Brief summary of the deconvolution problem

A brief summary of denoising and deconvolving calcium imaging data using FOOPSI approaches.

FOOPSI

A general form of the FOOPSI problem

```
\text{minimize}_{\bm{c}, \bm{s}} ~ \frac{1}{2} \|\bm{y}-\bm{c}\|_2^2 + \lambda
\|\bm{s}\|_1
~
\text{subject to } c_t = \sum_i^p \gamma_{i=1} c_{t-i}+s_t \text{ with } s_t=0
\text{ or } s_t\geq s_{min}
```

where $|\cdot|1$ means [L-1 norm](#) and $|\cdot|2$ means [L-2 norm](#). $\bm{y}$ is the noisy raw calcium trace and $\bm{c}$ is the desired denoised trace. We use an auto-regressive (AR) $c_t = \sum_i^p \gamma{i=1} c{t-i}+s_t$ model to model $\bm{c}$ and the corresponding spiking activity $\bm{s}$. A more general model of $\bm{c}$ and $\bm{s}$ is $\bm{c} = \bm{s} \ast \bm{h}$, where $\ast$ means convolution. The convolution kernel $\bm{h}$ has a interpretation of calcium responses of a neuron after firing a spike.

The AR model is equivalent to special kernel $\bm{h}$. For example, AR-1 model ($p=1$) correspond to

```
h(t) = \exp(-t/\tau)
```

and AR-2 ($p=2$) model correspond to

```
h(t) = \exp(-t/\tau_d) - \exp(-t/\tau_r).
```

Details of these AR model can be found in the literatures [1][2][3].

Different formulations of FOOPSI

We described the most general format of FOOPSI above. In practice, people use different formulations of FOOPSI for their own data. Here I brief summarize different formulations.

original FOOPSI [1]

The original FOOPSI solves the following problem

```
\text{minimize}_{\bm{c}, \bm{s}} ~ \frac{1}{2} \|\bm{y}-\bm{c}\|_2^2 + \lambda
\|\bm{s}\|_1
~
\text{subject to } c_t = \sum_i \gamma_{i=1}^p c_{t-i}+s_t
```

where λ has to be pre-specified by users and γ needs to be estimated as well.

This formulation has the simplest format, but the parameter selection is a pain.

constrained FOOPSI [2]

```
\text{minimize}_{\bm{c}, \bm{s}} ~ \|\bm{s}\|_1
~
\text{subject to } c_t = \sum_i \gamma_{i=1}^p c_{t-i}+s_t \text{ and } \|\bm{y}-
\bm{c}\|_2^2 = \sigma^2T
```

This formulation is equivalent to solving a original FOOPSI problem except that it chooses λ automatically by constraining the residual sum of squares (RSS) to be σ^2T, where σ is the noise level and can be estimated from the raw data.

thresholded FOOPSI [3]

Both the upper two approaches uses L-1 penalty to enforce the sparsity of the spiking signal $\bm{s}$. The advantage of these two formulations is that the optimization problems are convex. However, the estimated $\bm{s}$ usually contains many false positive spikes with small amplitudes, and these false positives affect the accurate estimation of the amplitudes of those true spikes. We can definitely remove them by thresholding $\hat{\bm{s}}$ in the downstream analysis, but it won't help with the estimation of true spikes in this step.

Thus a thresholded-FOOPSI was proposed to integrate the thresholding step into deconvolution by setting a hard threshold of spike size $s{min}$. This $s{min}$ is usually chosen as 3σ, where the noise level σ is estimated from the raw data.

```
\text{minimize}_{\bm{c}, \bm{s}} ~ \frac{1}{2} \|\bm{y}-\bm{c}\|_2^2
~
\text{subject to } c_t = \sum_i \gamma_{i=1}^p c_{t-i}+s_t \text{ with } s_t=0
\text{ or } s_t\geq s_{min}
```

This other automatic way of choosing s_{min} is to use a similar approach of constrained FOOPSI by letting RSS equal to σ^2T.

why OASIS

FAST and *ONLINE*

For the original FOOPSI and the constrained FOOPSI problem, they were usually solved using generic convex optimiation solvers to solve, which is slow and has to use the complete time series of $\bm{y}$ to optimize. OASIS is optimized for solving FOOPSI problems by utilizing the special structure of AR model. The algorithm is FAST and, more importantly, it allows ONLINE processing of the data.

As for the thresholded FOOPSI, it is not convex. Thus we don't have algorithm to solve this problem at the beginning. When OASIS was developed, we made very tiny changes to the code and it outputs reasonable good results. Also there is no guarantee on the global optimums, the results are usually good in both simulated data and real data.

Besides solving the above optimization problems, the implementation of OASIS also include supports for automatic optimizing parameters in those problems, such as λ, γ and the baseline b.

References

[1] Vogelstein, J.T., Packer, A.M., Machado, T.A., Sippy, T., Babadi, B., Yuste, R. and Paninski, L., 2010. Fast nonnegative deconvolution for spike train inference from population calcium imaging. *Journal of neurophysiology*, *104*(6), pp.3691-3704.

[2] Pnevmatikakis, E.A., Soudry, D., Gao, Y., Machado, T.A., Merel, J., Pfau, D., Reardon, T., Mu, Y., Lacefield, C., Yang, W. and Ahrens, M., 2016. Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron*, *89*(2), pp.285-299.

[3] Friedrich, J., Zhou, P. and Paninski, L., 2017. Fast online deconvolution of calcium imaging data. PLoS computational biology, 13(3), p.e1005423.