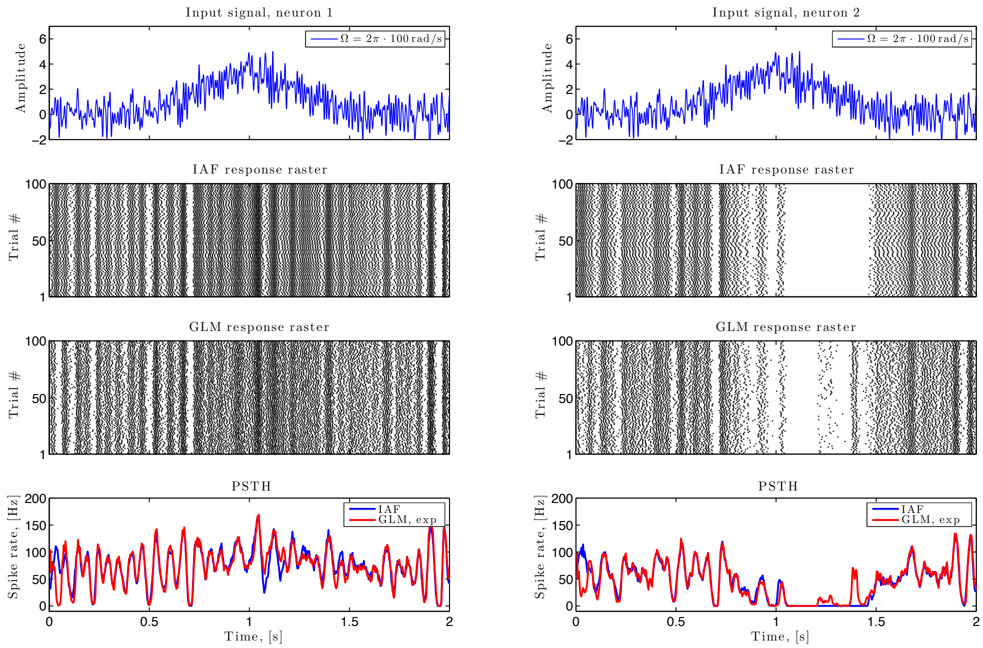

Figure S9: **Another example of a stimulus for which the GLM prediction breaks down.**

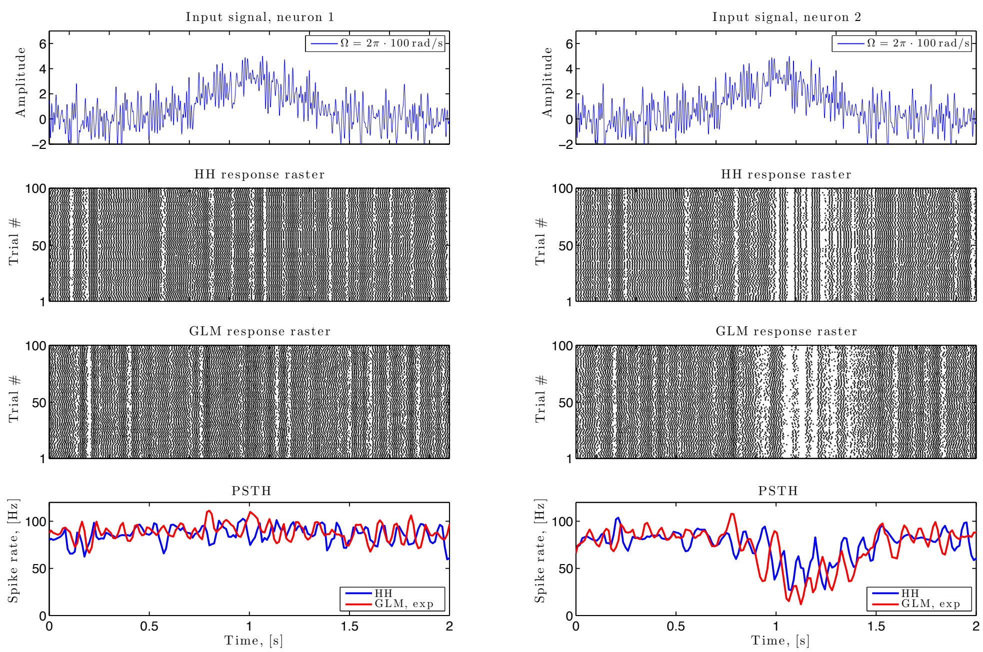

Figure S10: **Similarly, the GLM prediction breaks down when using a Hodgkin-Huxley neuron instead of IAF neuron.**

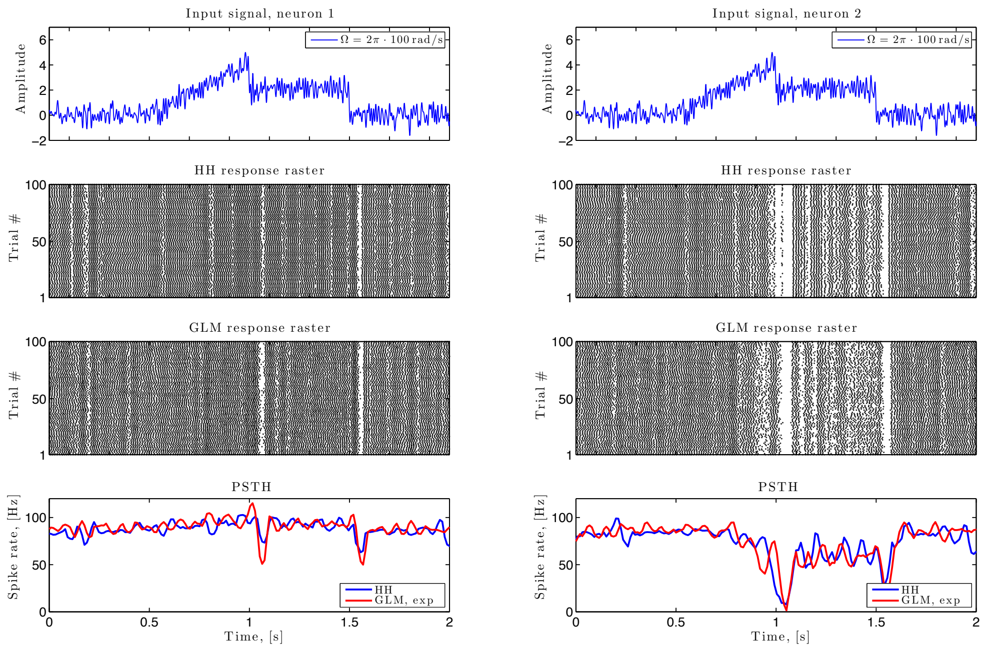

Figure S11: **Another example of the same.**

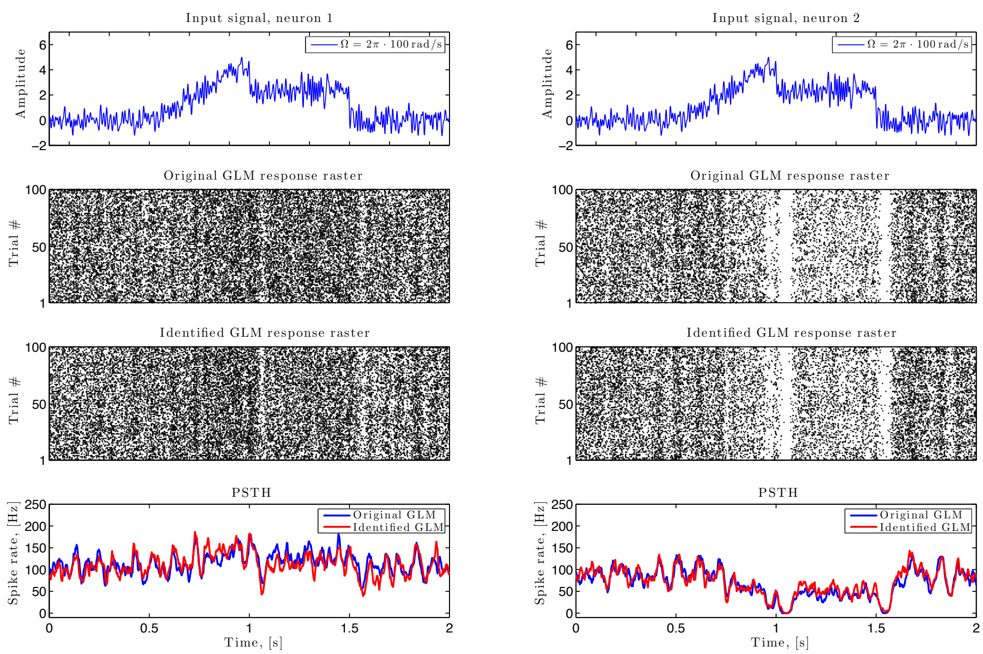

Figure S12: **The GLM prediction does not break down if the underlying model of spike generation consists of a nonlinearity and a Poisson spike generator.**

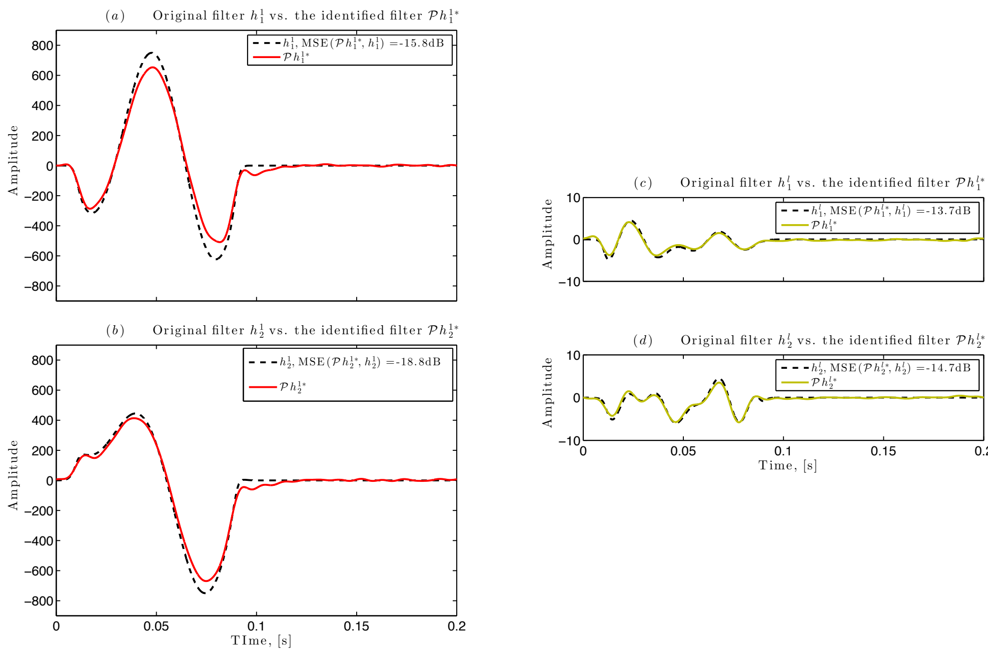

Figure S13: **Kernels identified by the proposed methodology when the spike generator is a Hodgkin-Huxley model.**

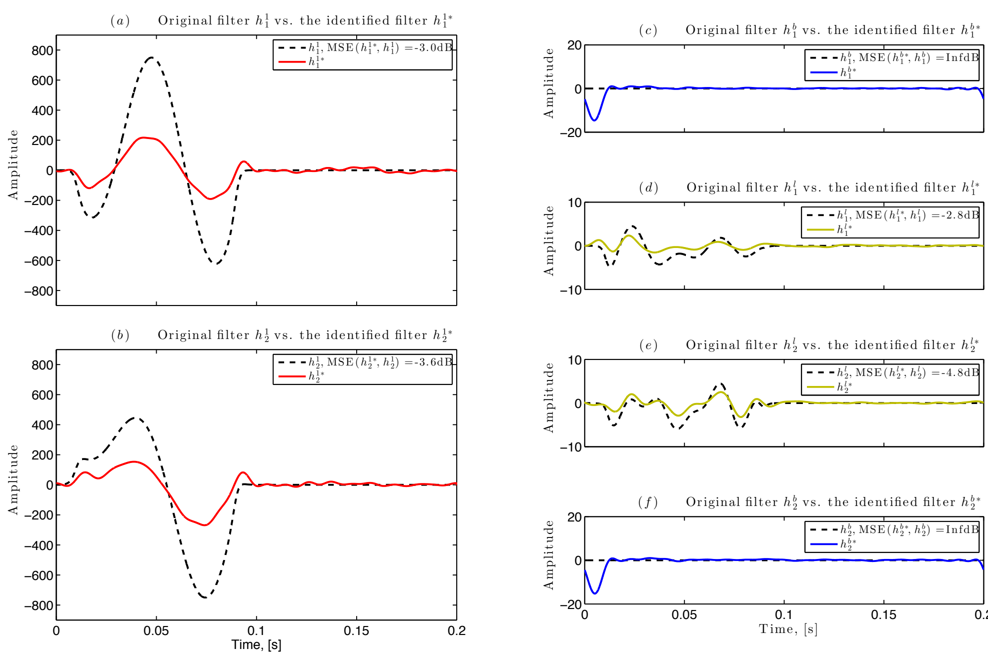

Figure S14: **Kernels identified by the GLM when the spike generator is a Hodgkin-Huxley model.**

References

Pillow, J. (2010). Generalized Linear Model (GLM) point process model for spike trains. `http://pillowlab.cps.utexas.edu/code_GLM.html` (Retrieved April 15, 2013).

Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J., and Simoncelli, E. P. (2008). Spatio-temporal correlations and visual signaling in a complete neuronal population. Nature, 454:995–999.