

Infering spikes from Calcium Fluorescence Signal

January 10, 2008

1 The Model

We assume that each spike produces a jump in the Calcium Fluorescence. Between the jumps, the Calcium Fluorescence signal decays exponentially between the jumps. A Gaussian noise is added to the observed Calcium signal. The task is to infer the underlying spike train from the observed noisy Calcium fluorescence signal.

Although the model for Calcium dynamics and noise is overly simplistic, it is a reasonable starting point and a good benchmark against which more complicated models are evaluated.

The spike train can be written as:

$$r_t = \sum_{i=1}^n \delta(t - t_i) \quad (1)$$

where t_i is the spike time and n is the total number of spikes. The dynamics of the Calcium fluorescence can be described as a first order linear differential equation:

$$\tau_{Ca} \frac{d[Ca]}{dt} = -[Ca] + A r_t \quad (2)$$

where τ_{Ca} is the time constant of calcium decay (typically between a few hundred milliseconds to a few seconds) and A is the magnitude of jump due to each spike. The observed noisy signal can be written as:

$$S_t = [Ca]_t + \xi \quad (3)$$

2 Deconvolution method

Solving the equation (2) and plugging the results into equation (3) results in:

$$S_T = \left(A \int_0^T r_t e^{(t-T)/\tau_{Ca}} dt \right) + \xi \quad (4)$$

which can be interpreted as the spike train convoluted by the exponentially decaying Calcium filter, plus a gaussian noise. If we discretize the equation with Δt time steps and treat the variables S_T and r_t as vectors, the equation reads:

$$\vec{S} = \mathbf{K} \vec{r} \Delta t + \xi \quad (5)$$

The matrix K in the above equation is constructed as follows: Suppose that the (discretized) Calcium filter is the vector f :

$$f = (A, A e^{\Delta t/\tau_{Ca}}, A e^{2\Delta t/\tau_{Ca}}, \dots, A e^{n\Delta t/\tau_{Ca}}) \quad (6)$$

The matrix \mathbf{K} can be constructed from the vector f as:

$$\mathbf{K} = \begin{pmatrix} f_1 & 0 & 0 & \dots & 0 \\ f_2 & f_1 & 0 & \dots & 0 \\ f_3 & f_2 & f_1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ f_n & f_{n-1} & f_{n-2} & \dots & f_1 \end{pmatrix} \quad (7)$$

So the term $\mathbf{K} \vec{r} \Delta t$ in equation (5) implements the discretized convolution of the spike train and the Calcium filter. Assuming that ξ is gaussian white noise, the problem has a simple analytic solution. If we write the likelihood of this model using the log of the Gaussian probability density function we have:

$$\log p(\vec{r} | \vec{S}, \mathbf{K}) = c_1 - c_2 \sum_t (S - \mathbf{K} \vec{r})^2 \quad (8)$$

where c_1 and c_2 are constants that do not depend on \vec{r} and \mathbf{K} . The vector \vec{r} that maximizes the above likelihood is clearly:

$$\hat{r} = (\mathbf{K}^T \mathbf{K})^{-1} \mathbf{K}^T \vec{S} \quad (9)$$

What explained so far is the core of the so-called deconvolution methods that have been used by Yaksi and Friedrich [1] and others for inferring the spikes from the Calcium fluorescence signal. Of course, prior to deconvolution, different noise elimination methods are necessary to make the signal cleaner, while preserving the spiking information as much as possible [1].

3 Constrained deconvolution

In the deconvolution method, it is assumed that the parameters of the Calcium filter (A and τ_{Ca}) are known for all the neurons. However, these parameters vary cell to cell and are not known *a priori*. Thus, a more optimized algorithm for inferring the spikes should treat the parameters of Calcium filter as unknown variables. Moreover, if we take into account a few properties about the spiking activity of the neurons, the algorithm may improve considerably and potentially go beyond the sampling rate of the Calcium signal. The relevant properties of spike trains are:

1. The spikes are sparse i.e. there is no spike in most of the time bins of the spike train.
2. The spikes are, obviously, positive values.
3. The spikes are point processes i.e. the bins of the spike train are either zero or one.

In order to incorporate the first property of the spike train into the the convolution method, we can simply add a sparseness penalty to our log likelihood function (equation (8)). The second property can be addressed by constraining the entries of \vec{r} to be positive. Together, these result in the following optimization problem:

$$\hat{r} = \arg \max_{r \geq 0} \left(\sum_t \left(- (S - \mathbf{K} \vec{r})^2 - \lambda |\vec{r}| \right) \right) \quad (10)$$

The second term in the right hand side operates as a sparseness penalty and λ is a fixed coefficient. Although the optimization problem doesn't have an analytical solution anymore, it is still convex and can be solved with the standard numerical methods. However, the non-negativity constraint and the non-differentiability of $|\vec{r}|$ makes the convergence very slow. In order to make the optimization faster and more efficient, the non-negativity constraint can be replaced by another penalty function. This new penalty function should tend to infinity as r tends to zero, so that r can never get negative. Such a penalty function is called "barrier function" in the optimization literature. By relaxing the non-negativity constrain, we can also drop the absolute value function ($|\cdot|$) from the sparseness penalty and the whole problem would be convex and smooth:

$$\hat{r} = \arg \max \left(\sum_t \left(- (S - \mathbf{K} \vec{r})^2 - \lambda_1 \vec{r} - \lambda_2 f(\vec{r}) \right) \right) \quad (11)$$

where λ_1 and λ_2 are constant coefficients and $f(\cdot)$ is the barrier function. A suitable choice of the barrier function is $f(x) = -\epsilon \log(x)$ with ϵ being a

small coefficient.

The third property of the spike train mention above (being a point process) can potentially increase the resolution of the algorithm further than the sampling rate of the Calcium fluorescence signal. The work in this direction is currently in progress.

An advantage of the barrier function is that it serves also as a non-linear noise filter for free, so that no filtering of the data is needed before the optimization.

After finding the first estimate of the spike train (\vec{r}^*), the same procedure can be repeated for improving the estimate for the Calcium filter (\vec{f}). In order to do this estimation, we treat the spike train as a known parameter and construct the convolution matrix based on it. Instead of the sparseness penalty and the barrier function, a smoothness penalty can be used for the estimation of the Calcium filter. Taken together, each iteration of the estimation procedure improves the estimated values of the spike train and the Calcium filter until they converge to their optimal value. Thus, an important advantage of this method is that it does not assume that the parameters of the calcium signal are given *a priori*.

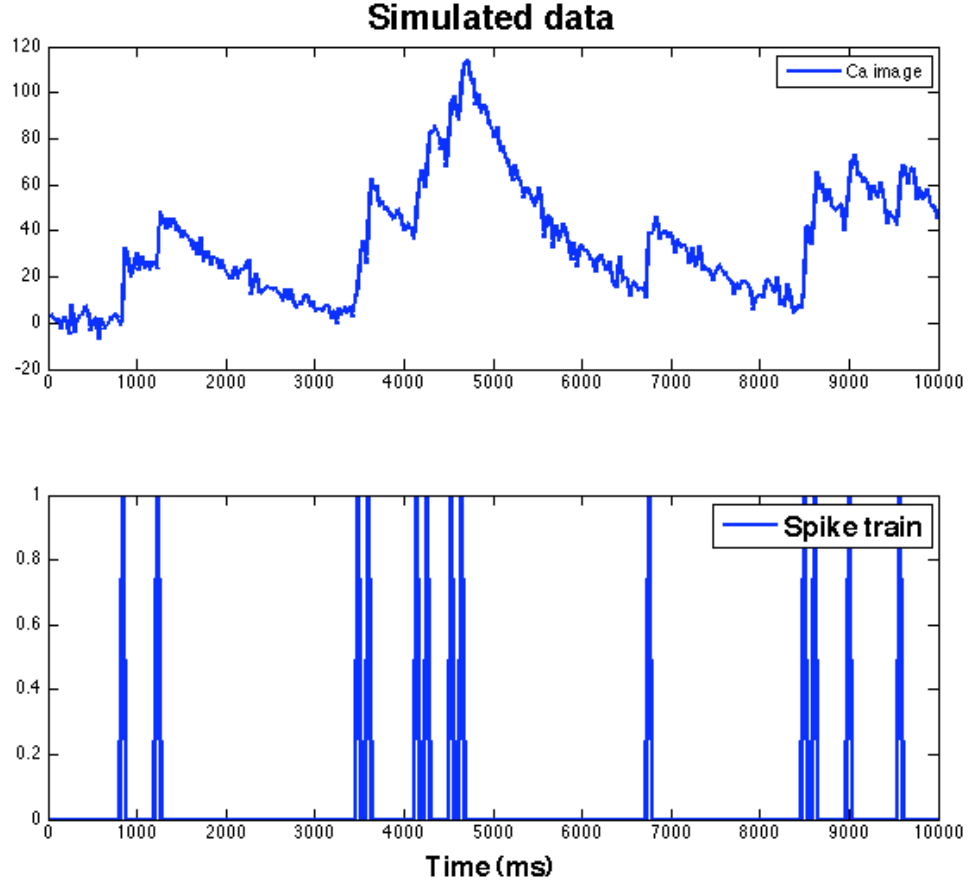
4 Results

Here, we compare the performance of the Constrained Deconvolution method (CDC) against the simple deconvolution (DC) and also a simple thresholding (TH). As it is common in the experimental studies, for the thresholding method we first calculate the $\frac{\Delta S}{S}$ vector from the Calcium signal to enhance the contrast. Then we apply a threshold to detect abrupt changes that are supposedly due to a spike. We compare these three methods for both simulated and recorded data.

Figure 1 shows the Calcium fluorescence signal and the corresponding spike train obtained from a numerical simulation. Figure 2 shows the inferred spike train by the standard methods. As can be seen (figure 2, top) the DC method results in a very noisy estimate of the spike train. The TH method also does a poor job in finding the timing of the spikes. It is also unable to find the number of spikes that occur in a frame of the Calcium movie. Figure 3 shows the result of the CDC method. As can be seen both the spike train and the Calcium filter can be estimated much more faithfully in comparison to the previous methods.

Figure 4 shows the recorded Calcium signal and the associated spike train of a cortical neuron in Layer II/III of the visual cortex of mouse. The frame

Figure 1: Simulated data



rate of the movie is 200 milliseconds and the used dye is Fura 2. As can be observed, the neuron shows bursting activity during the recording period. Figure 5 shows the estimated spike trains by the conventional methods and Figure 6 shows the results of the CDC method. Here again the CDC method has a much better performance compared to the standard methods.

5 Further Directions

By adding the third constraint discussed above the resolution of the CDC method can be higher than the frame rate of the Calcium fluorescence movie. A proposed statistical method for this purpose is the "Projection Pursuit" algorithm. The work in this direction is currently in progress.

Saturation of the dye does not seem to be a confounding factor in this sample of recorded neuron, although it shows high-frequency bursting ac-

Figure 2: Standard methods

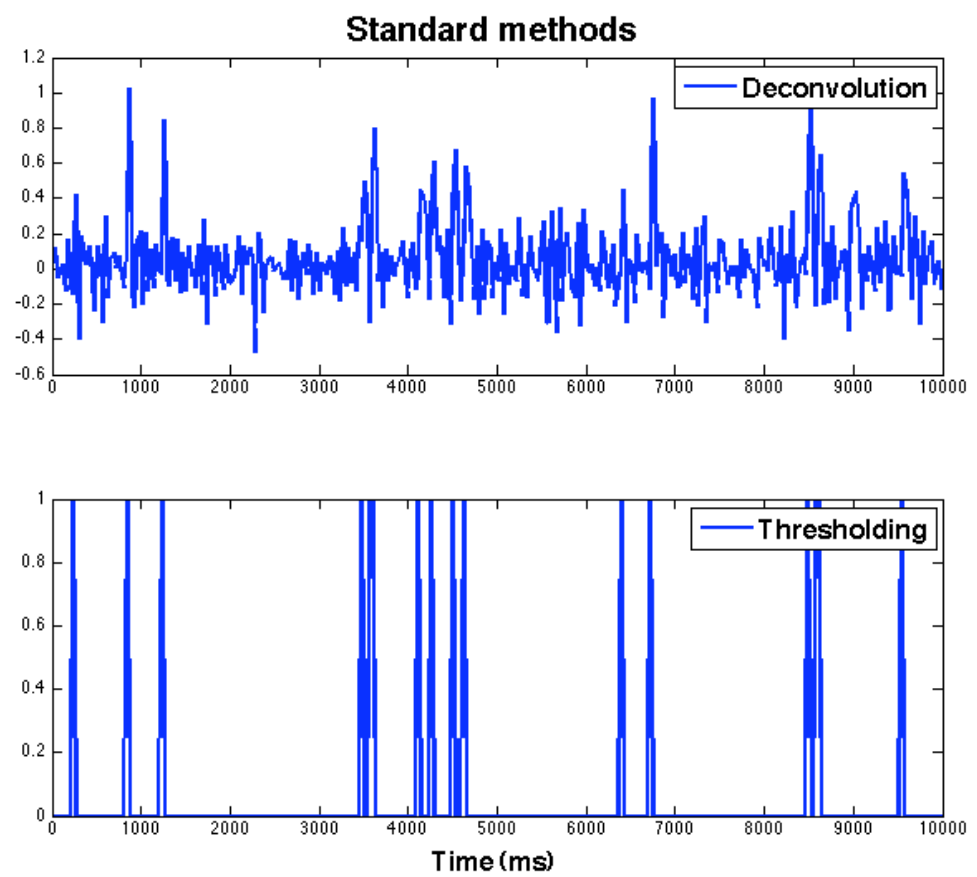


Figure 3: Constraint Deconvolution

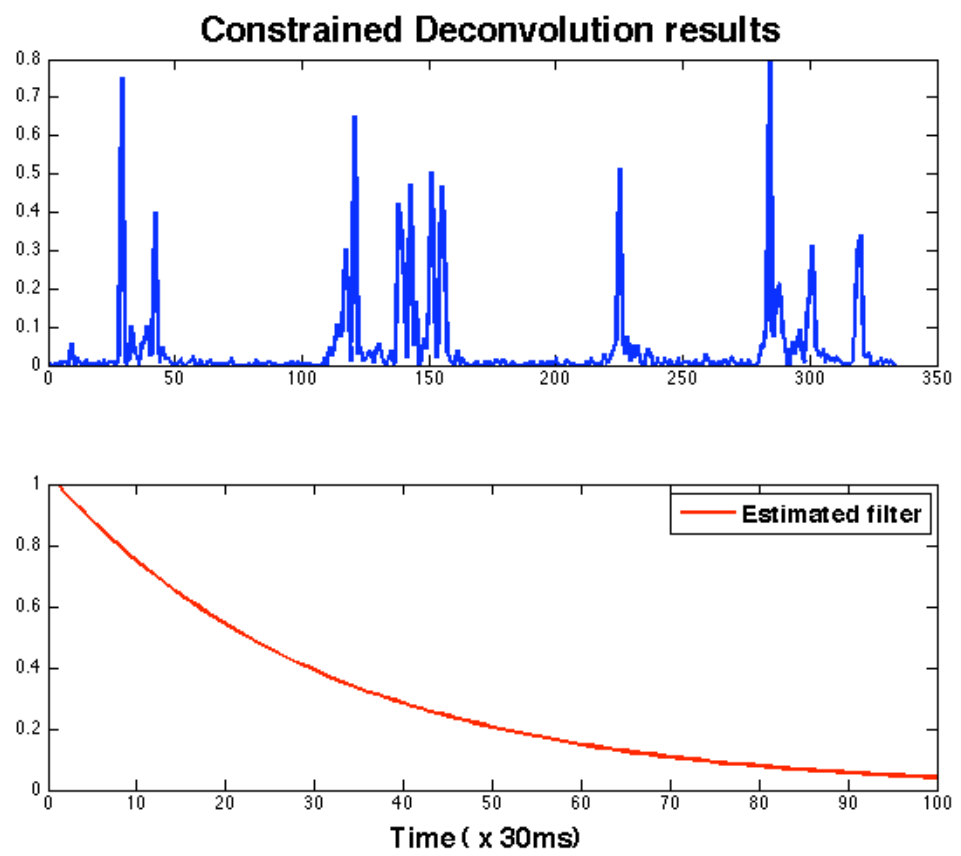


Figure 4: Experimental data

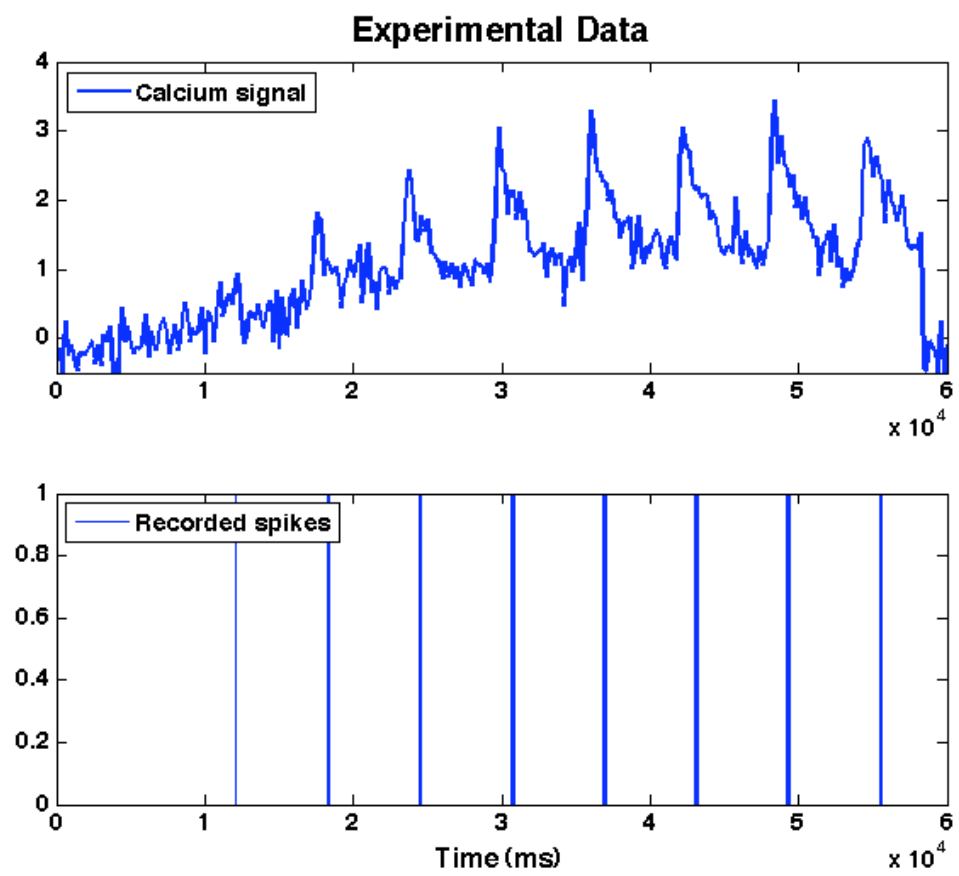


Figure 5: Standard methods

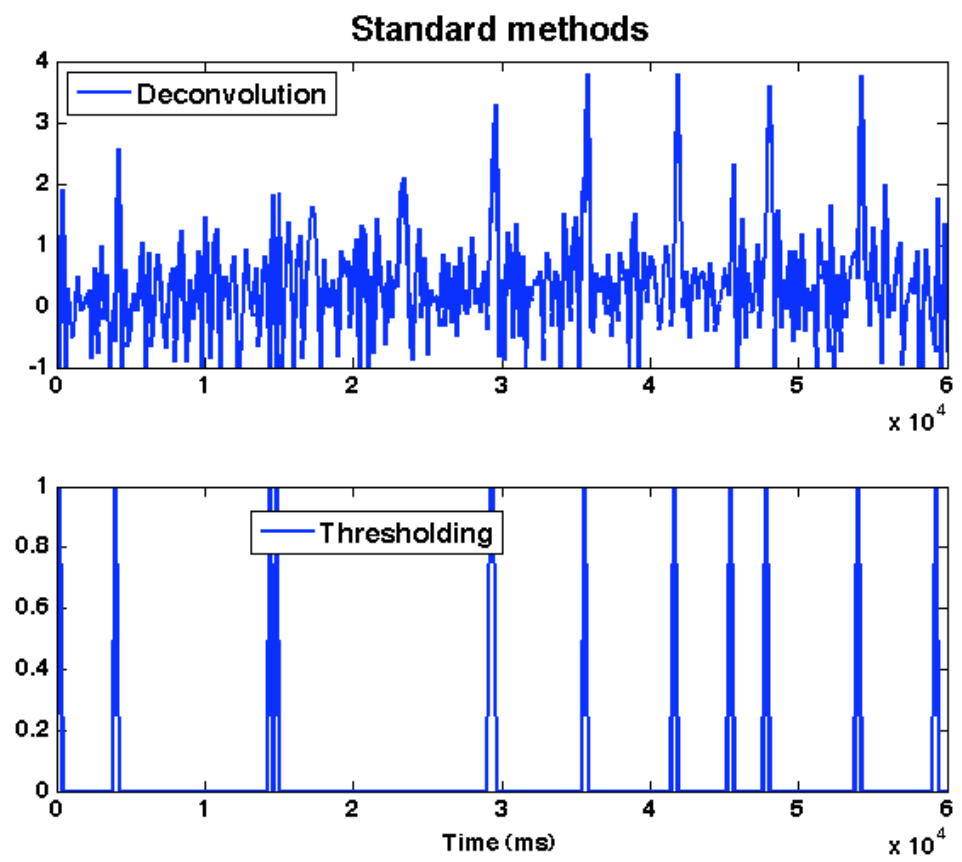
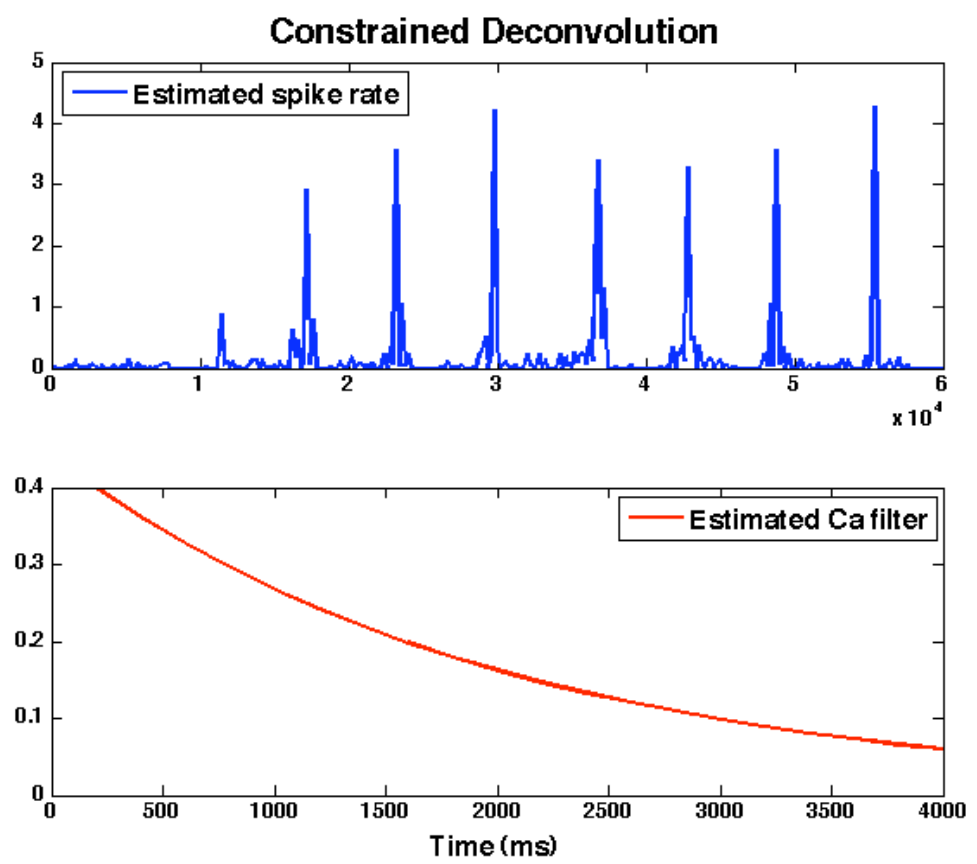


Figure 6: Constraint Deconvolution



tivity. Nevertheless, incorporating this property of the dye into the method can improve the accuracy of the estimates. This can be done by applying a saturating function on equation 5, prior to the numerical optimization. The caveat would be loosing the convexity of the problem which can result in local maxima. However, with a suitable initial guess (near the global maximum) the local maxima can be avoided. The current linear model can be used as the initial guess in that case.