joshua vogelstein <joshuav@gmail.com>

JN-01073-2009 Fast non-negative deconvolution for spike train inference from population calcium imaging

6 messages

dlinden@jhmi.edu <dlinden@jhmi.edu> **Mon, Jan 4, 2010 at 10:06 AM**
To: joshuav@jhu.edu
Cc: dlinden@jhmi.edu

Dear Josh,

Please find attached the comments of expert reviewers for your manuscript entitled "Fast non-negative deconvolution for spike train inference from population calcium imaging." Although the reviewers were generally positive, they indicated significant concerns that require your attention before the manuscript can be accepted. A final decision concerning the acceptance of your manuscript will be made after the original reviewers have examined the revisions and provided me their assessment of the revised manuscript and your responses to the reviewers' critiques. Please be aware that this invitation does not guarantee eventual acceptance of your manuscript.

The revision is due by 13th Jul 2010, six (6) months from today. If you do not request an extension to this deadline or submit a revised manuscript within 6 months, we will consider the manuscript withdrawn from submission to Journal of Neurophysiology. However, if you wish to submit the manuscript to another journal within the 6-month time frame, you must officially withdraw your manuscript from the journal Journal of Neurophysiology. Please contact at mcapers@the-aps.org if you wish to withdraw your manuscript.

:

To successfully submit a revision, follow the instructions below.

THE FILES THAT YOU UPLOAD NOW WILL BE THE FILES USED FOR PUBLICATION IF THE MANUSCRIPT IS ACCEPTED.

To submit a revision, first, create your responses to reviewers in a DOC, RTF or PDF format; create your article file - clean version (all text in black) in DOC, RTF or LaTeX format (PDF will not be accepted) and your article file - -redlined version in a DOC, RTF or PDF format. When you have all your new files ready, use the web address below to start the submission for your revision:
http://jn.msubmit.net/cgi-bin/main.plex

Validate the metadata, especially authors' names, the title, and the abstract.

FILES:
- Provide your article file without figures in a DOC or RTF format.
- Upload either a single file or separate files for your figures in a PDF, EPS, PPT, or TIF format (PDF preferred).

Use the REMOVE Files tab to remove files that were uploaded in error or that are no longer needed.

Under File Types, use the following categories:
- Article file - Article file only. Turn off Track Changes for clean copy file
- Figures
- Redlined Article File
- Responses to reviewers - Supporting Material (this will not be part of the merged PDF sample, but it will be viewed by the reviewers.)

For help preparing figures, go to:
http://www.the-aps.org/publications/i4a/figures/index.htm

Use the FILE ORDER tab to verify the order of your files.

Notes about files:
The References must be listed in alphabetical order. Neglecting to put references in alphabetical order will delay the publication of your paper. An easy way to ensure the proper format of the References is to use a citation management software of your choice, available at: http://www.the-aps.org/publications/journals/styles.htm

Reviewers like to see a text file with changes tracked. You can upload that file as "Red Lined Version." It will not appear in the merged PDF to be reviewed but it will be available to reviewers as a supplemental file.

The title and author information will be prepopulated. If you wish to make changes to this information, click the heading labeled MANUSCRIPT INFORMATION.

Click the heading labeled VALIDATE to validate all your files. If you get an error, click the link that reads "Fix."

Hit the SUBMIT button to make the record go immediately to the associate editor assigned to the manuscript.

In case there are changes in authorship (addition, deletion, or change of order), a Change of Authorship Form must be obtained from:
http://www.the- aps.org/publications/journals/pub_author_form.PDF
and all the authors must sign it. The signed form should then be faxed to (301) 634-7243.

Regards,

David

David Linden
Journal of Neurophysiology
Editor-in-Chief

PLEASE NOTE: AS OF 1 JANUARY 2007, ALL ACCEPTED MANUSCRIPTS (EXCLUDING LETTERS TO THE EDITOR) WILL AUTOMATICALLY BE PUBLISHED ONLINE IN ARTICLES IN PRESS SHORTLY AFTER ACCEPTANCE.

Reviewer #1 (Comments to the Author (Required)):

This submission by Vogelstein and colleagues presents techniques for extracting spiking information from calcium imaging data. A particularly noteworthy feature is the attention that the authors have placed on developing computationally tractable algorithms. The submission is timely, as calcium imaging of large populations is becoming more common. The exposition and illustrations are clear and very detailed, allowing others to readily apply these methods. While there are now several solutions to this problem, I suspect that the methods presented here will become the preferred solution (or the major foundation for even newer developments) for many labs.

The manuscript is already very good, and I have no major issues with it. Three "medium" issues:
1. In the paragraph following Eq. 10b, the discussion of optimizing these in the following paragraph considers only an enumeration approach with constraints on integer n_t; the difficulties of this approach are used to motivate the exponential approximation. However, this ignores the possibility that Eq. 10b could be expanded by Stirling's formula, and (as with the exponential approximation) allow continuous n_t. Naturally, the argument

of the exponential would no longer be quadratic, and would contain log terms. However, the authors' barrier method also introduces log terms into the penalty function, and so it is not obvious that this would be any more challenging numerically. Indeed, with Stirling's approximation it would appear that the barrier method would not be necessary. The authors may want to explore, or at least discuss, this possibility. (If they are particularly interested in the case where the number of spikes in each frame is low, it is possible that the exponential approximation is better.)

2. The estimation of beta seems to be perhaps the most important parameter. The nonnegativity, clearly a useful constraint, becomes less powerful if beta is estimated as being too low, because the algorithm would then assign a "dense" spike train (i.e., with multiple spikes/significant probability of spiking in each frame). Fig 7 contains perhaps the example that comes closest to realizing this outcome. I'm curious to know why this isn't found in other cases, given that the authors have not imposed any form of sparseness constraint that I can detect. Is this a consequence of the exponential approximation, which is more accurate for spike rates that are low? Or is there some other reason that the examples produce sparse spiking?

3. It seems there is no figure showing real experimental data analyzed with the overlapping spatial filter? Is this because the authors lack a data set with "ground truth" spiking for this case? That would be reasonable grounds for not showing this type of analysis, but they should be aware that people who apply these methods will surely want to run their code on this type of data set. Some comments about the realities of running this algorithm, and comparison with the recent work of Mukamel et al (ref. 30), could be helpful.

Minor points:
First sentence of the abstract: "are gaining" -> "is gaining"

p.2, end of 1st paragraph in "Data driven generative model": main-text description of histogram in Fig. 1 has the solid/dashed reversed from what is shown in the figure.

Some of the figure captions could be more clearly marked as simulations. It could be helpful to have a convention of starting all captions with either "(Simulation)" or "(Experimental data)."

Reviewer #2 (Comments to the Author (Required)):

Calcium imaging provides an increasingly popular method for measuring neural activity in large populations of neurons both in brain slices and in vivo. While calcium signals largely reflect firing of action potentials (AP) in neurons, estimation of the underlying firing pattern from the calcium indicator fluorescence recordings alone remains challenging. Recently, a number of groups, including the authors of this manuscript, have presented various algorithms for the reconstruction of APs from calcium imaging. The current manuscript presents yet another approach for achieving this goal.
In line with previous work (Vogelstein et al., Biophys. J. 97(2):636), the authors pursue an optimization strategy to estimate the most likely firing pattern giving rise to an observed calcium signal. Unlike previous approaches, the authors here present an analytical solution to the problem, which they achieve by approximating the 'true' Poisson distribution underlying spiking events with an exponential distribution. Since the presented algorithm relies on analytical rather than numerical methods, it is guaranteed to converge on the optimal solution very quickly. This is a novel and original development, which is also attractive because of the possibility to implement it in a computationally efficient way.

While the authors provide a detailed, thorough and understandable description of the novel algorithm in the Methods section, their evaluation of the algorithm using simulated and measured data is insufficient. Specifically, a quantitative analysis of the performance of the algorithm is completely lacking. It is therefore impossible to judge - based on the results presented here - in how far and in what particular aspects the new algorithm represents an improvement compared to previous methods.

Specific comments:
Methods
1. The authors should comment on the error expected from approximating a Poisson distribution with an

exponential distribution.

Results
2. The main results are not well enough supported by a quantitative analysis of the inference algorithm. Statements like "very efficiently" (pg. 9, line 5 from bottom), "clearly .. more closely resembles" (pg. 9, line 3 from bottom), "poor inference quality" (pg. 10, line 10), "it is clear that .. it outperforms..) (pg. 10, line 24) or even "significantly more accurate" (Fig. 2 legend) need to be substantiated by a quantitative evaluation and a statistical comparison of the algorithm's performances. Such an evaluation is entirely missing in the manuscript.

3. A key evaluation criterion is the difference between the estimated and the real spike trains (in terms of number of spikes and of spike time accuracy). For example, in Figs. 2 and 3 on artificial data it is not obvious at all, whether the fast filter or the Wiener filter performs better, since both seem to predict spike occurrences well within the temporal resolution provided (taking the peaks of the Wiener filter). It also remains completely unexplored how the performance of the inference algorithms depends on critical parameters such as the noise level, the amplitude of single-spike evoked calcium transients, and the data sampling rate. A systematic examination of these dependencies would be very helpful.

4. Likewise the analysis of the performance of the fast filter on experimental data is lacking a convincing quantitative analysis: In Figs. 4 and 6 false positives are apparent in the fast filter results, but they are neither mentioned nor further evaluated. There is also no mentioning of how the background 'spiking' level of the filter results in Figs. 7, 9, and 10 is dealt with and what its origin is. Obviously, a simple thresholding procedure would fail to eliminate all false positives (or negatives). A statistical analysis has to be performed to substantiate the claim that 'fast filter significantly outperforms the Wiener filter' (Fig. 4 legend). Most importantly, however, there is no direct comparison of the estimated and the real spike times. Many of the traces, in particular those for bursts of spikes, would need to be shown expanded time scale in order to be able to visually inspect the quality of spike inference. It is not even clear how often spike are correctly assigned
to the correct acquisition frame and it is also impossible to see how many spikes are contained within the bursts in Figs. 4,5, and 7. The authors should display zoomed figure panels with larger time intervals.

5. Essential information is missing in the results section. What was the speed of fluorescence recordings? What was the range of baseline noise levels, which apparently varied between experiments and/or cells? How large were noise levels in comparison to single spike-evoked transients? How reliable could single spikes be inferred? What were the filter parameters in Figs. 4ff? What were the initial settings of algorithm parameters in Fig. 3 and how did the estimated parameter values compare with the true ones? What were the algorithm parameter settings in the applications to the experimental data? Since sometimes the recording durations used for parameter estimation and display were different (Fig. 5), these important time window as well as all results for estimated parameters need to be given for all Figure data.

6. Instead of providing a quantitative evaluation of the algorithm's performance the authors elaborate in section 3.2 on several extensions, for which no data are shown and which anyway do not seem to give any improvements. These additions could be just mentioned in the Discussion.

7. Section 3.3.4. Again it is NOT obvious that "the SMC filter outperforms the fast filter on biological data" (pg. 13, line 7 from bottom, referring to Fig. 7). This statement needs to be based on a thorough quantification.

Discussion
8. The authors claim that their approach works on "all the in vivo and in vitro data analyzed", yet no in vivo results are presented in the manuscript.

9. Overall, the discussion is relatively short and the comparison with other existing methods could be extended. The possible extensions provided in section 3.3 could be rather discussed in the Discussion.

Minor points
10. Units of fluorescence traces should be stated as relative percentage fluorescence changes DF/F, not normalized between 0 and 1 (or at least both versions should be given). DF/F is the standard way to present such calcium indicator data and allows a comparison with published work.

11. Does the filter also work in non-sparse firing regimes? Up to what firing frequency is reliable AP

reconstruction possible?

:

---

joshua vogelstein <joshuav@jhu.edu> **Mon, Jan 4, 2010 at 11:20 AM**
To: Liam Paninski <liam@stat.columbia.edu>

pretty good news. i'll send another email in a sec with my thoughts with regard to each specific comment....
[Quoted text hidden]

--
If it makes you feel better, please remember to consider humanity before doing stuff. Otherwise, please just have a nice day.
http://jovo.joshyv.me

Sent from Miami, Florida, United States

---

liam@stat.columbia.edu <liam@stat.columbia.edu> **Mon, Jan 4, 2010 at 11:32 AM**
To: joshua vogelstein <joshuav@jhu.edu>

nice. i think most of these comments are fair and on-point. you want to send me (and maybe tim) a to-do list? (speaking of to-do lists, any progress on the slow rise time or new spatial model?)

happy new year -
L
[Quoted text hidden]

---

joshua vogelstein <joshuav@jhu.edu> **Tue, Jan 5, 2010 at 8:07 PM**
To: Liam Paninski <liam@stat.columbia.edu>, "tim.machado" <tim.machado@gmail.com>

1. In the paragraph following Eq. 10b, the discussion of optimizing these in the following paragraph considers only an enumeration approach with constraints on integer n_t; the difficulties of this approach are used to motivate the exponential approximation. However, this ignores the possibility that Eq. 10b could be expanded by Stirling's formula, and (as with the exponential approximation) allow continuous n_t. Naturally, the argument of the exponential would no longer be quadratic, and would contain log terms. However, the authors' barrier method also introduces log terms into the penalty function, and so it is not obvious that this would be any more challenging numerically. Indeed, with Stirling's approximation it would appear that the barrier method would not be necessary. The authors may want to explore, or at least discuss, this possibility. (If they are particularly interested in the case where the number of spikes in each frame is low, it is possible that the exponential approximation is better.)

This is a cool suggestion. I'll mention it in the discussion, as another possible analytic approximation. Using Stirling's approximation does not clearly admit a distribution, which breaks the generative model aspect of the exponential approximation. Also, Stirling's approximation does not make sense when n=0, 1, 2, as it yields negative numbers. Thus, the barrier would still be necessary (more so), as it imposes the non-negativity constraint.

2. The estimation of beta seems to be perhaps the most important parameter. The nonnegativity, clearly a useful constraint, becomes less powerful if beta is estimated as being too low, because the algorithm would then assign a "dense" spike train (i.e., with multiple spikes/significant probability of spiking in each frame). Fig 7 contains perhaps the example that comes closest to realizing this outcome. I'm curious to know why this isn't found in other cases, given that the authors have not imposed any form of sparseness constraint that I can detect. Is this a consequence of the exponential approximation, which is more accurate for spike rates that are low? Or is there some other reason that the examples produce sparse spiking?

The exponential prior acts a sparsifying regularizer. I'll elaborate on this. We really don't mention this much at all.

3. It seems there is no figure showing real experimental data analyzed with the overlapping spatial filter? Is this because the authors lack a data set with "ground truth" spiking for this case? That would be reasonable grounds for not showing this type of analysis, but they should be aware that people who apply these methods will surely want to run their code on this type of data set.

Tim, you wanna apply the spatial stuff to our 12 in vitro trials? If not, i'll just add a note in the discussion that testing this in vitro is in progress.

Some comments about the realities of running this algorithm, and comparison with the recent work of Mukamel et al (ref. 30), could be helpful.

I'll mention that Mukamel might be useful to initialize the multineuron spatial stuff.

Minor points:
First sentence of the abstract: "are gaining" -> "is gaining"

i'll fix.

p.2, end of 1st paragraph in "Data driven generative model": main-text description of histogram in Fig. 1 has the solid/dashed reversed from what is shown in the figure.

i'll fix.

Some of the figure captions could be more clearly marked as simulations. It could be helpful to have a convention of starting all captions with either "(Simulation)" or "(Experimental data)."

i'll put simulated or in vitro in the beginning of each sentence (it is in each sentence already).

Reviewer #2 (Comments to the Author (Required)):

While the authors provide a detailed, thorough and understandable description of the novel algorithm in the Methods section, their evaluation of the algorithm using simulated and measured data is insufficient. Specifically, a quantitative analysis of the performance of the algorithm is completely lacking. It is therefore impossible to judge - based on the results presented here - in how far and in what particular aspects the new algorithm represents an improvement compared to previous methods.

yeah, i expected this comment. i made a figure a while back comparing the fast filter with the wiener filter as a function of sigma and lambda*dt. the problem with the fig was that it was not clear what metric to use. clearly, d(nhat,n), where d(.) is some metric. MSE & AUC were both ok, but neither was robust to time-bin shifts. any ideas?

Specific comments:

Methods
1. The authors should comment on the error expected from approximating a Poisson distribution with an exponential distribution.

In the paragraph before Eq (12), I'll elaborate on this approximation: when it is good, that it imposes sparsity (unlike Gaussian), admits a generative model (unlike Stirling's)

Results
2. The main results are not well enough supported by a quantitative analysis of the inference algorithm. Statements like "very efficiently" (pg. 9, line 5 from bottom), "clearly .. more closely resembles" (pg. 9, line 3 from bottom), "poor inference quality" (pg. 10, line 10), "it is clear that .. it outperforms..) (pg. 10, line 24) or even "significantly more accurate" (Fig. 2 legend) need to be substantiated by a quantitative evaluation and a statistical comparison of the algorithm's performances. Such an evaluation is entirely missing in the manuscript.

After Fig 3, I'll add a fig showing d(nhat,n) for fast and wiener as a function of sigma and lamda*dt.

3. A key evaluation criterion is the difference between the estimated and the real spike trains (in terms of number of spikes and of spike time accuracy). For example, in Figs. 2 and 3 on artificial data it is not obvious at all, whether the fast filter or the Wiener filter performs better, since both seem to predict spike occurrences well within the temporal resolution provided (taking the peaks of the Wiener filter). It also remains completely unexplored how the performance of the inference algorithms depends on critical parameters such as the noise level, the amplitude of single-spike evoked calcium transients, and the data sampling rate. A systematic examination of these dependencies would be very helpful.

A bit of discussion around Figs 2&3 could explain that one could threshold to take the peaks of the Wiener filter, but there would be no way of determining how many spikes were in each burst. Further, it would be unclear how to choose this threshold (ie, as a function of the percentiles, what about individual bumps from the wiener filter). so, while possible that post-processing the wiener filter could improve the results, it is not clear how, and not necessary for the fast filter. I'll add thresholded Wiener filter into the quantification fig, and see whether it is worth adding (last time i checked, it was not).

4. Likewise the analysis of the performance of the fast filter on experimental data is lacking a convincing quantitative analysis: In Figs. 4 and 6 false positives are apparent in the fast filter results, but they are neither mentioned nor further evaluated.

i'll mention them.

There is also no mentioning of how the background 'spiking' level of the filter results in Figs. 7, 9, and 10 is dealt with and what its origin is.

not sure what he means here, but i'll indicate in the captions that spikes are evoked.

Obviously, a simple thresholding procedure would fail to eliminate all false positives (or negatives). A statistical analysis has to be performed to substantiate the claim that 'fast filter significantly outperforms the Wiener filter' (Fig. 4 legend). Most importantly, however, there is no direct comparison of the estimated and the real spike times.

the d(nhat,n) fig will have errorbars, showing that improvement is significant.

Many of the traces, in particular those for bursts of spikes, would need to be shown expanded time scale in order to be able to visually inspect the quality of spike inference. It is not even clear how often spike are correctly assigned to the correct acquisition frame and it is also impossible to see how many spikes are contained within the bursts in Figs. 4,5, and 7. The authors should display zoomed figure panels with larger time intervals.

probably worthwhile for figs 4, 5 and 7. i thought about this before, but never did it. i'll try making them and see how it goes

5. Essential information is missing in the results section. What was the speed of fluorescence recordings?

I'll add frate rate to methods.

What was the range of baseline noise levels, which apparently varied between experiments and/or cells? How large were noise levels in comparison to single spike-evoked transients? How reliable could single spikes be inferred? What were the filter parameters in Figs. 4ff?

I'll add estimated parameters for Fig 4, 5 and 7, in the fig caption.

What were the initial settings of algorithm parameters in Fig. 3 and how did the estimated parameter values compare with the true ones?

I'll add initial params and true params in Fig 3 caption

What were the algorithm parameter settings in the applications to the experimental data? Since sometimes the recording durations used for parameter estimation and display were different (Fig. 5), these important time window as well as all results for estimated parameters need to be given for all Figure data.

I'll add duration and frame rate to all fig captions.

6. Instead of providing a quantitative evaluation of the algorithm's performance the authors elaborate in section 3.2 on several extensions, for which no data are shown and which anyway do not seem to give any improvements. These additions could be just mentioned in the Discussion.

I'll move section 3.3.1 -- 3.3.3 to discussion, and reword as appropriate.

7. Section 3.3.4. Again it is NOT obvious that "the SMC filter outperforms the fast filter on biological data" (pg. 13, line 7 from bottom, referring to Fig. 7). This statement needs to be based on a thorough quantification.

I'll quantify and report. Tim - perhaps you want quantify results for all 12 data sets? Or, better, we could quantify results on some 2P data, for which smc-filter tends to outperform the fast-filter?

Discussion
8. The authors claim that their approach works on "all the in vivo and in vitro data analyzed", yet no in vivo results are presented in the manuscript.

I'll add "(data not shown)" after "in vivo"

9. Overall, the discussion is relatively short and the comparison with other existing methods could be extended. The possible extensions provided in section 3.3 could be rather discussed in the Discussion.

i will move.

Minor points
10. Units of fluorescence traces should be stated as relative percentage fluorescence changes DF/F, not normalized between 0 and 1 (or at least both versions should be given). DF/F is the standard way to present such calcium indicator data and allows a comparison with published work.

I don't really agree here, since dF/F contains no more information than 0,1. Nonetheless, i'll acquiece, unless it looks really ugly, in which case i'll just add in the caption the range of dF/F

11. Does the filter also work in non-sparse firing regimes? Up to what firing frequency is reliable AP reconstruction possible?

this will be address in the fig in which we vary lambda*dt

****

So, i'll get to work on most of this stuff.

Tim - let me know how you feel about the stuff i laid out for you

Tim & Liam - i think it is worth changing from F_t = a*(C_t + b) to F_t = a*C_t + b, as i think it is more clear, and makes more sense for the vector observation case. also, the algorithms for inferring {a,b} in the latter case are more standard.

Tim & Liam - let me know if you have any further questions/complaints. If not, i'll probably finish this stuff sometime this week...

cheers,
j

ps - happy new year

[Quoted text hidden]

---

liam@stat.columbia.edu <liam@stat.columbia.edu> **Tue, Jan 5, 2010 at 9:14 PM**
To: joshua vogelstein <joshuav@jhu.edu>
Cc: "tim.machado" <tim.machado@gmail.com>

sounds all good, thanks.

i think it's definitely worth adding some real spatial data if we have it.

also worth adding the slow-rise time extension.

spike train metrics - how about convolve spike train with a smoothing (eg gaussian) filter, then compute L2 metric?

stirling's approx - this is going to come down to a gaussian approximation, basically, i think. you want to write this out a bit and see if it goes anywhere interesting?

i might drop the mention of the in vivo stuff. i agree with the reviewer, it is a bit odd to mention it and not show any real results.

as usual, send updated figs over first, then we can worry about the text later.
L

[Quoted text hidden]

---

joshua vogelstein <joshuav@jhu.edu> **Tue, Jan 5, 2010 at 9:18 PM**
To: liam@stat.columbia.edu

Cc: "tim.machado" <tim.machado@gmail.com>

ok on all accounts. give me a few days on the figs...

[Quoted text hidden]

Sent from Baltimore, MD, United States
